Welcome Messages

On behalf of the ISPACS steering committee, I would like to cordially welcome you to ISPACS 2019 held during Dec. 3-6 at Beitou Hot Spring Resort, Taipei, Taiwan.

The ISPACS is a premium international forum for leading researchers especially from Asia-Pacific basin in the highly active fields of theory, design and implementation of signal processing and communication systems. The ISPACS has a long history. The ISPACS has been held in nine countries including USA. The first conference of ISPACS was held in Taipei, Taiwan in 1992. After 27 years, it is the third time for this significant event to be held in Taipei. As always, ISPACS aims to serve as a platform for all professionals to discuss the latest findings and state-of-the-art technologies, create opportunities for young scholars to participate in international academic activities, as well as stimulate commercial activities and industrial development. In addition, ISPACS 2019 provides a nice chance to participants from all around the world to visit the beautiful city, Taipei.

I would like to show my sincere appreciations to all the organizing committee members led by Prof. Jing-Ming Guo for their great contributions to ISPACS 2019. Welcome all participants to ISPACS 2019!

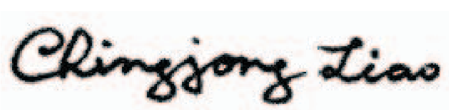
Akira Taguchi
Chair, International Steering Committee of ISPACS
Tokyo City University, Japan

As the president of National Taiwan University of Science and Technology, I have the great honor to welcome leading experts, professors, friends and colleagues gathering here. It is our pleasure to host the event of the 2019 International Symposium on Intelligent Signal Processing and Communication Systems. I would like to thank all of you who work so hard as the educator of technology pioneer that stand at the front line of intelligent signal processing and communication systems issues. Your generous support has made significant contribution to Taiwan and the world.

Taiwan Tech was formerly known as the National Taiwan Institute of Technology, the first higher education institution of its kind within our nation's technical and vocational education system, seeking to develop highly trained engineers and managers. Over the last 40-plus years since its foundation, NTUST has focused on developing academic research, teaching, and learning services with an innovative spirit. The Department of Electrical Engineering was founded in 1978 with the mission of providing the society with high-quality education pertinent to the electrical engineering discipline, in response to the rapid growth of science and technology industries. To pursue excellence in research, teaching, and service in the area of electrical engineering, graduate programs were established in 1979 (MS) and 1982 (PhD). By incorporating a wide selection of advanced courses and opportunity of conducting independent research, students were trained to possess in-depth frontier knowledge, high technical skills, and planning ability that are vital to their future careers in industry or academia. It is much honored to have our Department of Electrical Engineering to make such great contribution to Taiwan, as well as to ISPACS 2019.

Again, I would also like to express my sincere gratitude to all of your fabulous dedication. Wish you can enjoy the conference and Taipei in the following days.

Best Regards,



Ching-Jong Liao

Honorary Chair of ISPACS 2019

President, National Taiwan University of Science and Technology, Taiwan

I would like to express a warm welcome to all of you joining the 2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2019) held in Beitou, Taipei, Taiwan on December 3-6, 2019.

Since 1992, the ISPACS has become one of the major symposia in the signal processing and communication systems field. ISPACS provides excellent opportunities for researchers from academia and industries all around the world to report and discuss the latest applications on signal processing and communication systems technologies and innovations.

In view of the current technologies, artificial intelligence has made its way into many areas and become indispensable. To improve human welfare, we set up the main theme of ISPACS 2019 as "*Impact of Artificial Intelligence: From Reality to Imagination, from Technologies to Applications*". Recently, the development and application of artificial intelligence has grown rapidly. Thus, ISPACS 2019 aims to keep on exploring and exchanging up-to-date techniques and findings, integrate and commercialize this research so as to bring the results into full play.

On behalf of the organizing committee, I would like to express the appreciation for your endeavors and contributions to this symposium to make it meaningful. On top of that, Beitou is a beautiful city with interesting things to do and enchanting places to visit. It is particularly famous for hot spring. We believe you will have marvelous experiences and memories at ISPACS 2019!

Sincerely Yours,

Jing-Ming Guo

General Chair of ISPACS 2019

Distinguished Professor, Department of Electrical Engineering,

National Taiwan University of Science and Technology, Taiwan

On behalf of the Technical Program Committee of the 2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2019), we would like to take this opportunity to appreciate all of your participation in ISPACS 2019. We also want to express our gratitude to the Organizing Committee, especially the General Chair Prof. Jing-Ming Guo. This symposium would not have been possible without their guidance and effort.

This year, ISPACS 2019 attracted 295 submissions from 15 countries, from which 191 high-quality papers were accepted through a conscientious and careful review. Among 191 accepted papers, 161 papers are selected for oral presentation, and 30 papers for poster presentation.

During this four-day conference, we are going to experience different discussion sections and events, including 4 keynote speeches, 4 invited speeches, and various paper presentations through 28 lecture sessions and 2 poster sessions.

We hope that all participants will benefit from this symposium and gain new ideas for future research, while keeping abreast of the current development in the field of signal processing and communication systems.

Please enjoy the technical program and have a great time here in Beitou, Taipei!

Best Regards,

Prof. Chih-Hsien Hsia
Technical Program Chair of ISPACS 2019
National Ilan University, Taiwan

Prof. KokSheik Wong
Technical Program Chair of ISPACS 2019
Monash University Malaysia, Malaysia

Organizing Committee

Honorary Chair

President Ching-Jong Liao

National Taiwan University of Science and Technology, Taiwan

General Chair

Prof. Jing-Ming Guo

National Taiwan University of Science and Technology, Taiwan

General Co-Chairs

Prof. Jen-Shiun Chiang

Tamkang University, Taiwan

Prof. Kosin Chamnongthai

King Mongkut's University of Technology Thonburi, Thailand

Prof. Jar-Ferr Yang

National Cheng Kung University, Taiwan

Technical Program Chairs

Prof. Chih-Hsien Hsia

National Ilan University, Taiwan

Prof. KokSheik Wong

Monash University Malaysia, Malaysia

Technical Program Co-Chairs

Prof. Shang-Hong Lai

National Tsing Hua University, Taiwan

Prof. Mei-Juan Chen

National Dong Hwa University, Taiwan

Prof. Huanqiang Zeng

Huaqiao University, China

Prof. Wen-Huang Cheng

National Chiao Tung University, Taiwan

Prof. Chia-Yu Yao

National Taiwan University of Science and Technology, Taiwan

Advisory Board Liaison

Prof. Akira Taguchi

Tokyo City University, Japan

Special Session Chairs

Prof. Jason C. Hung

National Taichung University of Science and Technology, Taiwan

Prof. Hsuan-Ming Feng

National Quemoy University, Taiwan

Prof. Lin Zhang

Tongji University, China

Prof. Miin-Jong Hao

National Kaohsiung First University of Science and Technology, Taiwan

Prof. Hsiao-Chin Chen

National Taiwan University of Science and Technology, Taiwan

Award Chairs

Prof. Kai-Kuang Ma

Nanyang Technological University, Singapore

Prof. Kaoru Arakawa

Meiji University, Japan

Prof. Hitoshi Kiya

Tokyo Metropolitan University, Japan

Prof. Tihao Chiang

Ambarella Taiwan Ltd., Taiwan

Prof. Trio Adiono

Institut Teknologi Bandung, Indonesia

Registration Chairs

Prof. Duan-Yu Chen

Yuan Ze University, Taiwan

Prof. Ting-Lan Lin

National Taipei University of Technology, Taiwan

Publication Chairs

Prof. Jian-Jiun Ding

National Taiwan University, Taiwan

Prof. Heri Prasetyo

Universitas Sebelas Maret, Indonesia

Finance Chair

Prof. Tien-Ying Kuo

National Taipei University of Technology, Taiwan

Exhibition Chairs

Prof. Chi-Chia Sun

National Formosa University, Taiwan

Prof. Chih-Yuan Yao

National Taiwan University of Science and Technology, Taiwan

Prof. Yen-Lin Chen

National Taipei University of Technology, Taiwan

Local Arrangement Chairs

Prof. Chung-An Shen

National Taiwan University of Science and Technology, Taiwan

Prof. Kai-Lung Hua

National Taiwan University of Science and Technology, Taiwan

Web Chair

Mr. Sankar Srinivasan

National Taiwan University of Science and Technology, Taiwan

Technical Program Committee

Hao San
Tokyo City University, Japan

Chi-Chia Sun
National Formosa University, Taiwan

Yi-Hsien Wang
Chinese Culture University, Taiwan

KokSheik Wong
Monash University Malaysia, Malaysia

Hung-Wei Wu
Kun Shan University, Taiwan

Chia-Yu Yao
National Taiwan University of Science and Technology,
Taiwan

Huanqiang Zeng
Huaqiao University, China

Lin Zhang
Tongji University, China

Reviewers

Della Fitrayani Budiono
Kosin Chamnongthai
Che-Wei Chang
Bo-Ruei Chen
Chien-Hung Chen
Ching-Yi Chen
Hao-Hui Chen
Hsiao-Chin Chen
Hsin-Liang Chen
Pei-Yin Chen
Wei-Yu Chen
Yi-Ling Chen
Yongyong Chen
Yuan-Ho Chen
Yung-Yao Chen
Siew Cheng Lai
Wai Khuen Cheng
Ling-An Chiang
Feng-Tsun Chien
Yih-Shyh Chiou
Hung-Wei Chiu
Shen Shou Max Chung
Jian-Jiun Ding
Yongzhao Du
Shih-Hau Fang
Tz-Heng Fu
Minoru Fukumi
Zeng-Wei Hong
Shintami Chusnul Hidayati
Chih-Hsien Hsia
Juichien Hsiao
Wen-Liang Hsue
Kai-Lung Hua
Jeng-Tze Huang
Chin-Ya Huang
Sin-Ye Jhong
Rafael Kaliski
Wen-Cheng Lai
Chao-Hsien Lee
Cheng-Shian Lin
Chun-Yueh Lin
Guo-Shiang Lin
Jheng-Han Lin
Szu-Yin Lin
Ting-Yu Lin
Yi-Ru Lin
Peizhong Liu
Chung-Hsin Liu
Ye Luo
Wei-Lung Mao
Narongrit Mekloi
Norwati Mustapha
Boon-Yaik Ooi
Kang-Chun Peng
Nicodimus Retdian
Hao San
Sankarasrinivasan Seshathiri
Iwan Setiawan
Shan-Hsiang Shen
Horng-Yuan Shih
Huang-Chia Shih
Joni Welman Simatupang
Kuo-Ho Su
Andreas Simon Suganda
Chi-Chia Sun
Yasuhiro Takahashi
Cheng-Yuan Tang
Kea-Tiong (Samuel) Tang
Chien-Cheng Tseng
Shih-Huan Tseng
Min-Haw Wang
Shyang-Yuh Wang
Yi-Hsien Wang
Pei Voon Wong
Wing Wong
Hung-Wei Wu
Jian-Ming Wu
Yong Xia
Sendren Sheng-Dong Xu
Ching-Hao Yang
Jiachen Yang
Wei-Bin Yang
Fu-Ju Yang
Chia-Yu Yao
Hsiang-Yuan Yeh
Daxin Zhu

Program at a Glance

Time / Date	Dec. 3 (Tue.)
15:00-18:00	Registration (Lobby, 2F)
18:30-20:00	Welcome Reception (Pearl Banquet Hall, 2F)

Time / Date	Dec. 4 (Wed.)				
Venue	Pearl Banquet Hall	The Sky Room	The East and West Room	Jade, Stone and Spring Room	Outside the Jade, Stone and Spring Room
08:00-08:45	Registration (Lobby, 2F)				
08:45-09:00	Opening Ceremony				
09:00-10:00	Keynote Speech I *Prof. Ioannis Pitas*				
10:00-10:20	Coffee Break				
10:20-12:10	Invited I / Oral 1 Invited Speech I / Multimedia and Systems I	Oral 2 Advanced Information Technique and Its Applications I	Oral 3 Signal Processing I	Oral 4 Applications of Deep Learning Based Pixel-Wise Classification Techniques	Poster Session I
12:10-13:30	Lunch				
13:30-14:30	Oral 5 Recent Advances in Autonomous Vehicles and Connected Cars I	Oral 6 Deep Learning and Its Applications	Oral 7 Communication Systems I	Oral 8 RF and Microwave Techniques for Advanced Communication Systems	
14:30-15:00	Coffee Break				
15:00-16:30	Invited II / Oral 9 Invited Speech II / Circuits and Systems	Oral 10 Advanced Information Technique and Its Applications II	Oral 11 Communication Systems II	Oral 12 Signal and Image Processing for Advanced Technology I	
16:30-18:00	Oral 13 Recent Advances in Autonomous Vehicles and Connected Cars II	Best Paper Award	Oral 14 Communication Systems III	Oral 15 Smart Sensors and Intelligent Video Analytics for AI and IoT Applications	

Time / Date	Dec. 5 (Thu.)				
Venue	Pearl Banquet Hall	The Sky Room	The East and West Room	Jade, Stone and Spring Room	Outside the Jade, Stone and Spring Room
08:00-09:00	Registration (Lobby, 2F)				
09:00-10:00	Keynote Speech II *Director* *Tihao Chiang*				
10:00-10:20	Coffee Break				
10:20-12:10	Invited III / Oral 16 Invited Speech III / Signal Processing II	Oral 17 Analog Circuits and Their Applications	Oral 18 Multimedia and Systems II	Oral 19 Analog ICs and Antenna Design for 5G Communication	Poster Session II
12:10-13:30	Lunch				
13:30-14:30	Keynote Speech III *Mr. Pat Hsu*				
14:30-15:00	Coffee Break				
15:00-16:30	Oral 20 AI and the Interdisciplinary Research	Oral 21 Communication Systems IV	Oral 22 Signal and Image Processing for Advanced Technology II	Oral 23 Intelligent Signal Processing Technique and Its Applications	
16:30-18:00	Oral 24 VLSI	Oral 25 Communication Systems V	Best Student Paper Award	Oral 26 Intelligent Visual Perception	
18:30-21:00	Banquet (Ballroom, 6F, Asia Pacific Hotel Beitou)				

Time / Date	Dec. 6 (Fri.)			
Venue	Pearl Banquet Hall	The Sky Room	The East and West Room	Jade, Stone and Spring Room
08:00-09:00	Registration (Lobby, 2F)			
09:00-10:00	Keynote Speech IV *Prof. Pierre Moulin*			
10:00-10:20	Coffee Break			
10:20-12:10	Invited IV / Oral 27 Invited Speech IV / Signal Processing III	Oral 28 Signal and Image Processing for Advanced Technology III		
12:10-13:30	Lunch (For tour participants only)			
13:30-18:00	City Tour			

Keynote Speech I

09:00-10:00 / Pearl Banquet Hall

Chair: Jen-Shiun Chiang, Taiwan

Deep Learning and Computer Vision for Multiple Drone Media Production

Prof. Ioannis Pitas, Greece

Invited Speech I / Oral 1-Multimedia and Systems I

10:20-12:10 / Pearl Banquet Hall

Chair: Jen-Shiun Chiang, Taiwan
Suchada Tantisatirapong, Thailand

Invited I

Automated Music Scoring System Based on Deep-Leaning Method for Japanese Traditional Instruments Tsugaru-Shamisen

Juichi Kosakaya, Japan

O1-1

Classification of Environmental Sounds Using Convolutional Neural Network with Bispectral Analysis

Katsumi Hirata, Takehito Kato and Ryuichi Oshima, Japan

O1-2

Restoration of Compressed Picture Based on Lightweight Convolutional Neural Network

Tien-Ying Kuo, Yu-Jen Wei and Chang-Hao Chao, Taiwan

O1-3

Automatic Damage Recovery of Old Photos Based on Convolutional Neural Network

Tien-Ying Kuo, Yu-Jen Wei, Ming-Jui Lee and Tzu-Hao Lin, Taiwan

O1-4

Wireless Video Transmission Over Multiuser MIMO Systems with Fair Power Allocation

Yuta Yoshikawa, Koji Tashiro, Masayuki Kurosaki and Hiroshi Ochi, Japan

O1-5

Comparative Study of Masking and Mapping Based on Hierarchical Extreme Learning Machine for Speech Enhancement

Ryandhimas Zezario, Join Sigalingging, Tassadaq Hussain, Indonesia; Jia-Ching Wang, Yu Tsao, Taiwan

O1-6

Fractal Dimension Based Color Texture Analysis for Mangosteen Ripeness Grading

Montri Phothisonothai and Suchada Tantisatirapong, Thailand

Oral 2- Advanced Information Technique and Its Applications I

10:20-12:10 / The Sky Room

Chair: Chih-Hsien Hsia, Taiwan
Yung-Yao Chen, Taiwan

O2-1

A Study on the Applying of Digital Information and Visually Creative Thinking for Environment Design Education

Chi-Jui Tsai, Shu-Hsuane Chang and Kuo-An Wang, Taiwan

O2-2

A New Variation of Singular Value Decomposition

Heri Prasetyo, Didi Rosiyadi and Iwan Setiawan, Indonesia

O2-3

Thermal-Based Pedestrian Detection Using Faster R-CNN and Region Decomposition Branch

Yung-Yao Chen, Sin-Ye Jhong, Guan-Yi Li and Ping-Han Chen, Taiwan

O2-4

H-BTC Database: A Brief Review on Halftone Based Block Truncation Coding (H-BTC) Images

Jing-Ming Guo and Sankarasrinivasan Seshathiri, Taiwan

O2-5

Function Block-Based Robust Firmware Update Technique for Additional Flash-Area/Energy-Consumption Overhead Reduction

Jisu Kwon and Daejin Park, South Korea

Scientific Program Dec. 4, 2019

Oral 3- Signal Processing I

10:20-12:10 / The East and West Room
Chair: Minoru Komatsu, Japan
Ting-Yu Lin, Taiwan

O3-1
Adaptive Step-Size Recursive Least Biphase Errors Algorithm
Shin'Ichi Koike, Japan

O3-2
Proposal of BSS Method to Separate the Respiratory Sound and the Heart Sound
Yuki Kubota, Minoru Komatsu and Hiroki Matsumoto, Japan

O3-3
Hammerstein Spline Adaptive Filtering Based on Normalised Least Mean Square Algorithm
Sethakarn Prongnuch, Suchada Sitjongsataporn and Theerayod Wiangtong, Thailand

O3-4
Large-Cell Wireless Train Radio Communications Employing Narrowband Multiple Single Carrier Modulation Schemes for High-Speed Railways
Takuya Takahashi and Hiroshi Kubo, Japan

O3-5
A Preprocessing for Sound Source Separation Using Complex Weighted Sum Circuits
Shun Nishimaki and Kenji Suyama, Japan

O3-6
Feedback Active Noise Control using Linear Prediction Filter for Colored Wide-Band Background Noise Environment
Riku Takasugi, Yosuke Sugiura and Tetsuya Shimamura, Japan

O3-7
Chord Label Estimation from Acoustic Signal Considering Difference in Electric Guitars
Nozomiko Yasui, Masanobu Miura and Tetsuya Shimamura, Japan

Oral 4- Applications of Deep Learning Based Pixel-Wise Classification Techniques

10:20-12:10 / Jade, Stone and Spring Room
Chair: Jian-Jiun Ding, Taiwan
Huanqiang Zeng, China

O4-1
A Benchmark for Homework Tidiness Assessment
Hanxiao Wu, Zhenyu Zhang, Zhichao Zheng, Fei Shen, Weiwei Zhang, Jianqing Zhu and Huanqiang Zeng, China

O4-2
Super-Resolution via Wavelet Transform and Advanced Learning Techniques
Yi-Wen Chen and Jian-Jiun Ding, Taiwan

O4-3
Learning Based SLIC Superpixel Generation and Image Segmentation
Chieh-Sheng Chang, Jian-Jiun Ding and Heng-Sheng Lin, Taiwan

O4-4
Automatic Chinese Handwriting Verification Algorithm Using Deep Neural Networks
Chi-Chang Lee and Jian-Jiun Ding ,Taiwan

O4-5
DDSnet: A Deep Document Segmentation with Hybrid Blocks Architecture Network
Jing-Ming Guo, Li-Ying Chang and Hao-Hsuan Lee, Taiwan

O4-6
Using Multiple Fully Convolutional Networks Fusion for Floor Area Detection of Mobile Robots
Cheng-Jian Lin, Yu-Chi Li and Chin-Ling Lee, Taiwan

O4-7
Learning Based Noise Identification Techniques Using Time-Frequency Analysis and the U-Net
Chih-Hao Wang, Jian-Jiun Ding, Chieh-Sheng Chang and Liang-Yu Ouyang, Taiwan

Oral 5- Recent Advances in Autonomous Vehicles and Connected Cars I

13:30-14:30 / Pearl Banquet Hall
Chair: Lau Phooi Yee, Malaysia
Sankar Srinivasan, Taiwan

O5-1
Autonomous Vehicle Trajectory Planning and Control Based on Traffic Motion Prediction
Yuho Song, Dongchan Kim and Kunsoo Huh, South Korea

O5-2
Implementation and Evaluation of CNN Based Traffic Sign Detection with Different Resolutions
Ying-Chi Chiu, Huei-Yung Lin and Wen-Lung Tai, Taiwan

O5-3
Traffic Light Detection Using Convolutional Neural Networks and Lidar Data
Tien-Wen Yeh, Ssu-Yun Lin and Huei-Yung Lin, Taiwan

O5-4
Real-Time Forward Collision Alert System Using Raspberry Pi
Wai Chun Phoon and Phooi Yee Lau, Malaysia

Oral 6- Deep Learning and Its Applications

13:30-14:30 / The Sky Room
Chair: Sin-Ye Jhong, Taiwan
Wei-Wen Hsu, Taiwan

O6-1
Bounding Box Based Annotation Generation for Semantic Segmentation by Boundary Detection
Xiaolong Xu, Fanman Meng, Hongliang Li, Qingbo Wu, Yuwei Yang and Shuai Chen, China

O6-2
Hardware Architecture of Emotion Recognition from Speech Features Using Recurrent Neural Network and Backpropagation through Time
Yashael Faith Arthanto, Joshua Gunawan, Teresia Savera Rosa Putri and Trio Adiono, Indonesia

O6-3
A Metropolis within Gibbs Algorithm for Estimating the Bayesian Reduced Reparameterized Unified Model
Mengta Chung, Taiwan

O6-4
A Design Framework for Hardware Approximation of Deep Neural Networks
Wei-Hung Lin, Hsu-Yu Kao and Shih-Hsu Huang, Taiwan

Oral 7- Communication Systems I

13:30-14:30 / The East and West Room
Chair: Hsin-Liang Chen, Taiwan
Mei-Juan Chen, Taiwan

O7-1
A Cooperation-Based Scheme for Secondary Users to Enhance the Utilization of TV Spectra
David Shiung, Ya-Yin Yang and Chi Ou-Yang, Taiwan

O7-2
Lattice-Superposition NOMA for Near-Far Users
Chi Wan Sung, Hong Kong; Kenneth Shum, China

O7-3
Compare of Vehicle Management over the Air and On-Board Diagnostics
Taehyoung Kim and Sungkwon Park, South Korea

O7-4
Investigation on Distributed Joint Optimization of User Association and ICIC Based on PF Criteria in Small Cell Deployments
Takuya Matsumoto and Nobuhiko Miki, Japan

Oral 8- RF and Microwave Techniques for Advanced Communication Systems

13:30-14:30 / Jade, Stone and Spring Room
Chair: Hao-Hui Chen, Taiwan
Hung-Wei Wu, Taiwan

O8-1
Bended Differential Stripline Using Timing-Offset Differential Signal
Chih-Chen Yeh, Wen-Ju Chen and Chun-Long Wang, Taiwan

O8-2
A Compact Triple Passband Bandpass Filter Using Stub Load-Uniform impedance Resonators
Kuan-Jen Lin, Zong-Fu Li, Tzu-Chun Tai, Hung-Wei Wu and Yeong-Her Wang, Taiwan

O8-3
A New Transparent Planar Reflector Antennas for Satellite DTV Applications
Yu-Ming Lin, Zong-Fu Li, Tzu-Chun Tai, Hung-Wei Wu, Shoou-Jinn Chang and Yeong-Her Wang, Taiwan

Invited Speech II / Oral 9- Circuits and Systems

15:00-16:30 / Pearl Banquet Hall
Chair: Ching-Yuan Yang, Taiwan
Jian-Chiun Liou, Taiwan

Invited II
Formulas for Defining Dual-Typhoon Interactions and Eco-Environmental Vulnerability Assessment Frameworks with Earth Observations & GIS
Yuei-An Liou, Taiwan

O9-1
Overflow-Free Realizations for LTI Digital Filters
Shuichi Ohno and Shenjian Wang, Japan

O9-2
Proposal of DTW Distance Calculator Using Neuron CMOS Inverter
Takanori Kurano, Masaaki Fukuhara, Kaiki Yamada, Ryo Yamamoto, Hayato Yagi and Nao Onji, Japan

O9-3
A Reduction of Current Consumption of a Hamming Distance Detector by Improvement of Current Mirror Circuit
Nao Onji, Takanori Kurano, Toyoki Saho, Riku Ohtsuka, Kaiki Yamada and Masaaki Fukuhara, Japan

O9-4
Study and Explore on the Energy Harvesting of the Solar Cell with DC/DC Converter PWM System
Jian-Chiun Liou and Zhen-Xi Chen, Taiwan

O9-5
Inkjet Technology Addressing and Precise Control of DNA Liquid
Jian-Chiun Liou and Zhen-Xi Chen, Taiwan

Oral 10- Advanced Information Technique and Its Applications II

15:00-16:30 / The Sky Room
Chair: Szu-Yin Lin, Taiwan
Heri Prasetyo, Indonesia

O10-1
A Semi-Supervised Learning Model Based on Convolutional Autoencoder and Convolutional Neural Network for Image Classification
Yu-Xuan Li and Hsiang-Yuan Yeh, Taiwan

O10-2
Reversible Data Embedding for Fingerprint Minutiae Templates Authentication
Yenlung Lai, Zhe Jin, Koksheik Wong and Ming Jie Lee, Malaysia

O10-3
Practice of a Two-Stage Model Using Support Vector Regression and Black-Litterman for ETF Portfolio Selection
Jung-Bin Li and Chuan-Yin Chen, Taiwan

O10-4
A Continuous Facial Expression Recognition Model Based on Deep Learning Method
Szu-Yin Lin, Yi-Wen Tseng, Chang-Rong Wu, Yun-Ching Kung, Yi-Zhen Chen and Chao-Ming Wu, Taiwan

O10-5
Mixing Binary Face and Fingerprint Based on Extended Feature Vector (EFV) Hashing
Ming Jie Lee, Zhe Jin, Malaysia; Minyi Li, Australia; Daniel Bo-Wei Chen, Taiwan

O10-6
Analysis of Chroma Pixel Value Prediction Using Luma Pixel Values
Yan-Jhu Wang, You-Sheng Guo, Taiwan; Chang-Sheng Deng, China; Ting-Lan Lin, Taiwan

Oral 11- Communication Systems II

15:00-16:30 / The East and West Room
Chair: Hideyuki Torii, Japan
Hsin-Liang Chen, Taiwan

O11-1
Mapping and Permutation Set Design for Spatial Permutation Modulation (SPM)
Che-Wei Lee, Hsu-Hsuan Tu and I-Wei Lai, Taiwan

O11-2
Hybrid Multiple Access Using Simultaneously NOMA and OMA
Hirofumi Suganuma, Hiroaki Suenaga and Fumiaki Maehara, Japan

O11-3
Low-Complexity Maximum Likelihood (ML) Decoder for Space-Time Block Coded Spatial Permutation Modulation (STBC-SPM)
Hsu-Hsuan Tu, Che-Wei Lee and I-Wei Lai, Taiwan

O11-4
Comparison of Inter-Cell and Co-Channel Interference Power between A-ZCZ and GMO-ZCZ Sequence Sets by Computer Simulation
Hideyuki Torii and Takahiro Matsumoto, Japan

O11-5
Single Carrier Block Transmission Schemes for Acoustic Communications and Their Field Evaluation Results
Ryuki Sano and Hiroshi Kubo, Japan

O11-6
A Differential Multiple Single Carrier Modulation Scheme for Underwater Acoustic Communications and Its Actual Evaluation Results
Rintaro Yoshii, Yoshiki Tsukamoto, Takuya Takahashi and Hiroshi Kubo, Japan

Oral 12- Signal and Image Processing for Advanced Technology I

15:00-16:30 / Jade, Stone and Spring Room
Chair: Akira Taguchi, Japan

O12-1
Image Quality Improvement of Underwater Images in Ideal HSI Color Space
Ibuki Yoshida and Akira Taguchi, Japan

O12-2
Generalized Differential Gray-Level Histogram Equalization
Hideaki Tanaka and Akira Taguchi, Japan

O12-3
Color Image Enhancement by Using Hue-Saturation Gradient
Hiromu Endo and Akira Taguchi, Japan

O12-4
Inverse Halftoning Using Autoencoder
Jing-Ming Guo and Shyang-Yih Wang, Taiwan

O12-5
Enhanced Image Quality Assessment Using the Ordered Dithering Technique
Jing-Ming Guo and Sankarasrinivasan Seshathiri, Taiwan

O12-6
Road Boundary Detection for Straight Lane Lines Using Automatic Inverse Perspective Mapping
Fabien Rakotondrajao, France; Kharittha Jangsamsi, Thailand

Oral 13- Recent Advances in Autonomous Vehicles and Connected Cars II

16:30-18:00 / Pearl Banquet Hall
Chair: Kosin Chamnongthai, Thailand

O13-1
Semantic Traffic Light Understanding for Visually Impaired Pedestrian
Chanagan Pongseesai and Kosin Chamnongthai, Thailand

O13-2
Slice Image Segmentation Using Deep Learning for Lane Line Detection
Jing-Ming Guo and Herleeyandi Markoni, Taiwan

O13-3
Design and Implementation of Data Collection and Driving Behaviour Analysis Based on SAE J1939
B.V.P Prasad, Jing-Jou Tang and Sheng-Jhu Luo, Taiwan

O13-4
Multi-Agent Deep Reinforcement Learning for Cooperative Driving in Crowded Traffic Scenarios
Jongwon Park, Kyushik Min and Kunsoo Huh, South Korea

O13-5
New Asymmetric Data Transmission Method for In-Vehicle Network
Juho Lee and Sungkwon Park, South Korea

Best Paper Award

16:30-18:00 / The Sky Room

Chair: Pichaya Tandayya, Thailand
Heri Prasetyo, Indonesia

BS-01
Heterogeneous Industrial IoT Integration for Manufacturing Production
Chao-Hsien Lee, Zheng-Lin Wu, Yun-Ting Chiu and Yi-Shu Chen, Taiwan

BS-02
Integrated Simulator for Evaluating Cooperative Eco-Driving System
Geonil Lee and Jae-Il Jung, South Korea

BS-03
Pilot Allocation for Iteration Reduction of Channel Estimation Using MF-Based Interpolation
Norisato Suga, Ryohei Sasaki and Toshihiro Furukawa, Japan

BS-04
A 0.8V 14bit 62.5kSPS Non-Binary Cyclic ADC Using SOTB CMOS Technology
Shuichiro Yamada, Toshiki Ohtsu, Minami Sasaki, Hao San, Tatsuji Matsuura and Masao Hotta, Japan

BS-05
An Evolutionary Computation Approach for Approximate Computing of PNN Hardware Circuits
Ching-Yi Chen, Che-Wei Chang and Zih-Ching Chen, Taiwan

Oral 14- Communication Systems III

16:30-18:00 / The East and West Room

Chair: Miin-Jong Hao, Taiwan
Ting-Yu Lin, Taiwan

O14-1
Novel Subblock Partitioning for PTS Based PAPR Reduction of OFDM Signals
Miin-Jong Hao, Hong-Han Yao and Shou-Sheu Lin, Taiwan

O14-2
Famileaf: Flowerpot Robot for Dementia Prevention
Nagisa Ishizumi and Manabu Gouko, Japan

O14-3
An Enhanced Method for Turbo Code Interleavers
Shou-Sheu Lin, Sun-Ting Lin and Yu-Cheng Li, Taiwan

O14-4
Performance Analysis of Power Outage Probability for Drone based IoT Connectivity Network
Sarun Duangsuwan, Anna Chusongsang and Sathaporn Promwong, Thailand

O14-5
Channel Selection Metric by the Number of Users and SNR in WLAN
Taiga Okutate, Ikuo Oka and Shingo Ata, Japan

Oral 15- Smart Sensors and Intelligent Video Analytics for AI and IoT Applications

16:30-18:00 / Jade, Stone and Spring Room

Chair: Chi-Chia Sun, Taiwan
Yen-Lin Chen, Taiwan

O15-1
Robot Knows Where Human Are Through Sensory Data Fusion
Shih-Huan Tseng, Taiwan

O15-2
AIoT Solution Survey Machine Learning on Low-Cost Microcontroller
Hoang-The Pham, Minh-Anh Nguyen and Chi-Chia Sun, Taiwan

O15-3
Searching ROI for Object Detection Based on CNN
Chia-Lin Wu, Chih-Yang Lin, Phanuvich Hirunsirisombut, Taiwan; Hui-Fuang Ng, Malaysia; Timothy K. Shih, Taiwan

Keynote Speech II

09:00-10:00 / Pearl Banquet Hall
Chair: Kosin Chamnongthai, Thailand

Ultra HD Computer Vision Processor for Autonomous Driving Applications
Sr. Managing Director Tihao Chiang, Taiwan

Invited Speech III / Oral 16- Signal Processing II

10:20-12:10 / Pearl Banquet Hall
Chair: Miin-Jong Hao, Taiwan
Katsumi Hirata, Japan

Invited III
Development of a Diagnostic-Aid Platform for Epileptic Electroencephalogram
Toshihisa Tanaka, Japan

O16-1
Blind Equalizer with Noise Reduction Function
Mitsuru Mashimo, Minoru Komatsu and Hiroki Matsumoto, Japan

O16-2
Verification of Regression Analysis of Muscle Fatigue Using Wrist EMG
Misato Matsushita, Momoyo Ito, Shin-Ichi Ito and Minoru Fukumi, Japan

O16-3
Distant Sound Source Suppression Based on Multichannel Nonnegative Matrix Factorization with Bases Distance Maximization Penalty
Kazuma Takiguchi, Arata Kawamura and Youji Iiguni, Japan

O16-4
Alveolar Fricative Consonants Detection with Easily Interpretable Feature for Speech Training
Kouki Furuya, Arata Kawamura and Youji Iiguni, Japan

O16-5
Adaptive Line Enhancer-Based Beat Noise Suppression for FM Radio in Motor Vehicle
Kosuke Hasada, Arata Kawamura and Youji Iiguni, Japan

O16-6
Adaptive Direct Blind Equalization under Noisy Environment
Minoru Komatsu, Nari Tanabe and Toshihiro Furukawa, Japan

Oral 17- Analog Circuits and Their Applications

10:20-12:10 / The Sky Room
Chair: Hao San, Japan
Nicodimus Retdian, Japan

O17-1
Low-Power Bandgap Reference with Soft Startup for Energy Scavenging Applications
Nicodimus Retdian, Japan

O17-2
A Broadband Low Noise Amplifier for High Performance Wireless Microphones
Shui-Chuen Lin, Chang-Lang Tai, San-Fu Wang, Taiwan; Yitsen Ku, United States; Zi-Wei Wang, Shih-Chun Liu, Taiwan

O17-3
Investigation of Hybrid ADC Combined with First-Order Feedforward Incremental and SAR ADCs
Yuichiro Kobayashi, Tatsuji Matsuura, Ryo Kishida and Akira Hyogo, Japan

O17-4
Examination of Incremental ADC with SAR ADC That Can Reduce Conversion Time with High Accuracy
Yuichiro Unno, Tatsuji Matsuura, Ryo Kishida and Akira Hyogo, Japan

O17-5
A Plus Type CC-Based Current-Mode Universal Biquadratic Circuit
Takao Tsukutani, Japan

O17-6
Multi-Output Octagonal MOSFET for the Common Device of Both Sensor and Circuit Design
Tomochika Harada, Japan

O17-7
Design of 10 GHz CMOS Optoelectronic Receiver Analog Front-End in Low-Cost 0.18 um CMOS Technology
Xiangyu Chen and Yasuhiro Takahashi, Japan

Oral 18- Multimedia and Systems II

10:20-12:10 / The Sky Room
Chair: Chao-Hsien Lee, Taiwan
Yoshimitsu Kuroki, Japan

O18-1
Comparison of Face Recognition Loss Functions
Hung-Yi Wu, Min-Hsiang Chang and Gee-Sern Jison Hsu, Taiwan

O18-2
Detection of Dangerous Objects by Pan-Tilt Camera
Yurika Fujii, Minoru Fukumi, Shin-Ichi Ito and Momoyo Ito, Japan

O18-3
Zero-Phase Impulsive Noise Suppression with Iterative Phase Reconstruction
Aki Ishibashi and Arata Kawamura, Japan

O18-4
Target Speech Signal Extraction Using Variance of Phase Difference
Eiji Saito and Arata Kawamura, Japan

O18-5
Machine Learning-Based Rate Control Scheme for High Efficiency Video Coding
Jui-Hung Hsieh, Jing-Cheng Syu and Hui-Lan Zhang, Taiwan

O18-6
The Regression Model of NOx Emission in a Real Driving Automobile
Hung-Ta Wen, Ming-An Li and Jau-Huai Lu, Taiwan

O18-7
Dictionary Learning on ℓ_1-Norm Fidelity for Non-Key Frames in Distributed Compressed Video Sensing
Tsugumi Oishi and Yoshimitsu Kuroki, Japan

Oral 19- Analog ICs and Antenna Design for 5G Communication

10:20-12:10 / Jade, Stone and Spring Room
Chair: Wen-Cheng Lai, Taiwan
Shen Shou Max Chung, Taiwan

O19-1
Design of SAR ADC to Light Charging for Optical Sensors Applications
Wen Cheng Lai, Taiwan

O19-2
Multi Band Antenna Design for Mobile Devices
Wen Cheng Lai, Taiwan

O19-3
Frontend with Antenna for EM Body Analysis
Wen Cheng Lai, Taiwan

O19-4
A Broadband CPW-Fed Equilateral Pentagonal Slot Antenna
Narongrit Mekloi, Thailand

O19-5
Preliminary Design of a Waveguide-Fed Milimeter Wave Metasurface Antenna with LCD Controlled Array Factor for 5G User Equipment
Shen Shou Max Chung and Shih-Chung Tuan, Taiwan

O19-6
Duty Ratio and Capacitance Analysis of AC/DC Converter without Current Control Circuit
Yuki Tanaka, Tatsuji Matsuura, Ryo Kishida and Akira Hyogo, Japan

O19-7
Frontend Design for FMCW MIMO Radar Sensor
Wen Cheng Lai, Taiwan

Keynote Speech III

13:30-14:30 / Pearl Banquet Hall
Chair: Chia-Yu Yao, Taiwan

Making 5G Use Case a Commercial Reality
Mr. Pat Hsu, Taiwan

Oral 20- AI and the Interdisciplinary Research

15:00-16:30 / Pearl Banquet Hall
Chair: Cheng-Shian Lin, Taiwan
Yi-Hsien Wang, Taiwan

O20-1
Deep Wavelets for Heart Sound Classification
Kun Qian, Japan; Zhao Ren, Germany; Fengquan Dong, China; Wen-Hsing Lai, Taiwan; Björn Schuller, Germany; Yoshiharu Yamamoto, Japan

O20-2
Combine Facebook Prophet and LSTM with BPNN Forecasting Financial Markets: The Morgan Taiwan Index
Wen-Xiang Fang, Po-Chao Lan, Wan-Rung Lin, Hsiao-Chen Chang, Hai-Yen Chang and Yi-Hsien Wang, Taiwan

O20-3
An Analysis of Combining Correlation Screening with Artificial Neural Network for FITX Futures Prediction
Fu-Ming Lai, Wen-Xiang Fang, Taiwan; Chong-Kai Fu, China; Fu-Ju Yang, Kuang-Hsun Shih and Yi-Hsien Wang, Taiwan

O20-4
Comparison of Forcasting Ability between Back-Propagation Network and ARIMA in the Prediction of Bitcoin Price
Chung-Chieh Chen, Jung-Hsin Chang, Cheng-Shian Lin, Jui-Cheng Hung, Fang-Cih Lin and Yi-Hsien Wang, Taiwan

Oral 21- Communication Systems IV

15:00-16:30 / The Sky Room
Chair: Kenichi Higuchi, Japan
Takahiro Matsumoto, Japan

O21-1

Single-Input-Single-Output System with Interference-Unaware Time-Based Receive Transformation under Cochannel Interference and Intersymbol Interference

Jui Teng Wang and Tsun-Chia Wang, Taiwan

O21-2

Time-Space MIMO System with Interference-Unaware Time-Space Receive Transformation under Cochannel Interference and Intersymbol Interference

Jui Teng Wang and Tsun-Chia Wang, Taiwan

O21-3

A Generation Method of a Two-Dimensional Optical ZCZ Sequence with the Zero-Correlation Zone (4n-2) * (4n-2)

Takahiro Matsumoto, Hideyuki Torii, Yuta Ida and Shinya Matsufuji, Japan

O21-4

Repetition Control Method Using Terminal Mobility for Uplink Grant-Free URLLC

Shinichi Ozaku, Yukiko Shimbo, Hirofumi Suganuma and Fumiaki Maehara, Japan

O21-5

BER Detection of A2G Wireless Communication in Rician K-Factor Fading Channel for Massive IoT Connectivity Network

Sarun Duangsuwan, Anna Chusongsang, Chakree Teekapakvisit and Sathaporn Promwong, Thailand

O21-6

Performance Evaluation of IDMA-Based Random Access Considering User Detection and Channel Estimation

Taiki Tomizawa, Yoshihisa Kishiyama and Kenichi Higuchi, Japan

Oral 22- Signal and Image Processing for Advanced Technology II

15:00-16:30 / The East and West Room
Chair: Ting-Lan Lin, Taiwan
Mitsuji Muneyasu, Japan

O22-1

Image Regularization with Morphological Gradient Priors Using Optimal Structuring Elements for Each Pixel

Shoya Oohara, Hirotaka Oka, Mitsuji Muneyasu, Soh Yoshida and Makoto Nakashizuka, Japan

O22-2

Atrial Fibrillation Detection in Spectrogram Based on Convolution Neural Networks

Jing-Ming Guo, Chiao-Chun Yang, Zong-Hui Wang, Chih-Hsien Hsia and Li-Ying Chang, Taiwan

O22-3

Design and Implementation of Ultra-Low-Latency Video Encoder Using High-Level Synthesis

Kosuke Fukaya, Kaito Mori, Seiji Mochizuki, Kousuke Imamura, Yoshio Matsuda and Tetsuya Matsumura, Japan

O22-4

A Convolutional Dictionary Learning Based ℓ_1Norm Error with Smoothed ℓ_0 Norm Regression

Kaede Kumamoto, Shinnosuke Matsuo and Yoshimitsu Kuroki, Japan

O22-5

Convolutional Sparse Dictionary Learning with Smoothed ℓ_0 Norm and Projected Gradient Descent

Kazuki Kitajima, Akira Sugano and Yoshimitsu Kuroki, Japan

O22-6

Improvements on Data Insertion Technique in Encrypted Image Against Lossy Compression

Ryoma Ito, Japan; Koksheik Wong, Simying Ong, Malaysia; Kiyoshi Tanaka, Japan

Oral 23- Intelligent Signal Processing Technique and Its Applications

15:00-16:30 / Jade, Stone and Spring Room
Chair: Ying-Ren Chien, Taiwan
Huang-Chang Lee, Taiwan

O23-1
An Efficient Speech Recognition Algorithm for Small Intelligent Electronic Devices
Zhichao Zheng, Xiaotao Lin, Weiwei Zhang, Jianqing Zhu and Huanqiang Zeng, China

O23-2
A Transformation for Polar Code BP Decoding
Yi-Ru Chen, Huang-Chang Lee and Mao-Chao Lin, Taiwan

O23-3
Subjective Interpupillary Distance of Measurement Technique
Jih-Yi Liao, Shih-Tsung Chang, Chao-Han Wu, Der-Chin Chen, Chung-Ping Chen and Cheng-Ke Hsu, Taiwan

O23-4
Cross Conditional Network for Speech Enhancement
Haruki Tanaka, Yosuke Sugiura, Nozomiko Yasui, Tetsuya Shimamura and Ryoichi Miyazaki, Japan

O23-5
RSSI Measurement with Channel Model Estimating for IoT Wide Range Localization Using LoRa Communication
Yi-Cheng Lin, Chi-Chia Sun and Kuo-Ting Huang, Taiwan

Oral 24- VLSI

16:30-18:00 / Pearl Banquet Hall
Chair: Chia-Yu Yao, Taiwan
Wei-Wen Hsu, Taiwan

O24-1
Baby Care System Design for Multi-Sensor Applications
Tsung-Hsun Wu and Pei-Yin Chen, Taiwan

O24-2
An All-Digital Clock Generator with Modified Dynamic Frequency Counting Loop and LFSR Dithering
Pao-Lung Chen, Taiwan

O24-3
Performance Evaluation of Heterogeneous Cluster for Satellite Data Processing
Sethakarn Prongnuch, Suchada Sitjongsataporn and Theerayod Wiangtong, Thailand

O24-4
Softsign Function Hardware Implementation Using Piecewise Linear Approximation
Chih-Hsiang Chang, En-Hui Zhang and Shih-Hsu Huang, Taiwan

Oral 25- Communication Systems V

16:30-18:00 / The Sky Room
Chair: Kenichi Higuchi, Japan
Sankar Srinivasan, Taiwan

O25-1
Evaluation of Wireless Body Area Network Utilizing Super Orthogonal Convolutional Code
Kento Takabayashi, Hirokazu Tanaka and Katsumi Sakakibara, Japan

O25-2
Probability Distribution Analysis of Backoff Time with Frozen Backoff in CSMA/CA
Koki Samoto, Ikuo Oka and Shingo Ata, Japan

O25-3
Phase Correction for Automatic Modulation Classification Using Iterative Closest Point
Wataru Machida, Kei Ichijo, Yosuke Sugiura and Tetsuya Shimamura, Japan

O25-4
Inter-Cell Coordinated Transmission Power Control for IDMA-Based Random Access
Masayuki Kawata, Tsubasa Miyata, Yoshihisa Kishiyama and Kenichi Higuchi, Japan

O25-5
Low Latency HARQ Method Using Early Retransmission Before Channel Decoding Based on Superposition Coding
Koka Miura, Yoshihisa Kishiyama and Kenichi Higuchi, Japan

Best Student Paper Award

16:30-18:00 / The East and West Room
Chair: Akira Taguchi, Japan
Huanqiang Zeng, China

SA-01
Machine Learning Model with Technical Analysis for Stock Price Prediction: Empirical Study of Semiconductor Company in Taiwan

Po-Chao Lan, Wei-Ling Kung, Yao-Lun Ou, Chun-Yueh Lin, Wen-Cheng Hu and Yi-Hsien Wang, Taiwan

SA-02
A Spur-Suppression Technique for Frequency Synthesizer with Pulse-Width to Current Conversion

Po-Yu Hsieh, Shao-Yu Shu and Ching-Yuan Yang, Taiwan

SA-03
Action Conditioned Response Prediction with Uncertainty for Automated Vehicles

Hayoung Kim, Gihoon Kim, Jongwon Park, Kyushik Min, Dongchan Kim and Kunsoo Huh, South Korea

SA-04
Image Classification by Multilayer Feature Extraction Based on Nuclear Norm Minimization

Tomoya Hirakawa, Shohei Kubota and Yoshimitsu Kuroki, Japan

SA-05
Embedded Implementation of Human Detection Using Only Color Features on the NVIDIA Xavier

Masahiko Tsuyama, Takuro Oki, Shingo Kobayashi, Risako Aoki, Ryusuke Miyamoto, Hiroyuki Yomo and Shinsuke Hara, Japan

Oral 26- Intelligent Visual Perception

16:30-18:00 / Jade, Stone and Spring Room
Chair: Li-Ying Chang, Taiwan

O26-1
Irreversible Privacy-Preserving Images Holding Spatial Information for HOG Feature Extraction

Masaki Kitayama and Hitoshi Kiya, Japan

O26-2
XOR-Ed Based Friendly-Progressive Secret Sharing

Heri Prasetyo and Joni Welman Simatupang, Indonesia

O26-3
Influence of Significant Target on Image Quality Assessment via EEG

Tianyan Wu, Lihuo He, Hongxia Cai, Wen Lu and Xinbo Gao, China

Keynote Speech IV

09:00-10:00 / Pearl Banquet Hall
Chair: Jing-Ming Guo, Taiwan

Forgery Detectors for Adversarial Machine Learning
Prof. Pierre Moulin, USA

Invited Speech IV / Oral 27- Signal Processing III

10:20-12:10 / Pearl Banquet Hall
Chair: Chien-Cheng Tseng, Taiwan
Minoru Komatsu, Japan

Invited IV
Intellectual Property Protection of Deep Learning Models
Chee Seng Chan, Malaysia

O27-1
Designing High-Performance Green Filters Using Downsampling Techniques
David Shiung and Wen-Long Chin, Taiwan

O27-2
Proposal of Extracting Pulse Wave During Driving a Car Based on Frequency Domain BSS
Kakeru Ishikawa, Minoru Komatsu and Hiroki Matsumoto, Japan

O27-3
A Supervised Learning Method for the Design of Linear Phase FIR Digital Filter Using Keras
Chien-Cheng Tseng and Su-Ling Lee, Taiwan

O27-4
Improvement of Noise Suppression Performance of SEGAN by Sparse Latent Vectors
Minami Sakuma, Yosuke Sugiura and Tetsuya Shimamura, Japan

O27-5
Novel Deblocking Method for Cropped Video
Kenta Hashimoto and Seiichi Gohshi, Japan

O27-6
Design of Graph Filter Using Spectral Transformation and Window Method
Chien-Cheng Tseng and Su-Ling Lee, Taiwan

Oral 28- Signal and Image Processing for Advanced Technology III

10:20-12:10 / The Sky Room
Chair: Naoto Sasaoka, Japan
Yoshimitsu Kuroki, Japan

O28-1
Two-Stage Fingerprint Recognition System Based on Robust Binary Invariant Feature and Multiple Image Quality Metrics
Jing-Ming Guo, Taiwan; Shao-En Lee, Sierra Leone; Li-Ying Chang, Taiwan

O28-2
Pre-Inverse Active Noise Control System with Virtual Sensing Technique for Non-Stationary Path
Kohei Matsuhisa, Keisuke Okano, Naoto Sasaoka and Yoshio Itoh, Japan

O28-3
A Nail Image Analysis Method to Evaluate Accumulated Stress Using Fuzzy Reasoning
Kazuki Shimamoto, Shin-Ichi Ito, Momoyo Ito and Minoru Fukumi, Japan

O28-4
Distributed Compressed Video Sensing Based on Convolutional Sparse Coding
Tomohito Mizokami and Yoshimitsu Kuroki, Japan

O28-5
Personal Authentication by Walking Motion Using Kinect
Chunyu Guo, Japan

O28-6
Data Retrieval from Printed Image Using Image Features and Data Embedding
Takuhiro Nishikawa, Mitsuji Muneyasu, Yuuki Nishida, Soh Yoshida, Japan; Kosin Chamnongthai, Thailand

O28-7
Study on Discrimination of Finger Motions based on EMG Signals
Xiao Shan and Minoru Fukumi, Japan

Poster Session I Dec. 4, 2019

<u>P1-1</u>
Residual Concatenated Network for ODBTC Image Restoration
Alim Wicaksono, Heri Prasetyo, Indonesia; Jing-Ming Guo, Taiwan

<u>P1-2</u>
Trend Prediction of Influenza and the Associated Pneumonia in Taiwan Using Machine Learning
Sing-Ling Jhuo, Mi-Tren Hsieh, Ting-Chien Weng, Mei-Juan Chen, Chieh-Ming Yang and Chia-Hung Yeh, Taiwan

<u>P1-3</u>
A Modified Structural Similarity Index with Low Computational Complexity
Woei-Tan Loh and David Bong, Malaysia

<u>P1-4</u>
Quality Enhancement of DDBTC Decoded Image
Joni Welman Simatupang and Heri Prasetyo, Indonesia

<u>P1-5</u>
A 10-bit 100-MHz Current-Steering DAC with Randomized Thermometer Code Calibration Scheme
Hsin-Liang Chen, Ming-Han Hsieh and Jem-Shiun Chiang, Taiwan

<u>P1-6</u>
Reconfirm Gestalt Principles from Scan-Path Analysis on Viewing Photos
Hsien Chih Chuang and Da Lun Tang, Taiwan

<u>P1-7</u>
Controlling Myopia Progression in Children by the Rotary Prism Eye Exercise Device
Der-Chin Chen, Jih-Yi Liao, Wei-Hsin Chen, Po-Ting Liu and Shih-Tsung Chang, Taiwan

<u>P1-8</u>
Merge Mode-Based Data Embedding in SHVC Compressed Video
Kok Sheik Wong, Lie Lin Pang and Ryoma Ito, Malaysia

<u>P1-9</u>
Unpaired Object Transformation Based on Generative Adversarial Networks
Jing-Ming Guo, Yun-Fu Liu and Chang-Chun Chu, Taiwan

<u>P1-10</u>
An End to End Single Image Dehazing System Based on Dense Block and Hybrid Loss Function
Jing-Ming Guo and Yen-Chia Su, Taiwan

<u>P1-11</u>
Nighttime Image Dehazing Based on Improved Erosion Dark Channel and Multi-Scale Clipping Limit Histogram Equalization
Jing-Ming Guo and Chia-Hsiang Lin, Taiwan

<u>P1-12</u>
Reconstruction of Ordered Dithering Halftone Image
Jing-Ming Guo, Hung Le and Sankarasrinivasan Seshathiri, Taiwan

<u>P1-13</u>
User Trajectory Analysis within Intelligent Social Internet-of-things (SIoT)
Guang Xing Lye, Wai Khuen Cheng, Teik Boon Tan, Zeng-Wei Hong, Malaysia; Yen-Lin Chen, Taiwan

<u>P1-14</u>
Crowd Behavior Classification Based on Generic Descriptors
Pei Voon Wong, Norwati Mustapha, Lilly Suriani Affendey, Fatimah Khalid, Malaysia; Yen-Lin Chen, Taiwan

<u>P1-15</u>
Actived Edge Strength for Image Quality Assessment
Minjuan Gao and Xuande Zhang, China

<u>P1-16</u>
SRGNet: A GRU Based Feature Fusion Network for Image Denoising
Wenhao Wang, Cheng Pang, Zhenbing Liu, Rushi Lan and Xiaonan Luo, China

P2-1
A 11.75-bit Hybrid Sturdy MASH-21 Delta-Sigma Modulator for Audio Applications
Ting-Hui Chang and Chia-Yu Yao, Taiwan

P2-2
Using Inversion-Mode MOS Varactors and 3-Port Inductor in 0.18-um CMOS Voltage Controlled Oscillator
Ming-Xuan Li, Ching-Yi Jiang, Yao-Wen Hsu, Yu-Ying Pan and Hao-Hui Chen, Taiwan

P2-3
Capacitance Minimization Clock Synthesis with Blockage-Avoiding Hybrid-Structure Network
Chun-Wei Ho and Shao-Yun Fang, Taiwan

P2-4
Latency Constraint-Aware Scheduler for NVMe Solid State Drives
Cheng-Yu Li, Jih-Hsiang Cheng and Ya-Shu Chen, Taiwan

P2-5
MMW Receiver Front-End for Noninvasive Glucose Measurement
Ming-Yu Yen, Fang-Yu Zhou, Wen-Ling Chang and Hsiao-Chin Chen, Taiwan

P2-6
Rehabilitation Seat Cushion System
Ming-Ta Ke, Po-Cheng Su, Ya-Hsin Hsueh, Yen-Chin Lin, Hsiang-Lung Huang, Yu-Jhang Wu, Jyun-Jhe Chen and Yi-Ting Zhong, Taiwan

P2-7
Automatic Spine Vertebra segmentation in CT images Using Deep Learning
Ping-Cheng Wu, Teng-Yi Huang and Chun-Jung Juan, Taiwan

P2-8
Trigeminal Neuralgia Alleviation on Demand with an CMOS SoC Using Current-Mode Pulsed Radio-Frequency Stimulation
Hung-Wei Chiu, Kun-Ying Yeh, Shey-Shi Lu and Mu-Lien Lin, Taiwan

P2-9
Multi-Mode Halftoning Using Stochastic Clustered-Dot Screen
Jing-Ming Guo, Yun-Fu Liu and Shih-Chieh Lin, Taiwan

P2-10
Efficient Finger-Vein Recognition System Based on Fast Binary Robust Independent Elementary Feature Combined with Multi-Image Quality Assessment Verification
Chih-Hsien Hsia, Jing-Ming Guo and Chong Sheng Wu, Taiwan

P2-11
A New Contactless Deception Detection System with Hybrid Facial Features
Chih-Hsien Hsia, Jing-Ming Guo, Taiwan; Li-Wei Hsiao, Sierra Leone

P2-12
Cervical Image Segmentation Using U-Net Model
Yao Liu, Bing Bai, China; Hua-Ching Chen, Taiwan; Peizhong Liu, China; Hsuan-Ming Feng, Taiwan

P2-13
Using QR Code Labels to Enhance OCR for Capturing Legacy Machines' Data
Suk-Ling Lai, Boon-Yaik Ooi, Malaysia; Yen-Lin Chen, Taiwan

P2-14
Probabilistic Image Quality Assessment: An Economics Point of View
Guangtao Zhai, China

P2-15
Generative Adversarial Network-Based Image Super-Resolution with a Novel Quality Loss
Xining Zhu, Lin Zhang, Lijun Zhang, Xiao Liu, Ying Shen and Shengjie Zhao, China

P02-16
A Heart Rate Monitoring and Activities Recognition System for Badminton Training
Yu-Wen Huang, Shu-Hua Chang, Yu-Chih Tsai and Yuan-Hsiang Lin, Taiwan

Classification of Environmental Sounds Using Convolutional Neural Network with Bispectral Analysis

Katsumi Hirata
Dept. of Innovative Electrical and Electronic Engineering
National Institute of Technology (KOSEN), Oyama College
Oyama, Tochigi, Japan
hirata@oyama-ct.ac.jp

Takehito Kato
Dept. of Mechanical Engineering
National Instatitute of Technology (KOSEN), Oyama College
Oyama, Tochigi, Japan
kato_t@oyama-ct.ac.jp

Ryuichi Oshima
Dept. of Archtecture
National Instatitute of Technology (KOSEN), Oyama College
Oyama, Tochigi, Japan
o-shima@oyama-ct.ac.jp

***Abstract*—To realize a useful acoustic environmental recognition system, we propose a new method that classifies sound signals using a slice bispectrogram, which is a third-order version of a spectrogram. The classified sound was used as input data for a convolutional neural network. We conducted a fundamental classification experiment using UrbanSound8k, which was an open dataset consisting of 10 classes of environmental sounds. Our proposed method gave high accuracy and stability. Furthermore, a relationship between the accuracy and non-Gaussianity of sound signals was confirmed.**

Keywords—environmental sound, bispectrum, spectrogram, slice bispectrogram, convolutional neural network

I. Introduction

Certain systems can function well only if they are able to recognize the sound environment as humans do. These systems include self-driving cars, autonomous robots, and the systems that support the hearing impaired. In this research, we focus on the sound classification and aim to develop a method that automatically classifies various environmental sounds. Researchers have proposed a few methods for the classification of environmental sounds. One common method is the use of a convolutional neural network, which is a deep learning technique [1]. It is very important to create input data that best describes the characteristics of the original data because the performance of a method based on neural networks depends on the input data [2]. Time-frequency analysis, such as the spectrogram, is effective for time-varying sounds. Although the power spectrum used in the spectrogram is not able to describe higher-order statistics, higher-order spectra allow us to analyze higher-order signal components and to extract useful characteristics. Therefore, we propose a method that uses a third-order spectrogram, which is based on bispectral analysis. In this paper, we explain the *slice bispectrogram* (SBS) and discuss the effectiveness of the method by evaluating the experimental results using the UrbanSound8k sound dataset [3, 4].

II. Bispectrogram

A. Bispectral Analysis

A power spectrum $P(k)$ of a random signal $x(n)$ is estimated from the expectation of the second-order periodogram using the Wiener–Khinchin theorem,

$$P(k) = E[|X(k)|^2], \tag{1}$$

where $X(k)$ is the Fourier spectrum of $x(n)$, and $E[\,]$ denotes the expectation. Similarly, the estimate of a bispectrum $B(k_1, k_2)$ is obtained as the following equation because the bispectrum is defined by the Fourier transform of the third-order auto-correlation function [5]:

$$B(k_1, k_2) = E[X(k_1)X(k_2)X^*(k_1 + k_2)] \tag{2}$$

where X^* denotes the complex conjugate of X. It represents the physical linear dependency among the components on the three frequencies k_1, k_2 and $k_1 + k_2$.

B. Slice Bispectrogram

Certain kinds of environmental sounds are time varying, that is, they are non-stationary. For such signals, it is not suitable to apply the power spectral and bispectral analyses mentioned above because the spectra represent the averaged spectral density during the analyzed interval. Using the short-time Fourier transform is effective for non-stationary signals as follows:

$$X(m, k) = \sum_{n=0}^{N-1} w(n - m)x(n) \exp(j2\pi kn/N), \tag{3}$$

where $w(n - m)$ is a window function centered on the time m. Therefore, this equation represents a spectrum around each time m. Applying $X(m, k)$ to (1) instead of $X(k)$, we obtain the short-time power spectrum, i.e. *spectrogram* as follows:

$$S_2(m, k) = E[|X(m, k)|^2]. \tag{4}$$

We refer to this spectrogram as the *power spectrogram* (PS) to distinguish this from another spectrogram mentioned subsequently. Similarly, applying (3) to (2), we obtain the short-time bispectrum as follows:

$$B(m, k_1, k_2) = E[X(m, k_1)X(m, k_2)X^*(m, k_1 + k_2)], \tag{5}$$

and it cannot be displayed as a two-dimensional image. This is why we use the SBS

$$S_3(m, k) = |E[X(m, k)X(m, k)X^*(m, 2k)]|, \tag{6}$$

which is the slice version of the amplitude for the short-time bispectrum at $k_1 = k_2 = k$ [6].

III. Convolutional neural network

Our proposed method uses a convolutional neural network to classify the environmental sounds. The layer structure of the neural network used in our experiments is illustrated in Fig. 1. The network consists of five convolutional layers, three pooling layers, and a fully connected layer. Each convolutional layer is followed by normalizing and applying the ReLU function to effectively extract characteristics; a dropout operation is applied before the fully connected layer. After training the network using a large amount of spectrogram 2-D images generated from labeled environmental sounds, the images from the unknown class of sounds are input to the first convolutional layer, and a class ID corresponding to the sound signal is given as the output.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

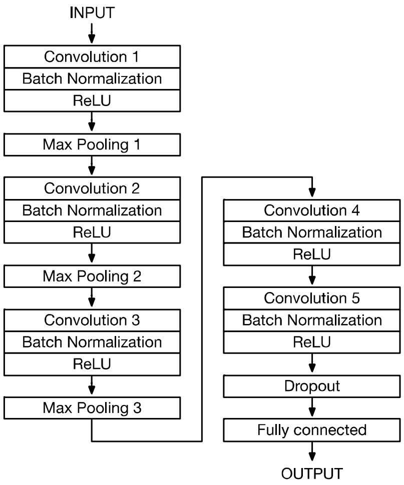


Fig. 1. Structure of a convolutional neural network.

IV. Classification experiments

We conducted experiments using either a PS or an SBS as the input for the neural network. We also discuss the effectiveness of our SBS method and by comparing and evaluating the classification accuracies.

A. Conditions

An environmental sound dataset UrbanSound8k was used as both the training and the test data. The dataset contained 8732 labeled sound excerpts of urban sounds from 10 classes, as shown in Table III [3, 4]. The experiments were implemented using MATLAB. The main conditions of the experiment are shown in Table I. Every sound signal in the dataset was randomly allocated to one of the 10 folders. In the experiments, we conducted a 10-fold cross validation as follows. The network was trained on data from 9 of the 10 folders and tested on data from the remaining folders. This process was repeated 10 times (each time using a different set of 9 out of the 10 folders for the training and using the remaining set for testing). Subsequently, we evaluated the performance using average accuracies and standard deviations over all 10 tests for each method. We used both PS and SBS.

TABLE I. Conditions of the experiment

Sampling frequency	8 kHz
Duration of each signal	4 sec
Number of points (pixels) of each spectrogram image	Time: 128 points Frequency: 128 points
Convolution layer 1	Filter size: 3×3, Number of filters = 12
Convolution layer 2	Filter size: 3×3, Number of filters = 24
Convolution layers 3, 4, 5	Filter size: 3×3, Number of filters = 48
Pooling layers 1, 2, 3	Size: 3×3
Dropout	50 %
Fully connected layer	10 output

B. Results

The averaged accuracies and their standard deviations for each method are shown in Table II. The accuracies for both methods reached more than 70%; SBS was achieved at a little higher accuracy and more stably than the conventional PS. Table III shows the accuracies for each class of sound, and the values are the rates at which sounds are classified correctly.

TABLE II. Total Accuracy of classifications [%]

	PS method	SBS method
Average	70.1 %	70.7 %
Standard deviation	6.4 %	5.6 %

TABLE III. Accuracy of classifications for each class [%]

ID, Class of sound	PS method	SBS method
0 Air conditioner	46.5 %	46.9 %
1 Car horn	78.1 %	81.2 %
2 Children playing	78.9 %	78.9 %
3 Dog bark	80.1 %	79.2 %
4 Drilling	74.7 %	77.3 %
5 Engine idling	58.9 %	55.4 %
6 Gun shot	88.0 %	91.1 %
7 Jackhammer	53.3 %	57.8 %
8 Siren	76.0 %	74.4 %
9 Street music	81.7 %	82.7 %

SBS is based on third-order statistics; therefore, it must represent the non-Gaussian characteristics that the conventional PS cannot represent. To clarify whether using SBS is effective for non-Gaussian signals, we examined the non-Gaussianity of signals and the effects of classification accuracy. Skewness is a measure of the asymmetry of a signal about the mean. The skewness of Gaussian signal is 0, and its absolute goes higher for more non-Gaussian signals. Fig. 2 shows the relative accuracies of SBS and PS and the averaged skewness for every class. The relative accuracy trends correspond to the skewness with the exception of ID3 and ID5. Therefore, although SBS had some effectiveness for classifying the non-Gaussian sounds, it was not effective despite its high skewness values, such as ID3 or ID5. We will need to investigate this in our future studies.

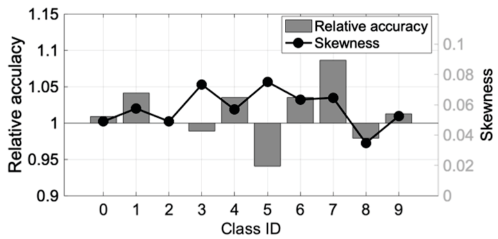


Fig. 2. Relationship between accuracies using SBS and skewness.

V. Conclusion

In this paper, we introduced SBS to classify environmental sounds using convolutional neural network having a high accuracy. Fundamental experiments were conducted to clarify the validity and effectiveness of the proposed method. Our results prove the high accuracy of the proposed method. In the future, we propose to further study the effectiveness of SBS in clarifying the properties of objects.

References

[1] K. J. Piczak, "Environmental sound classification with convolutional neural networks," 2015 IEEE international workshop on machine learning for signal processing, Sept. 2015.

[2] M. Huzaifah, "Comparison of time-frequency representations for environmental sound classification using convolutional neural networks," ArXiv Prepr. ArXiv170607156, 2017.

[3] "Urbansound8k dataset", https://urbansounddataset.weebly.com/urbansound8k.html (Last accessed at Jun. 20, 2019).

[4] J. Salamon, C. Jacoby, and J. P. Bello, "A dataset and taxonomy for urban sound research," MM '14 Proceedings of the 22nd ACM international conference on Multimedia, pp.1041-1044, Nov. 2014.

[5] C. L. Nikias and A. P. Petropulu, "Higher-order spectra analysis: a nonlinear signal processing framework," Prentice Hall, 1993, pp.7-30

[6] V. Swarnkar, U. Abeyratne, and C. Hukins, "Objective measure of sleepiness and sleep latency via bispectrum analysis of EEG," Medical and & biological engineering & computing, 48, pp.1203-1213, Dec. 2010.

Restoration of Compressed Picture Based on Lightweight Convolutional Neural Network

Tien-Ying Kuo, Yu-Jen Wei, Chang-Hao Chao
Dept. of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, R.O.C.
tykuo@ntut.edu.tw, {yjwei, chchao}@image.ee.ntut.edu.tw

***Abstract*-- In this work, we propose a deep learning-based method to improve the quality of JPEG images. Our proposed network predicts the compression loss of the JPEG image for compensating and restoring the image quality. To solve the color bleeding artifacts often found in JPEG image, our network considering it in our model and objective functions to restore the color channels. Our network is much lighter by using fewer parameters compared to other work, whereas our method can still provide satisfactory and well-restored images for JPEG images as demonstrated in our experiments. Even with the additional handling on the color channels, the number of parameters in our network model is still kept low around 224k.**

Keywords- JPEG, compression artifacts removal, convolutional neural network, residual network , dilated convolution

I. Introduction

Reducing media storage and network transmission is critical as the resolution of image applications become higher than before. To achieve this goal, the technique of image compression will play the role. Image compression, especial the lossy one, could achieve higher compression ratio, but it would cause the irreversible loss of information and may result in observable artifacts affecting the perceptional quality.

To restore the lost quality of the images after compression, many recent methods were developed based on deep learning technique [1, 2] . For these methods, the number of network layers usually has to be made deep enough to obtain acceptable restoration results. But this will add the network parameters and places burdens on the operating devices, especially, more on the low-end ones. In addition, the models and training methods proposed by the deep learning-based image restoration work [1, 2] only targeted for the luminance channel of images. Though the visual impact on the luminance quality indeed is the largest and dominant, the artifacts occurred in the color channels can still be noticeable and affect the quality of visual perception with color bleeding artifacts.

Thus, the objective of this work is to solve the above-mentioned problems. We try to design a model with a fewer number of parameters. The model can further remove the artifacts in color channels of the image to achieve better visual quality. Unlike some of the literature work, we targeted at a single model with a single set of model parameters to deal with different JPEG compression qualities. This will save the parameter usage, but on the other hand, it will make the design more difficult.

II. Proposed Method

A. Network Architecture

Our designed method adopts the ConvNet architecture to predict the residual value of all YCbCr color channels between the compressed and the original images, and then put the predicted residues back to the compressed image for restoring the image. The network architecture of our proposed method is illustrated in Fig. 1. where the orange block indicates the convolution layer and the green blocks represents the recursive networks.

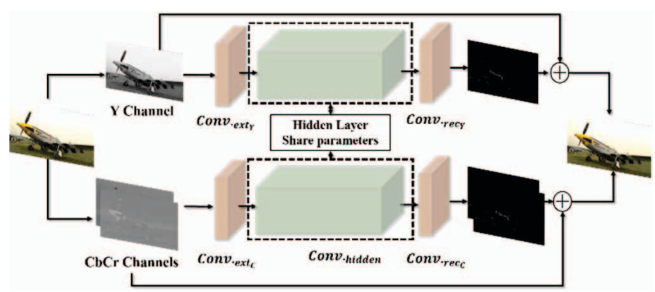
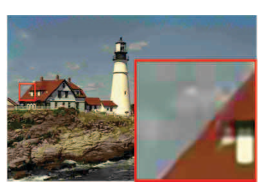


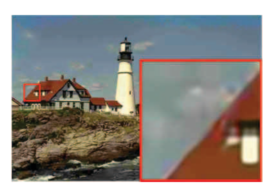
Fig. 1 Flow chart of the proposed method

This model can be roughly segmented into three sub-nets: the feature extraction layers including $Conv._{ext_Y}$ and $Conv._{ext_C}$, the hidden layers $Conv._{hidden}$, and the feature reconstruction layers including $Conv._{rec_Y}$ and $Conv._{rec_C}$.

(a) Compressed Image with Lighthouse2



(b) Image with Y channel restored only with Lighthouse2



(c) Image with both Y and CbCr channels restored with Lighthouse2

Fig. 2 Comparison of compressed images (QF=10) and restored images

Since the luminance is much more important than chrominance, so we use $Conv._{ext_Y}$ to extract the 64-channel features of the luminance channel, and $Conv._{ext_C}$ to retrieve a total of 64-channel features from the CbCr two channels. Two convolutional layers both use a kernel size of 3×3 to extract features.

The intermediate net contains six convolutional layers, and each convolutional layer uses 64 filters with 3×3 kernel size activated by PReLU. Our intermediate net processes the features in the luminance using three times of recursion, whereas there is no recursion in the less-sensitive Chroma channels for saving computational complexity.

Finally, the feature reconstruction net generates the residual image from $Conv._{rec_Y}$ and $Conv._{rec_C}$, in which architecture is the same as the $Conv._{ext_Y}$ and $Conv._{ext_C}$.

B. Extension of Receptive Field of Model

In order to have a good performance using a limited number of layers, we will integrate the dilated convolution to enlarge the receptive field in our model. The dilated convolution is similar to a standard convolution but different in the interval of the kernel of the filter, namely the dilated coefficient. We designed our dilated coefficients for the convolutional layers in three sizes. By adopting the dilated convolution, the receptive field of our model will be raised to 101×101 on the luminance channel, and from 16×16 to 37×37 on the color channels.

C. Objective Function

We use L2 loss as the objective function for model training. Compared with other literature methods which only deal with the artifacts in luminance, our experiment will train simultaneously in both luminance and color channels and achieve better results. Therefore, we need to consider both the luminance and the chroma in the loss function. So, we used (1) as the loss function of the model, λ is set to 0.25.

$$Loss = Loss_Y + \lambda Loss_C \tag{1}$$

D. Optimizer

Wilson [3] pointed out that Adam is prone to converge poorly and miss the optimal solution at the end of training. SGD, in contrast, is difficult to set up in the early stages of training, but it is more likely to find the best model if it is set up properly, so we adopt it as our loss optimizer.

E. Preprocessing and Augmentation of Datasets

We used the JPEG encoder [4] to compress images. Our model is trained with 400 images and 900 images from the BSDS500 [5] and DIV2K [6], respectively. The remaining 100 images of the BSDS500 are used as the validation set to adjust the hyper-parameters. We took 80×80 image patches as input when training the model.

III. Results

We used LIVE1 [7] dataset to test the performance of our model. In the analysis of complexity, we use the number of floating-point operations (FLOPs) instead of run time as a measure to be device-independent. In quality performance, we use PSNR and SSIM to make a judgment of the recovered image quality.

We compare our method with the models of the literature which can deal with different compression qualities using only a model and a set of model parameters. As shown in Table 1, the FLOPs, the number of parameters and the quality performance of each model were analyzed separately, the ones with top performance are denoted in red and blue as the first and second rank respectively. The results shown in Table 1 indicate that our model efficiency is similar to that of other models but with much lower computational complexity and fewer parameters.

Table 1. Comparison with other work on quality of image, number of parameters and FLOPs. (GFLOP = Giga FLOP)

		JPEG	DnCNN [1]	Zhan [2]	Ours
Average GFLOPs		X	**237.43**	614.94	317.25
# of parameters		X	667k	961k	**224k**
QF	Metrics	LIVE1 dataset (29 images)			
10	**PSNR(dB)**	27.77	29.20	**29.38**	29.35
	SSIM (N=11)	0.773	0.813	**0.818**	0.816
20	**PSNR(dB)**	30.07	31.59	**31.78**	31.74
	SSIM (N=11)	0.851	0.880	**0.884**	0.883

Fig. 2 shows the visual result of the image restoration examples produced by our proposed model. The artifacts of the compressed image are reduced greatly as compared to original compressed images on luminance. It improves even more when the color channel is also restored using our model.

IV. Conclusion

We use convolutional neural networks to learn the residual map between original and compressed images, and then put the predicted residues back to the compressed image for restoring the image. Unlike some literature work using multiple models or multiple sets of parameters of a single model, our proposed model uses a single model with a single set of model parameters to cope with different compression quality factors of JPEG images. The chrome is also processed in our model using the share-weights approach to further improve the restoration results without increase too much complexity.

Acknowledgment

This work was supported by Ministry of Science and Technology (grant # MOST 108-2221-E-027-034-)

References

[1] D. Zhang, D. Meng, L. Zhao, and J. Han, "Bridging saliency detection to weakly supervised object detection based on self-paced curriculum learning," arXiv preprint arXiv:1703.01290, 2017.

[2] W. Zhan, X. He, S. Xiong, C. Ren, and H. Chen, "Image deblocking via joint domain learning," Journal of Electronic Imaging, vol. 27, no. 3, p. 033006, 2018.

[3] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, "The marginal value of adaptive gradient methods in machine learning," in Advances in Neural Information Processing Systems, 2017, pp. 4148-4158.

[4] M. Version, "9.0. 0 (R2016a)," MathWorks *Inc., Natick, MA, USA,* 2016.

[5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 5, pp. 898-916, 2011.

[6] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2017, vol. 3, p. 2.

[7] H. Sheikh, "LIVE image quality assessment database release 2," http://live. ece. utexas. edu/research/quality, 2005.

Automatic Damage Recovery of Old Photos Based on Convolutional Neural Network

Tien-Ying Kuo, Yu-Jen Wei, Ming-Jui Lee, Tzu-Hao Lin
Dept. of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, R.O.C.
tykuo@ntut.edu.tw, {yjwei, mjlee, thlin}@image.ee.ntut.edu.tw

***Abstract*— Most of the methods for repairing old photos today are to manually process them using image editing software, such as Photoshop. The time of manual repairing is directly proportional to the damage degree of the photo, which is time consuming and laborious. Therefore, this paper proposes a two-stage convolution network to automatically repair damaged old photos. The first stage will detect the damaged areas of the photos, and the second stage will repair these areas. The experiment results demonstrates our method can successfully detect and repair the damage of the photos.**

Keywords—Damaged old photos, Image repariment, Image restoration, Image inpainting, Convolutional Neural Network.

I. INTRODUCTION

Old images are often subject to human damage or environmental influences during the preservation process. In order to recover the damaged areas, the most common method at present is to repair them artificially one by one through the photo editing software.

With the advent of image inpainting algorithms [1-3], the damaged areas can be automatically repaired based on the predefined damage marker of the image, whereas the integrity and consistency of the original image are still maintain, which can save a lot of repairing time. The damage marks required by these algorithms can be divided into semi-automatic [1, 2] and fully automatic methods [3] depending on whether or not to manually mark the damaged area. Semi-automatic method [1, 2] is the most common way at present. The area to be damaged is manually marked and then repaired by algorithm. Although it can save the repair time, the process of marking the area is still laborious. The fully automatic method is less common. The literature work [3] can automatically detect and then repair the damaged area, but it can only detect the simple and obvious damage, and the repairing ability still has room for improvement.

In order to solve the above problems, this paper proposes a fully automatic algorithm using Convolutional Neural Network (CNN) architecture. The algorithm proposed in this paper will automatically mark the damaged area in the image to save the time of manual marking, and then repair the damaged area without changing the consistency to the original photos, such as the overall tone of the photo.

II. PROPOSED METHODS

This section will introduce the two-stage model structure and design concept presented in this paper. Fig. 1 shows the schematic diagram of the architecture proposed in this paper. The first stage is used to detect and mark the location of the damage in the image. The second stage is to repair these areas and produce a recovered image.

A. Network Architecture

The two-stage model in this paper uses the same model architecture (except the output layer) but different parameter sets. The model structure is shown in Fig. 2. It can be divided into three parts, namely the feature extraction layer, the intermediate network layer and the feature reconstruction layer. The extraction layer and the reconstruction layer are all 3x3 convolutional layers with 64 channels, and the middle layer is three convolutional blocks designed by us to analyze the features acquired by the extraction layers.

B. Convolutional blocks

The purpose of the convolution block is to take the features of the extraction layer for analysis. The convolutional block of this paper combines two concepts of parallel structure and channel attention to improve the performance of the model

We design two branches in our convolutional blocks based on the parallel structure of GoogleNet [4]. The first branch is a two-layer 3x3 convolutional layer with 64 channels and ReLU activation functions, followed by channel attention. The other branch is channel attention mechanism, through which we can find out more important features from both shallow and deep layers respectively.

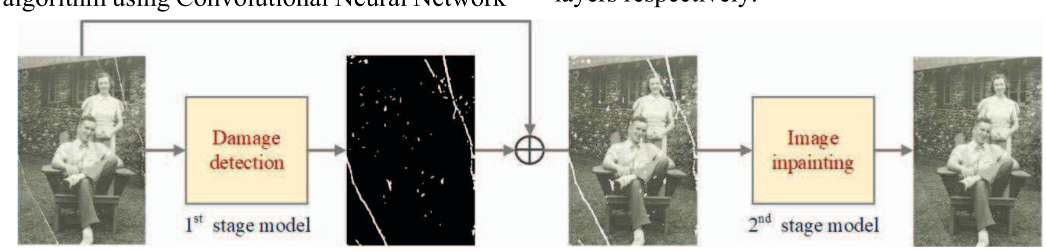


Fig. 1 Schematic architecture for proposed method

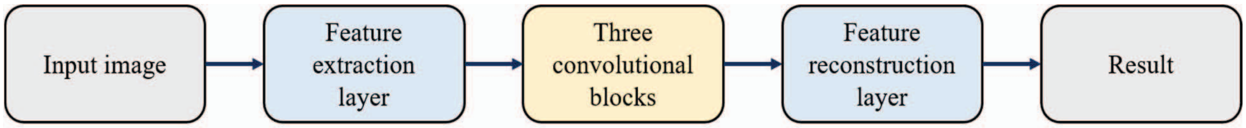


Fig. 2 Structure of proposed model

Channel attention originates from the inability of the human eye to receive all information while viewing an image, and can only process information on certain regions. SENet [5] applied this concept to image classification model respectively, to find out the most representative parts of the extracted features. This paper also adopts the concept similar to SENet [5] because the damage part usually catches human vision. This concept gives higher weights to more important features while reducing the influence of less important features.

Next, we concatenate the features of two branches, and then analyze these features through subsequent networks.

C. Short connection

Our model combines the shortcut connection concept of ResNet [6] to pass the feature information of the shallow network to the deep network for reconstruction, avoiding the phenomenon of gradient disappearance and maintaining the stability of the model training process.

III. Experiments

A. Settings for Training

A total of 190 old images with different contents were collected from the Internet, of which 100 were undamaged and 90 were damaged. Since the first stage was to detect the damaged area, we randomly picked out 50 damaged images as the training set and manually labeled the damaged areas of them, thereby generating the training data for the first stage model. The rest of damaged images are used to validate and test the experimental results. To create the second stage of the data set, we used 100 pieces of undamaged old photos and 50 damage masks to generate and generate 5000 pieces of training materials. Note that the technique of data augmentation is used in all of our dataset creation.

Because the tasks of the two-stage model are different, we choose different loss functions to evaluate the performance of the model. The first-stage model is to distinguish the damaged or undamaged areas of the old photos, binary cross entropy is used as the loss function. In the second stage of the model, MAE is used to train and judge the error between the model prediction and ground truth.

B. Evaluation

After the damaged area is completely repaired, the repair result may be slightly different from the original image, and it is not justifiable to use PSNR or SSIM to calculate the image difference. Therefore, the repaired results of this paper is evaluated subjectively.

Figure 3 shows the results of each stage. It can be seen from Fig. 3 (c) that most of the obvious damaged areas in the old photos have been detected. Fig. 3 (d) shows the image by applying our detection and repairing method. It can be observed that the area detected in Fig. 3 (c) has been repaired successfully and looks consistent with the undamaged areas.

(a) Damaged old photo

(b) Detected damage in first stage

(c) Overlaid image of (a) and (b)

(d) Repaired old photo

Fig. 3 Results of our proposed method

IV. Conclusions

We propose a two-stage network architecture to detect and repair damaged old photos automatically. Our network uses the same model architecture in each stage (except the output layer) but different parameter sets for damage detection and repairment. The concepts of parallel structure and channel attention is integrated in our network to enhance the performance. Experiments show that our method can accurately identify the distorted areas and successfully repair them.

Acknowledgment

This work was supported by Ministry of Science and Technology (grant # MOST 108-2221-E-027-034-)

References

[1] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, "Edgeconnect: Generative image inpainting with adversarial edge learning," *arXiv preprint arXiv:1901.00212,* 2019.

[2] Y. Liu, J. Pan, and Z. Su, "Deep Blind Image Inpainting," *arXiv: Computer Vision and Pattern Recognition,* 2017.

[3] E. Ardizzone, H. Dindo, and G. Mazzola, "Multidirectional scratch detection and restoration in digitized old images," *EURASIP Journal on Image and Video Processing,* vol. 2010, no. 1, p. 680429, 2010.

[4] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

[5] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132-7141.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

Wireless Video Transmission over Multiuser MIMO Systems with Fair Power Allocation

Yuta Yoshikawa, Koji Tashiro, Masayuki Kurosaki, and Hiroshi Ochi
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
{yoshikawa.yu, tashiro}@dsp.cse.kyutech.ac.jp, {kurosaki, ochi}@cse.kyutech.ac.jp

***Abstract*—With the spread of mobile devices, image and video transmission is in great demand. The use of multiuser multiple-input multiple-output (MU-MIMO) systems, where multiple data can be sent to multiple users, is highly expected. This paper proposes a novel power allocation scheme for MU-MIMO systems, which minimizes the variance of channel capacities of each spatial stream of all users to achieve user fairness in terms of video quality. In addition to the user fairness, we make the spatial streams which convey visually more important information be high-quality for achieving high video quality. Simulation results show that the proposed power allocation enables to equalize video quality even if the distances between a transmitter and users are different.**

***Index Terms*—Wireless Video Transmission, MU-MIMO Systems, User Fairness**

I. Introduction

The demand of image and video transmission has increased lately because of the dissemination of mobile devices. The use of multiple-input multiple-output (MIMO) systems, which increase the channel capacity owing to multiple antennas [1], is highly expected. In multiuser MIMO (MU-MIMO), an access point can transmit multiple data to multiple users. Recently, studies on the maximization of channel capacities of MU-MIMO have actively been conducted [2], [3].

In current standards of images and videos, the different parts of a code stream have different importance. Scalably encoded videos are composed of multiple units called layers. Each layer differs in the contribution to video quality. Even if the code stream has some errors, it is possible to improve video quality by protecting more significant layers [4].

In image and video transmission with MU-MIMO, it can happen that the video quality of the user in a poor communication environment remains low. Thus, it is important not only to improve sum throughput but also to keep the video quality of all users high as much as possible. This paper focuses on the user fairness in terms of video quality; that is, to equalize the video quality of all users. We present a fair power allocation (FPA) scheme, which minimizes the variance of channel capacities of each spatial stream of all users on condition that signals of more significant layers are sent through larger capacity spatial streams.

This work was supported in part by JSPS Grants-in-Aid for Scientific Research Grant Number JP17H01735 and by JSPS Grant-in-Aid for JSPS Research Fellow Grant Number 18J12075.

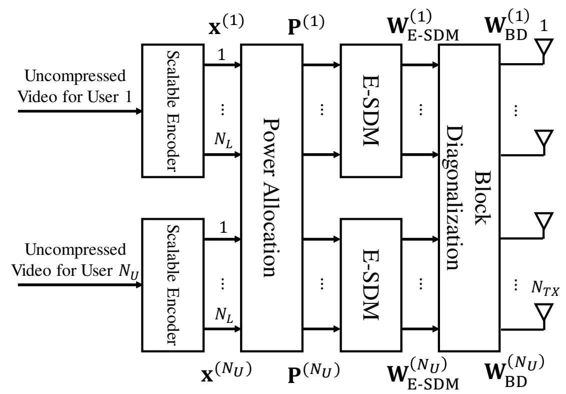


Fig. 1. Proposed transmitter which has N_{TX} transmit antennas and N_U users.

II. System Model

A system model of the proposed transmitter with the MU-MIMO system is shown in Fig. 1. The received signal vector for user i is expressed as

$$\mathbf{y}^{(i)} = \mathbf{H}^{(i)}\mathbf{W}^{(i)}_{\text{BD}}\mathbf{W}^{(i)}_{\text{E-SDM}}\mathbf{P}^{(i)}\mathbf{x}^{(i)} + \mathbf{z}^{(i)}, \qquad (1)$$

where $\mathbf{H}^{(i)}$ is a channel matrix, $\mathbf{W}^{(i)}_{\text{BD}}$ is a weight matrix of a block diagonalization [5], $\mathbf{W}^{(i)}_{\text{E-SDM}}$ is a weight matrix of eigenbeam space division multiplexing (E-SDM) [6], $\mathbf{P}^{(i)}$ is a power allocation matrix, $\mathbf{x}^{(i)}$ is a transmitted signal vector, and $\mathbf{z}^{(i)}$ is an additive Gaussian noise vector. The Scalable Encoder block generates a code stream which consists of N_L layers. Signals of each layer are transmitted through different spatial streams and the signals of more important layers are supposed to be transmitted through the higher-quality spatial streams. In the Power Allocation block, we determine a power allocation matrix based on the proposed method which we shall discuss later. In the E-SDM blocks, $\mathbf{W}^{(i)}_{\text{E-SDM}}$ is applied to each user to increase channel capacities. The Block Diagonalization block makes it possible to eliminate inter-user interference, where the MU-MIMO channel can be considered as multiple single-user MIMO (SU-MIMO) channels. In this paper, the number of spatial streams per user is equal to N_L as a prerequisite. The E-SDM blocks make it possible to regard a SU-MIMO channel as N_L parallel spatial streams.

III. A PROPOSED FAIR POWER ALLOCATION

We denote by $C_k^{(i)}$ and $C_{k,ave}$ a channel capacity of the k-th stream of user i and an average channel capacity of the k-th stream over all users, respectively. $C_k^{(i)}$ is expressed as

$$C_k^{(i)} = \log_2 \left(1 + \frac{\lambda_k^{(i)}}{P_z^{(i)}} P_k^{(i)} \right), \tag{2}$$

where $P_z^{(i)}$ is the additive Gaussian noise power for user i, $P_k^{(i)}$ is the power allocated to the k-th stream for user i, and $\lambda_k^{(i)}$ is the k-th largest eigenvalue of $\left(\mathbf{H}^{(i)}\mathbf{W}_{\mathrm{BD}}^{(i)}\right)^H \left(\mathbf{H}^{(i)}\mathbf{W}_{\mathrm{BD}}^{(i)}\right)$. To equalize video quality of all users, the variance of channel capacities of each spatial stream of all users should be zero:

$$\frac{1}{N_U} \sum_{i=1}^{N_U} \left\{ C_k^{(i)} - C_{k,ave} \right\}^2 = 0, \forall k \in \{1, 2, \cdots, N_L\}, \tag{3}$$

where N_U is the number of users. In order to enhance video quality, spatial streams where more significant signals are sent should be higher-quality. Therefore, the following equation should be established:

$$\begin{aligned} & w_1^{(i)} \left(\lambda_1^{(i)} P_1^{(i)} \right) = w_2^{(i)} \left(\lambda_2^{(i)} P_2^{(i)} \right) \\ & = \cdots = w_{N_L}^{(i)} \left(\lambda_{N_L}^{(i)} P_{N_L}^{(i)} \right), \forall i \in \{1, 2, \cdots, N_U\}, \end{aligned} \tag{4}$$

where $w_k^{(i)}$ is the weight value of the k-th stream for user i, $w_1^{(i)} < w_2^{(i)} < \cdots < w_{N_L}^{(i)}$, and $w_k^{(1)} = w_k^{(2)} = \cdots = w_k^{(N_U)}$. Since the total transmission power is limited, the following equation should be established:

$$\sum_{i=1}^{N_U} \sum_{k=1}^{N_L} P_k^{(i)} = P_{tot}, \tag{5}$$

where P_{tot} is the total transmission power. We determine $P_k^{(i)}$ to satisfy (3), (4), and (5), utilizing the fact that (3) means $C_k^{(1)} = C_k^{(2)} = \cdots = C_k^{(N_U)}, \forall k \in \{1, 2, \cdots, N_L\}$.

IV. SIMULATIONS

We use JPEG 2000 to encode and decode 100 color 1920×1080 images for each user. 1 bit per pixel source coding rate is used for every image and frame rate is 24 fps. The images consist of 2 layers. We assume quadrature phase shift keying modulation and i.i.d. Rayleigh fading channel. Convolutional codes whose coding rate is 2/3 are used for forward error correction. We use the path loss model defined by ITU-R [7] to assume indoor surroundings. The distance power loss coefficient is 3.1 and the number of floors between the transmitter and the user is 1. The transmit power is 10 dBm. The transmitter which has 4 antennas sends images to 2 users and each user has 2 antennas. Let the weight values be $\left(w_1^{(1)}, w_2^{(1)}, w_1^{(2)}, w_2^{(2)}\right) = (1, 35, 1, 35)$. The bandwidth used by this system is 80 MHz and the carrier frequency is 5.2 GHz. We make use of a minimum mean square error decoder to estimate the transmitted signals. We assess the quality of the received images with its peak signal-to-noise ratio (PSNR).

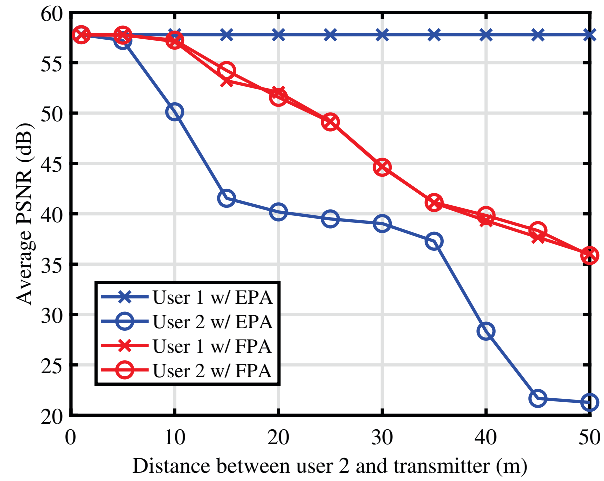


Fig. 2. Average PSNR curves with FPA and UPA schemes.

Fig. 2 shows the average PSNR versus distance between user 2 and the transmitter curves with the proposed FPA scheme and a conventional equal power allocation (EPA) scheme. The distance between user 1 and the transmitter is always one meter in the both schemes. With EPA, the PSNR curve of user 1 keeps a high constant value, while the counterpart of user 2 drops off very rapidly. By contrast, the PSNR curves with FPA of both user 1 and user 2 similarly drop off. In other words, video quality of all users is almost equal even when a user is far from the transmitter. The PSNR difference between user 1 and user 2 with EPA is up to 36 dB, while the counterpart with FPA is up to 1 dB.

V. CONCLUSION

This paper has proposed the power allocation scheme for wireless video transmission over MU-MIMO systems. The proposed power allocation scheme is based on the user fairness in terms of video quality. Simulation results indicate that it is possible to equalize video quality even if the distances between a transmitter and each user are different.

REFERENCES

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[2] Y. H. Yang, S. C. Lin, and H. J. Su, "Multiuser MIMO downlink beamforming design based on group maximum SINR filtering," *IEEE Trans. Sig. Process.*, vol. 59, no. 4, pp. 1746–1758, Apr. 2011.

[3] Y. Yang and H. Yue, "Efficient beamforming method for downlink MU-MIMO broadcast channels," *Int. J. Electronics Commun.*, vol. 69, no. 3, pp. 636–643, Mar. 2015.

[4] K. Tashiro, L. Lanante, M. Kurosaki, and H. Ochi, "Joint transmission and coding scheme for high-resolution video streams over multiuser MIMO-OFDM systems," *IEICE Trans. Fundamentals*, vol. E100-A, no. 11, pp. 2304–2313, Nov. 2017.

[5] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt,"Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Sig. Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[6] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2172–2178, Nov. 2000.

[7] ITU-R Rec. P.1238-4, "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to 100 GHz," Mar. 2005.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Comparative Study of Masking and Mapping Based on Hierarchical Extreme Learning Machine for Speech Enhancement

Ryandhimas E. Zezario [1], Join W. C. Sigalingging[2], Tassadaq Hussain [3], Jia-Ching Wang [2], Yu Tsao [1]

[1]Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan
[2]Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan
[3]Taiwan International Graduate Program in Social Network and Human-Centered Computing, Institute of Information Science Academia Sinica, Taipei, Taiwan

***Abstract*—In this study, we compare the speech enhancement performance based on hierarchical extreme learning machine (HELM) with two distinct strategies: masking and mapping. Experimental results of the perceptual evaluation of speech quality (PESQ) show that for both of limited and sufficient amounts of training data, mapping-based HELM tends to be more effective to improve the performance of speech enhancement.**

I. INTRODUCTION

Speech enhancement (SE) aims to generate speech signals with improved quality and intelligibility by taking noisy speech input and eliminating unwanted noise components. Recently, data-driven based SE approaches have been popularly studied. Among them, deep-learning-based models can perform non-linear transformations in an effective manner and have been confirmed to provide outstanding SE results [1-3]. Deep learning-based SE approaches can be roughly divided into two categories based on the strategies to generate enhanced output: mapping [1, 2] and masking [3]. Although deep learning-based SE approaches (both mapping and masking) can provide high SE results, they require a large computation power for back-propagation training and a large amount of training data.

In our previous work [4], an alternative framework based on the hierarchical extreme learning machine (HELM) model was proposed to address the above two issues. The HELM model is different from other deep learning-based models because it performs estimations much faster and requires a smaller amount of training data to achieve satisfactory SE results. The HELM model consists of a stack of ELM-based sparse auto-encoders. Given the noisy speech as input, the output nodes of the final ELM-based sparse auto-encoder are transformed to generate enhanced speech. In [4], we adopted spectral mapping-based strategy (termed MP-HELM in the following discussion). In this study, we investigate a masking-based HELM (termed MS-HELM) SE approach. We adopted the ideal ratio mask (IRM) [3] as the regression target to train the MS-HELM model. In the experiment, we compared the SE performance of MP-HELM and MS-HELM using the same evaluation set. Experimental results show that MP-HELM outperforms MS-HELM when limited or sufficient training data are available. The remainder of this paper is organized as follows. Section II presents the MS-HELM SE system. Section III reports the experimental setup and results. Finally, the study is concluded in Section IV.

II. THE MS-HELM SE SYSTEM

Figure 1 shows the MS-HELM SE system, which can be divided into two stages: offline and online. It is different from the MP-HELM [4], which uses the clean spectral features as the learning target. The MS-HELM's learning target is IRMs, which is defined as:

$$\boldsymbol{T}_{IRM}(t,f) = \frac{\boldsymbol{S}(t,f)}{[\boldsymbol{S}(t,f)+\boldsymbol{N}(t,f)]}, \tag{1}$$

where $\boldsymbol{S}(t,f)$ and $\boldsymbol{N}(t,f)$ are clean and noise spectral features at time-frequency unit, respectively. Based on the entire set of training data, we can prepare T-pairs of noisy spectral features, $\boldsymbol{Y}(t,f)=\boldsymbol{S}(t,f)+\boldsymbol{N}(t,f)$, and IRM, $\boldsymbol{T}_{IRM}(t,f)$, where T is the total number of training features. Next, the noisy speech samples, $\boldsymbol{Y}(t,f)$, are used to estimate the stack of HELM-based sparse auto-encoders in an unsupervised learning manner. By inputting the entire set of training data, we obtain an output matrix $\boldsymbol{H}$ that is formed by the collective outputs of the final sparse auto-encoder. With $\boldsymbol{H}$, we can estimate the transformation matrix, $\boldsymbol{B}$, which maps the output to the matrix formed by IRMs.

$$\boldsymbol{T}_{IRM} = \boldsymbol{HB}, \tag{2}$$

where $\boldsymbol{T}_{IRM}$ denotes the IRM matrix formed by $\boldsymbol{T}_{IRM}(t,f)$, t=1...T and f=1...F, where F is the feature dimension. The transformation matrix, $\boldsymbol{B}$, can be computed as:

$$\boldsymbol{B} = \boldsymbol{H}^{+}\boldsymbol{T}_{IRM}, \tag{3}$$

where $\boldsymbol{H}^{+}$ indicates the pseudoinverse of $\boldsymbol{H}$.

In the online stage, a noisy speech waveform is first converted to form the matrix of spectral features ($\widehat{\boldsymbol{Y}}$) and then processed by the stack of HELM-based sparse auto-encoders to generate $\widehat{\boldsymbol{H}}$. Based on the estimated $\boldsymbol{B}$ from the offline stage, we can obtain $\widehat{\boldsymbol{T}}_{IRM}$, where

$$\widehat{\boldsymbol{T}}_{IRM} = \widehat{\boldsymbol{H}}\boldsymbol{B}, \tag{4}$$

Finally, we obtain the enhanced spectral features, $\widehat{\boldsymbol{S}}$, by

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

$$\hat{\boldsymbol{S}} = \hat{\boldsymbol{Y}} \otimes \hat{\boldsymbol{T}}_{\boldsymbol{IRM}}, \tag{5}$$

where $\otimes$ is the point wise multiplication. Finally, the enhanced spectral features along with the original phase information of noisy speech were used to reconstruct the enhanced speech waveform.

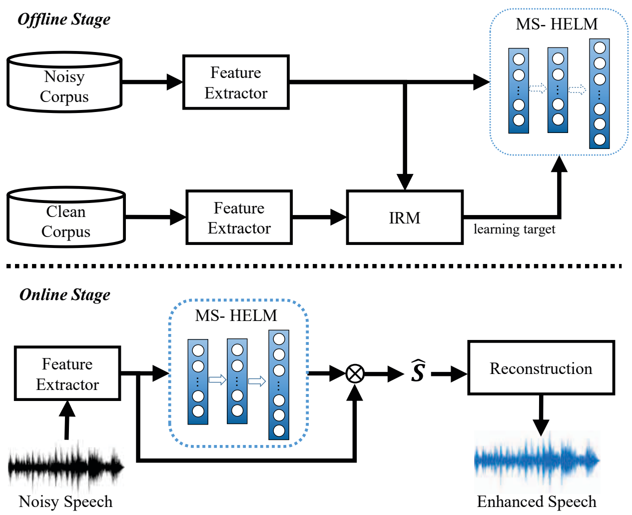


Fig. 1 The MS-HELM SE framework.

III. EXPERIMENTS

A. Experimental setup

This section presents the dataset, configurations of the SE system, and evaluation metric.

1) Speech data preparation

We used the TIMIT [5] dataset to evaluate the HELM-based SE systems. It consisted of 4620 training utterances. We selected the whole training utterances from the original TIMIT dataset as the training data and corrupted them with 90 noise types, at eight SNR levels (from -10 dB to 25 dB with steps of 5 dB). Then, we used three unseen noise types (engine, white, and street) to generate the testing data.

2) SE model setup

We followed the configuration setup used in [4], where [500 500 2000] architecture was used to train both MP-HELM and MS-HELM systems.

3) Objective evaluation metric

Perceptual evaluation of speech quality (PESQ) [6] was used as the evaluation metric to measure the performance of the SE systems. The PESQ score was in the range of 0–4.5. A higher score of PESQ indicates better speech quality of enhanced speech.

Table I. PESQ of MP-HELM and MS-HELM on TIMIT.

Method	**Amount (in %)**	Engine	White	Street
MP-HELM	100	**1.8080**	**1.9848**	**1.9018**
MS-HELM		1.6803	1.7011	1.7443
MP-HELM	50	**1.7958**	**1.9786**	**1.8946**
MS-HELM		1.6761	1.6960	1.7374
MP-HELM	5	**1.7918**	**1.9691**	**1.8815**
MS-HELM		1.6590	1.6946	1.7278

B. Experimental results

Table I lists the PESQ scores of MP-HELM and MS-HELM on TIMIT. To investigate the performance using different amounts of training utterances, we prepared three training sets containing 5%, 50%, and 100% of the original training utterances. From the table, we can see that MP-HELM outperforms MS-HELM at different amounts of training data.

C. Spectrogram analysis

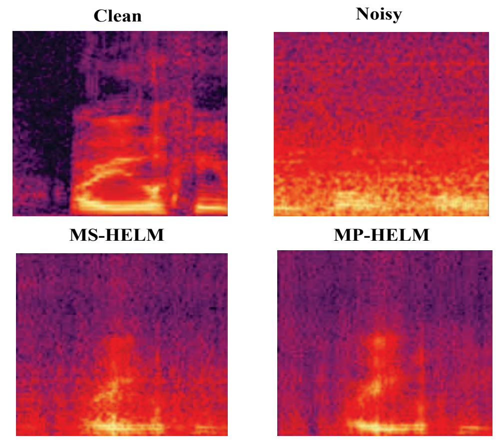


Fig. 2 Spectrograms of a clean utterance, with its noisy and the enhanced versions by MS-HELM and MP-HELM.

Figure 2 shows the spectrograms of a clean utterance with its noisy (engine noise), and the enhanced versions by MS-HELM and MP-HELM to visually analyze the SE results. From the figure, it is observed that both MP-HELM and MS-HELM can effectively reduce the noise components. Next, we note that MP-HELM can generate relatively better noise reduction performance as compared to MS-HELM.

IV. CONCLUSION

This study elaborated comparative analysis between MS-HELM and MP-HELM under unseen noisy environments. Experimental results indicated that MP-HELM tends to achieve better SE results than MS-HELM when limited or sufficient amounts of training data is available.

REFERENCES

[1] X. Lu, Y. Tsao, S. Matsuda, and C. Hori, "Speech enhancement based on deep denoising autoencoder," *in Proc. INTERSPEECH*, pp. 436–440, 2013.

[2] Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee, "A regression approach to speech enhancement based on deep neural networks," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 23, no. 1, pp. 7–9, 2015.

[3] A. Narayanan and D.L. Wang, "Ideal ratio mask estimation using deep neural networks for robust speech recognition," *in Proc. ICASSP*, pp. 7092-7096, 2013.

[4] T. Hussain, S. M. Siniscalchi, C.-C. Lee, S.-S. Wang, Y. Tsao, W.-H. Liao, "Experimental study on extreme learning machine applications for speech enhancement," *IEEE Access,* vol. 5, pp. 25542-25554, 2017.

[5] J. W. Lyons, "DARPA TIMIT acoustic-phonetic continuous speech corpus," *National Institute of Standards and Technology*, 1993.

[6] ITU-T, Rec. P.862, "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs," International Telecommunication Union-Telecommunication Standardisation Sector, 2001.

FRACTAL DIMENSION BASED COLOR TEXTURE ANALYSIS FOR MANGOSTEEN RIPENESS GRADING

Montri Phothisonothai[*], *Suchada Tantisatirapong*[**]
[*] International College, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[**] Faculty of Engineering, Srinakharinwirot University, Nakhon Nayok 26120, Thailand
[**] Corresponding author: suchadat@g.swu.ac.th

ABSTRACT

Mangosteen is one of the fruits that has an enormous export potential in Thailand. However, it contains numerous undesirable external as well as internal conditions which result in the shipment rejection and decrease the reliability of the export. Therefore, in this paper for the first time, we propose the method for mangosteen ripeness grading using the spatiotemporal properties of external rind texture analysis on the basis of fractal dimension (FD) approach for three classes: i.e., Glossy (GS), Medium Rough (MR) and Extreme Rough (ER). In this study, for the first time, the five stages of ripening have been extracted using FD based feature with Gaussian Mixture Model (GMM) classifier. The obtained results showed that the proposed method can perform the better results compared with the classical texture feature, i.e., average accuracy rates of 88.0%, 82.0%, and 90.0% for GS, MR, and ER classes, respectively.

Index Terms—*mangosteen ripeness, fractal dimension, image texture analysis, nondestructive, fruit image*

I. INTRODUCTION

The purple mangosteen (Garcinia mangostana) is a tropical fruit that is mostly grown in tropical rainforests and humid climates of Southeast Asia. For Southeast Asian countries such as Indonesia, Thailand, Malaysia and the Philippines, mangosteen is a primary export commodity. Especially, Thailand's mangosteen is popular all over the world. Mangosteen is also well-known tropical fruits grown as "*Queen of All Fruits*" [1]. At present, mangosteen is considered to be a high-quality product in which Thailand is the country that is the top exporter in the world with an annual market revenue of more than US$64 million [2] for exporting to East Asian countries and Middle Eastern countries. In terms of quality assurance, unlikely other fruits, it contains many undesirable external as well as internal condition which governs acceptance and popularity in the international market level. Inability to provide quality assurance for this fruit leads to shipment rejection and decreases the reliability of export. In addition, the traditional process of mangosteen grading for export purposes requires expertise from gardeners with long experience to visually inspect its rank or class of fruits. These is still a lack of accurate and practical automated inspection technology to assure the irregular qualities. Importantly, the preliminary grading of mangosteen fruit is measured by external factors such as color, shape, skin or rind blemishes which are also very important for consumer's acceptance [3]. Thus, texture analysis for mangosteen ripeness grading becomes a challenging issue in many applications for classification or segmentation using image texture analysis. Machine vision systems for visualizing have been widely used in the agricultural processing industry as well as an automation system. The rind defects inspection process (or diseases appearing) on the fruits are characterized by different texture patterns [4] in which these defects can be qualitatively detected by image processing. The stage of ripening is closely related to the variation of chromatic information or texture pattern which can be identified by physical properties and external appearance. These qualities have resulted in considerable efforts to control mangosteen fruit ripening, so as to achieve optimal fruit maturity for harvesting as well as to devise appropriate postharvest packaging and handling strategies. Understanding the physiological mechanisms underlying ripening processes is necessary to better predict and control ripening processes. For export purpose, it is very important at which stage fruit is harvested. If we are aware of the stage of ripeness of the fruit it is easy to estimate how far distance fruit can be exported to and how long it can be preserved. Mangosteen can be categorized into different ripening stages based on color changes. In our study, for the first time, we investigated ripeness grading of mangosteen with texture analysis based on fractal dimension (FD) according to the five stages of ripening. The proposed method can be described as the following parts:

II. MATERIAL AND PROPOSED SYSTEM

Mangosteen can be classified into 3 classes based on the knowledge of expert farmers, i.e., Glossy (GS)(1), Medium Rough (MR)(2), and Extreme Rough (ER)(3). Postharvest damage in fresh mangosteens at wholesale level in Thailand was investigated from April to October. Total of 37.1% of the production yield was rendered inedible by damage during this period; damages included fruit cracking, hardened rinds, rough surfaces, translucent flesh, gummosis and decay. Dealing with the three classes of grading and the time dependent firmness of unripe and ripe conditions made a challenging issue to understand the complex spatiotemporal prediction of variety ripens using image information analysis.

Mangosteen's rind texture classes in this experiment:
(1) GS: Fine smooth surface with very minimum or no skin bruises and purple fruit. The outer rind of the fruit is thick and rubbery. This class gets more consumer's preference and pricing is highest.
(2) MR: Coarse surface. This class gets lower consumer's preference and lower price than glossy class, but still acceptable for export.
(3) ER: Degree of coarseness is higher than medium rough class. The rind is unpleasant to touch. The fruit gains lower customer acceptance, is mainly sold in local market and is unable to export.

Fig. 1: Progression of five ripening GS stages and growth, $S_1, S_2, \ldots, S_n$, taking approximately 3-5 days for each stage and their most representative chromatic shading information changed gradually from green/red pad to dark purple according to the firmness of unripe to ripe timeline.

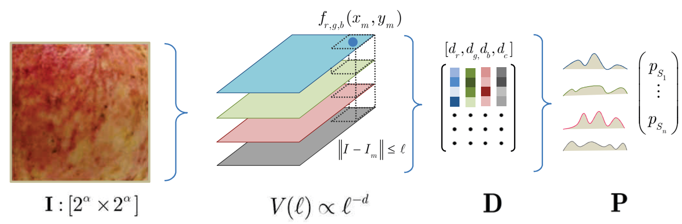


Fig. 2: Fundamental concept of the proposed FD based texture analysis method.

In this paper, we proposed the FD based texture analysis method by using multi FD descriptors for color (RGB cube) and gray-scale images as shown in Fig. 2. The FD feature vector distribution is then obtained to represent the probability based complex spatiotemporal parameter of mangosteen ripeness, $\mathbf{P}$. In order to present the change in terms of complex spatiotemporal pattern, we use the average FD value distributions from gray-scale component combined the RGB components to show the feature patterns [5]. From these results, we found that the distribution of FD changed gradually from stage of ripening versus the class among GS, MR, and ER classes in Fig. 3. However, for the stage S2 and S3 in GS and MR classes, the FD values were reduced due to the color became reddish purple which made more difficult to visually separate the class. The surface roughness classification using Gray Level Co-occurrence Matrix (GLCM) feature is compared with our proposed method. The FD values outperforms the GLCM feature, under training and validating the dataset with GMM classifier, as shown in Table 1.

Table 1: Comparison of average accuracy performed by FD and GLCM feature based on GMM classifier.

Class	FD vs GLCM*				
	S1	S2	S3	S4	S5
GS	95.0% (80.0%)	92.5% (77.5%)	85.0% (72.5%)	82.5% (77.5%)	85.0% (72.5%)
MR	87.5% (77.5%)	82.5% (72.5%)	77.5% (62.5%)	82.5% (67.5%)	80.0% (72.5%)
ER	90.0% (82.5%)	85.0% (75.0%)	87.5% (72.5%)	90.0% (80.0%)	97.5% (77.5%)

* The GLCM results are shown in bracket [6].

III. Experimental Results

To collect the samples, we visited mangosteen orchard at Wang Mai sub-district, Na Yai Am district, Chanthaburi province where the country's eastern region and rich in fruits. For the manually sorted mangosteen samples, an image is acquired using Canon G3X digital camera where a fruit is placed in a uniform background UdioBox-III lighting studio box. Original color image is manually cropped to acquire the four-view centered image of mangosteen fruits. Then, the dataset in the experiment, input image is composed of the region of interest (ROI) in the preprocessing phase: i.e., 24-bit resolution, image dimension of 512×512 pixels, collected 600 samples (40 samples/class/stage). The offline processing of the proposed method was implemented in Matlab software.

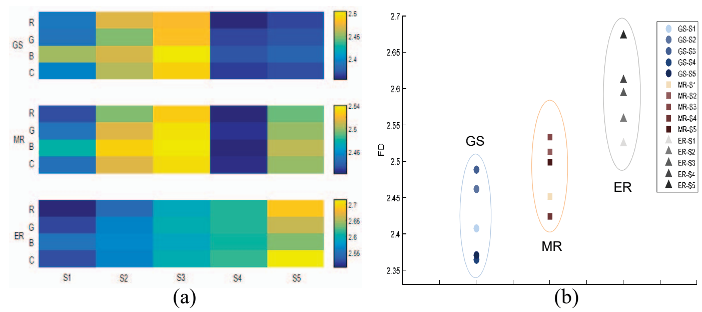


Fig. 3: Distribution of FD changes. (a) Spatiotemporal FD feature and (b) Class versus ripening stage.

As the experiment has been demonstrated, we proposed the fractal dimension analysis for grading the complex spatiotemporal feature for mangosteen ripeness. We found that the change of FD values corresponds to the mangosteen stage with an average of 87% accuracy. The proposed method provides promising results, as the FD can be potentially applied as a feature, allowing the feasibility of automated classification of mangosteen in agribusiness.

REFERENCES

[1] K. Phopin, S. Wanwimolruk, and V. Prachayasittikul, "Food Safety in Thailand. 3: Pesticide residues detected in mangosteen (Garcinia mangostana L.), queen of fruits," *J. Sci. Food. Agric.*, vol.97(3), pp.832–840, 2017.

[2] Thai Trade. "Export of Thaialnd:HS 0804503," reported on June 24, 2015. http://www.ops3.moc.go.th/infor/HS/export/export_commodity/report.asp

[3] K. Phopin, S. Wanwimolruk, and V. Prachayasittikul, "Food safety in Thailand. 3: Pesticide residues detected in mangosteen (Garcinia mangostana L.), queen of fruits," *J. Sci. Food. Agric.*, vol.97, no.3, pp.832– 840, 2017.

[4] D. Unay and B. Gosselin, "Artificial Neural Network-Based Segmentation and Apple Grading by Machine Vision," *IEEE Int. Conf. Img. Proc. (ICIP'05)*, pp.630–633, 2005.

[5] M. Ivanovici and N. Richard, "Fractal Dimension of Color Fractal Images," *IEEE Trans. Image Proc.*, vol.20, no.1, pp.227–235 , 2011.

[6] A. Acharya, M. Phothisonothai, and S. Tantisatirapong, "Surface Roughness Classification of Mangosteen with Gray Level Co-occurrence Matrix based Texture Analysis," *Int. Comp. Sci. & Eng. Conf. (ICSEC 2018)*, Chiang Mai, Thailand (DOI: 10.1109/ICSEC.2018.8712705)

A Study on the Applying of Digital Information and Visually Creative Thinking for Environment Design Education

Chi-Jui Tsai
†Department of Advertising and Strategic Marketing
Ming Chuan University
Taoyuan City, Taiwan
††Department of Industrial Education and Technology
National Changhua Normal University
Changhua City, Taiwan
tsaichi@mail.mcu.edu.tw

Shu-Hsuane Chang
†Department of Industrial Education and Technology
National Changhua Normal University
Changhua City, Taiwan
shc@cc.ncue.edu.tw

*Kuo-An Wang
†Department of Management Information Systems
Central Taiwan University of Science and Technology
Taichung City, Taiwan
††Department of Industrial Education and Technology
National Changhua Normal University
Changhua City, Taiwan
gawang@ctust.edu.tww

***Abstract*—The study discusses the feasibility strategies and creative planning ideas among urban image, digital life and digital art by analyzing the design technique of digital information and visually creative thinking using in the environment. I hope to induce some visual design forms which can be took as the reference on digital art and digital life education or the guide for students' teaching and operation.**

The text is about comprehending the environmental reform and conceptual planning in a city or a certain area that is based on (1)process of creative ideas, (2)construction of visually creative design ideas, (3)technique of visual design, and (4)application of cultural marketing and cultural elements. According to these stages, the researchers can build their own urban culture image for a place, region and city.

***Keywords*—Digital information、Visually Creative Thinking、Environment Design Education**

I. Introduction

From the perspective of the government's policy in urban environment, there are many visual creative plans not only being promoted from the top down, but also reaching a consensus from bottom up to urge the city's change and evolution, allowing residents to understand, feel, enjoy the the achievement and resource of the design. Urban space is like the container accommodating the citizens; hence this study proposes that design is focused on human beings. From three orientations "humanities and cultures", "society" and "nature", professional knowledge and design thinking of visual communication design, product design and interior design implemented as the technical as well as thinking foundation in creative life and integral design. Following are the executive principles in design thinking：(1)Mainly basing on human beings. (2)Importing visual communication design, product design and interior design concepts. (3)Involving life and artistic quality in urban society. (4)Integrating the rationality and the sensibility to creative the pleasure. (5)Bring visually display in space.

Visual design is a kind of cognition element which is the interaction between the designers' visual export and the viewers' import. From the messages and aesthetic experience from visual elements, the designers creative visual communication design through "space", "human" and "vision" concepts to catch viewers' eye and interaction as well as achieve the visual merchandising effect.

II. VISUAL CREATION DESIGN CONCEPT

To creative effect of visually display, combined design with stereography and dynamic, 2D and 3D, can bring the stereoscopic depth of perception in visually. Visual design expresses modelling sense, color sense; Stereoscopic design show sense of shadow and light, spacial perception, design texture; Dynamic design increase attention and pleasant. By the special manifestations of combination of vision and stereography (the design integrated with visual extension and 3D performance), and vision and dynamic (the design of the dynamic vision performed by the raster wall), it breaks through the stereotypes of visual graphic design and static perception, hence expresses the attraction of visual modelling and dynamic display.

III. VISUAL DESIGN TECHNIQUE

Based on the previous visual creation design concept and planning concept, and in order to achieve the design purpose mentioned above, following are synoptical design technique:

(1) Change the visually viewing angle: Ordinarily people used to watch things with eyes parallel or up and down. Consequently, designers apply the visual elements to ceiling, floor, wall, etc. that can change the custom of watching, break through the limit of watching by standing on the ground, and create the visually magic illusion that walls and floors involve another space.

(2) Mirroring and mirroring replace: the generation of mirroring can increase the spacial depth sense. With the visual creation design words, through the objects' transmission of reflection and mirroring, forming more symmetrical relation about attachment and penetration spatially and visually.

(3) Flipped design of swerve and corner: turning and corner is the transition of angle in action. Normally, environmental visual design would not be used at the turning and corner. Therefore, building corner surprise in urban, is not only presenting delightful visual design as well as promoting the application of the corner, could catch people's attention. It is no longer as the corner people pass by, hence the flipped design for the visual custom.

† 1st *Affiliation* †† 2nd *Affiliation* **Corresponding author*

(4) Reverse and differentiation design of crossing and penetration: in a way of jumping off the spatial reality, constructing design penetrating from walls and ground, reverse and differentiation break through people's cognitive rules and visual-Inertial in ordinary life experience. It brings the visual impact, impels psychological reaction and reach the surprising visual effect by producing variation, visual phenomenon and spatial illusion against the normality.

(5) Design of visual advertisement marketing: applying and integrating visual design on metro handle as a promotion scheme of advertisement marketing.

IV. CULTURE MARKETING AND CULTURE ELEMENTS DESIGN

Many cultural assets, which is deeply rooted in people's life and concept and presents distinctive Taiwan eastern culture style, have gone through the accumulation of time and humanities. By means of the identification with culture and emotional establishment, it initiates the symptom about Taiwan's image. Consequently, designers should find the design about culture elements and import to the visual creative planning from the dimensions of culture, and market with culture.

(1) Creative installation design in metro

Plant image displays the image of grow-up and fluency, which is a symbol of cultural power and integrity. Therefore, Taiwan eastern culture and urban night scenery can be presented by using the image of pine, plum blossom, bamboo, orchid and calligraphy as well as projecting lighting design about the city above the entrance of the station.

(2) Creative design and production of viaduct pier

In Taiwanese society culture and belief, there are many objects that can be the metaphor applied in design. For example, red eggs, New Year scroll, incense stick, etc. can be regarded as some kind of red; mountain and plants in Taiwan can be green; and river and ocean around the island can be blue.

In other words, visual creativity and styling of Taiwanese red can be found from Taiwanese culture, society, folk custom and belief; Taiwanese green can be found from Taiwanese or Taipei's natural environment; Taiwanese blue can be found from Taiwanese or Taipei's geographic location.

To design the viaduct pier via image and night lightning with Taiwanese red (culture, society, folk custom and belief), Taiwanese green (forest), and Taiwanese blue (ocean), designers can search for lots of appropriate subject matters from Taiwan environments and society.

(3) Visual promotion of outdoor space creativity

In form of space, outdoor walls and fences are relations of plastic covering and coating. Consequently, designers present the concept design and visual creativity by the figure of hands which can symbolize that people guard the environment.

(4) Communication software stickers

Cooperating with illustrators and create a series of communication software stickers about Taiwanese image or make dolls placed in a key position in DM.

(5) Cooperation with enterprise publicity resources

The enterprise like Chunghwa Telecom takes media and crowds as the promoting target. Cooperating with the enterprise could increase the promotion of the campaign and city image that can be operated in the static location such as the entrance of convenient store, night market and cinema, or in moving advertising medium such as taxi and bus.

Convenient store is fixed point propaganda. Propaganda in a point can construct a path of propaganda that can assemble as an area, hence increase opportunities and exposures to catch public's attention. Taxi and bus are the dynamic display that can form as the pathway network propaganda. The entrance of the night market and cinema can attract young people and foreigners to pay more attention on the visual messages and propaganda of the urban image and culture.

(6) Image registration

Cooperating with schools and make the micro cinema as the campaign video which can act in concert with the city's marketing group.

V. DISCUSSION AND CONCLUSION

Using good information vision and good visual creativity can effectively enhance the aesthetics of environmental design. Improve the aesthetic quality and aesthetic education of urban residents. The goal of this research is to enhance the image of the city through design techniques and the application of cultural elements of the city, and to promote the participation of urban residents in the form of design marketing. Therefore, in this year's proposal to participate in University Social Responsibility Projects, through the practice of visual design and application of cultural elements, citizens can establish their own cultural symbols and visual images of the city.

REFERENCES

[1] .Fukuda, S. (2005). Shigeo Fukuda Master Works. New York : Firefly Books.

[2] Block, J. R. (1992). Can you believe your eyes? :over 250 illusions and other visual oddities. Bristol, PA: Brunner/Mazel.

[3] Gopnik, Alison & Andrea Rosati. (2001). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. Developmental Science, 4 (2), 175-183.

[4] Arnheim, Rudolf., 1974, Art and visual perception: A Psychology of the Creative Eye, Berkeley, University of California Press.

[5] Jenny Preece ., 1998, A Guide to Usability Human Factors in Computing, U.S.A.,Basic Books, Inc.

[6] Wickens, Christopher D., 1992, Engineering Psychology and Human Performance, Harper Collins Publishers, Inc.

[7] Neisser,U., 1967, Cognitive Psychology, New York: Appleton-Century-Crofts.

[8] Neisser,U., 1967, Selective reading: A method for the study of visual attention, Presented at the Nineteenth International Congress of Psychology, London.

A New Variation of Singular Value Decomposition

Heri Prasetyo
Department of Informatics
Universitas Sebelas Maret
Surakarta, Indonesia
heri.prasetyo@staff.uns.ac.id

Didi Rosiyadi
Reseach Centerfor Informatics
Lembaga Ilmu Pengetahuan Indonesia
Bandung, Indonesia
didi.rosiyadi@lipi.go.id

Iwan Setiawan
Reseach Centerfor Informatics
Lembaga Ilmu Pengetahuan Indonesia
Bandung, Indonesia
iwan017@lipi.go.id

***Abstract*— This paper presents a new variation of Singular Value Decomposition (SVD) computation for image compression. It modifies the SVD operation with two additional pre-processing steps, i.e. Pixel Interleaving and Linearization process. An input image is firstly scrambled by Pixel Interleaving operation. This interleaved image is further linearized before applying SVD calculation. As documented in the Experimental Results Section, this simple approach yields a promising result in the image compression task compared to that of the classical SVD-based technique.**

***Keywords*—** *image compression, pixel interleaving, SVD*

I. Introduction

Let I be a grayscale image of size $M \times N$. This image is real-valued matrix, i.e. $I \in \mathbb{R}^{M\times N}$. The SVD operation decomposes this matrix into the following form:

$$I \Rightarrow U\Sigma V^T, \tag{1}$$

where $U \in \mathbb{R}^{M\times r}$ and $V \in \mathbb{R}^{N\times r}$ are two unitary matrices, referred as left and right singular vectors, respectively. $\Sigma \in \mathbb{R}^{r\times r}$ is singular value matrix consisting non-zero entries along its diagonal, i.e. $\Sigma_r = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_r\}$ with condition $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$. The symbol r denotes the rank of matrix I. The SVD owns some interesting properties making it suitable for several applications in the image processing and computer vision tasks such as image watermarking [1, 2], image denoising [3, 4], noise level estimation [5], etc. The SVD yields promising results on those aforementioned tasks [1-5].

The SVD has been proven to give a good performance on image compression system. It effectively performs data compaction under the maximum principal component setting. Specifically, the SVD-based image compression is defined as follow:

$$\hat{I} \Leftarrow \sum_{i=1}^{t} \sigma_i\, u_i v_i^T, \tag{2}$$

where t denotes the number of truncated or selected singular values. This value is commonly set with an arbitrary value smaller that the rank of matrix I, i.e. $t \leq r$. The symbols σ_i, u_i , and v_i are the singular value, left, and right singular vectors, respectively, at the i-th column. T represents the matrix transpose operator. By selecting an appropriate t value, the reconstructed image $\hat{I}$ is visually identical to that of the original image I. It implies that $\hat{I} \approx I$.

II. Proposed SVD-Based Image Compression

In SVD-based image compression, the value of t determines the quality of decompressed image. Since this value is fix or user predetermined, the only way to improve the performance of SVD-based image compression is by modifying the SVD computation itself. Herein, we modify the SVD computation by utilizing the pixel interleaving of an input image. Suppose that I be an input image of size $M \times M$. For simplicity, we only consider squared matrix of input image. The pixel interleaving uniquely scrambles the image I into another matrix $\tilde{I}$ as follow:

$$\mathcal{P}\{I\} \Rightarrow \tilde{I}, \tag{3}$$

where $\mathcal{P}\{\cdot\}$ denotes the pixel interleaving operator with specified (selected) image block of size $m \times m$. The choice of this image block size should satisfy the condition $M = m^2$. Fig. 1 shows the two grayscale images, each of size 256×256, while Fig. 2 depicts the pixel interleaving results. Each block of interleaved image contains some information about an input image I. De-correlating these image blocks affects on modifying the SVD computation.

The scrambled image $\tilde{I}$ is further linearized by concatenating all entries of each block into the column vector as follow:

$$X \Leftarrow [x_1, x_2, \dots, x_i, \dots, x_M], \tag{4}$$

where X and x_i are linearized version of scrambled image $\tilde{I}$ and the i-th column vector of X. Subsequently, the SVD decomposes the matrix X into the following form:

$$X \Rightarrow U_X \Sigma_X V_X^T, \tag{5}$$

where U_X, Σ_X, and V_X are three generated SVD matrices.

The image compression can be further performed by using the following operation:

$$\hat{X} \Leftarrow \sum_{i=1}^{t} \sigma_{Xi}\, u_{Xi} v_{Xi}^T, \tag{6}$$

where $\hat{X}$ denotes the decompressed matrix X by selecting the value of t with condition $t \leq r$. It should be noted that the matrix $\hat{X}$ is still in linearized version and pixel interleaved form. An additional step is executed to obtain correct decompressed image. This operation is defined as follow:

$$\hat{I} \Leftarrow \mathcal{P}^{-1}\{\hat{X}\}, \tag{7}$$

where $\mathcal{P}^{-1}\{\cdot\}$ denotes the operator of inverse linearization followed by the inverse of pixel interleaving. The matrix $\hat{I}$ is decompressed image obtained from the proposed method.

III. Experimental Results

This section reports some extensive experiments on the SVD-based image compression. Firstly, we evaluate the performance in terms of visual investigation. This evaluation only considers the quality of decompressed image based on visual inspection. Subsequently, we examine the quality of decompressed image based on the value of objective image quality assessments.

Two images as shown in Fig. 1 are turned as testing images. In the visual evaluation, we compare the performance of proposed method with the classical approach on SVD-based image compression. We simply set the number of truncated singular value as $t = 5, 10, 15, 20$. Fig. 3 shows the comparisons between the classical and proposed SVD-based image compression over various t values, from the first to the last rows are with $t = 5, 10, 15, 20$. As it can be seen from this figure, the proposed method offers better quality on the decompressed image compared to that of the classical SVD-based image compression under visual observation.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Two image quality assessments, i.e. Peak-Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM), are considered for objective comparison between the proposed method and classical approach. Higher value of these two metrics indicates better performance. In this experiment, we evaluate the image compression with $t = 1,2,\ldots,100$. Figs. 4 and 5 displays the performance comparisons between the proposed method and classical SVD-based image compression over various t values. It clearly reveals that the proposed method is better than the classical approach indicated with higher PSNR and SSIM scores.

REFERENCES

[1] J. M. Guo, and H. Prasetyo, “False-positive-free SVD-based image watermarking,” *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 1149-1163, Jul. 2014.

[2] J. M. Guo, D. Riyono, and H. Prasetyo, “Hyperchaos permutation on false-positive-free SVD-Based image watermarking,” *Multimed. Tools and Appl.*, 2018. https://doi.org/10.1007/s11042-018-6767-x.

[3] Q. Guo, C. Zhang, Y. Zhang, and H. Liu, “An efficient SVD-based method for image denoising,” *IEEE Trans. Circuits, Syst., Video Technol.*, vol. 26, no. 5, pp. 868-880, 2016.

[4] Y. M. Huang, H. Y. Yan, Y. W. Wen, and X. Yang, “Rank minimization with applications to image noise removal,” Information *Sciences*, vol. 429, pp. 147-163, 2018.

[5] W. Liu, and W. Lin, “Additive white Gaussian noise level estimation in SVD Domain for Images,” *IEEE Trans.* Image *Process.*, vol. 22, no. 3, pp. 872-883, 2013.

(a) (b)

Fig. 1. Two grayscale images used in this experiment: (a) Image #1, and (b) Image #2.

(a) (b)

Fig. 2. Examples of pixel interleaving results over image block size: (a) $\frac{M}{2} \times \frac{M}{2}$, and (b) $\frac{M}{4} \times \frac{M}{4}$.

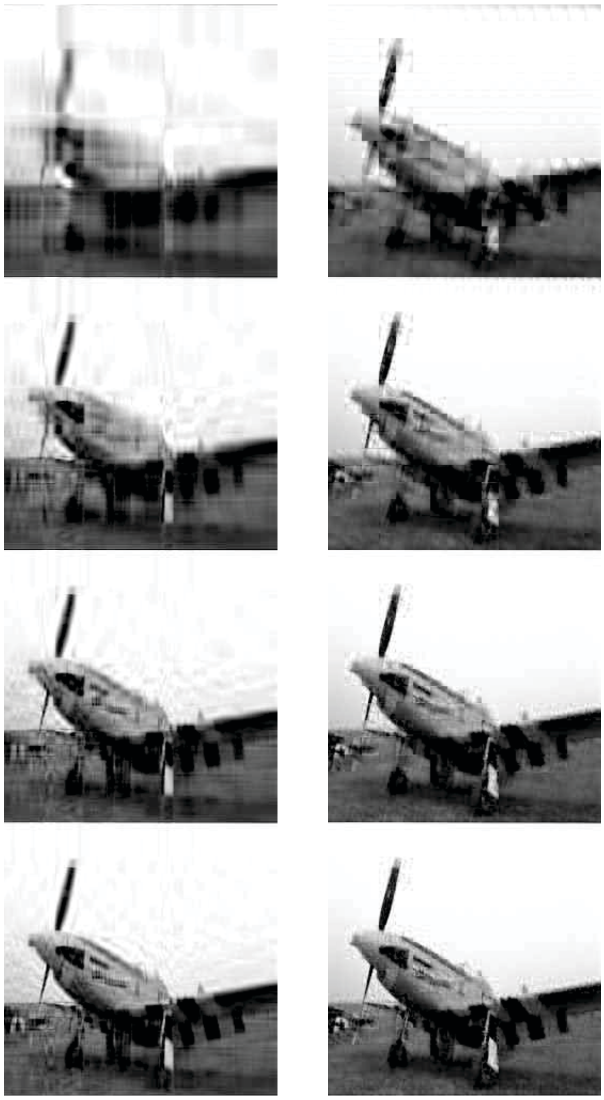
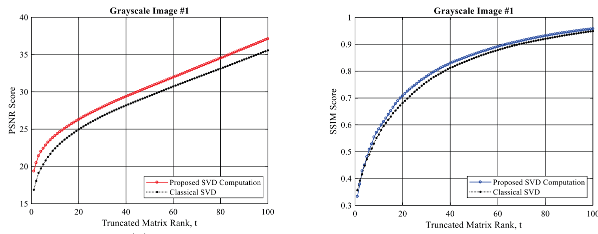
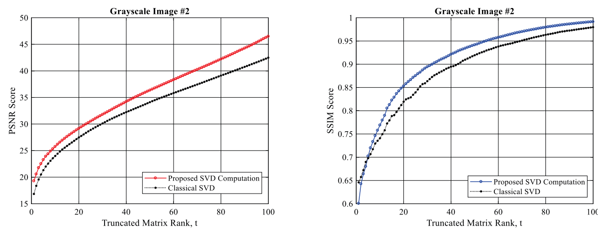


Fig. 3. Image compression results using: (left column) classical SVD-based image compression, and (b) proposed SVD-based image compression.



(a) (b)

Fig. 4. Performance evaluations on Image #1 over: (a) PSNR and (b) SSIM score.



(a) (b)

Fig. 5. Performance comparisons on Image #2 under: (a) PSNR and (b) SSIM value.

Thermal-Based Pedestrian Detection Using Faster R-CNN and Region Decomposition Branch

Yung-Yao Chen
Graduate Institute of Automation Technology
National Taipei University of Technology
Taipei, Taiwan
yungyaochen@mail.ntut.edu.tw

Sin-Ye Jhong
Graduate Institute of Automation Technology
National Taipei University of Technology
Taipei, Taiwan
a888999a100@gmail.com

Guan-Yi Li
Graduate Institute of Automation Technology
National Taipei University of Technology
Taipei, Taiwan
gwspotex@gmail.com

Ping-Han Chen
Graduate Institute of Automation Technology
National Taipei University of Technology
Taipei, Taiwan
t107618006@ntut.edu.tw

***Abstract*—In this paper, we present an infrared thermal-based pedestrian detection method that can be applied in nighttime intelligent surveillance systems. Pedestrian detection plays an important role in computer vision and automation industry applications, which include video surveillance, automotive robot, and smart vehicles. Recently, the improvement in deep learning techniques, such as convolutional neural networks (CNNs), have significantly increased the accuracy of pedestrian detection. Normally, the optical cameras, e.g. charge-coupled device cameras, are the device used to capture images. However, considering the dark environments and the luminance variation issues, infrared thermal camera would be an effective alternative solution to nighttime pedestrian detection. On the other hand, occlusion is one of the commonest problems, which makes nighttime pedestrian detection more challenging. To address the abovementioned problems, this work presents a pedestrian detection framework which consists of Faster R-CNN and a region decomposition branch. The proposed region decomposition branch allows us to detect wider range of the pedestrian appearances including partial body poses and occlusions. From the experimental results, this work demonstrates better detection accuracy than the currently developed CNN-based detection method because of combining the multi-region features.**

Keywords—pedestrian detection, infrared thermal imaging, region decomposition branch

I. INTRODUCTION

The goal of object detection is to find all the instances of one or more classes of objects given an image [1, 2]. Among object detection problems, pedestrian detection is one of the most essential issues because it has widespread applications. Nighttime pedestrian detection brings another level of challenges due to dramatical luminance variations [3, 4]. Although compared to the charge-coupled device (CCD) imaging, infrared imaging provides better capture ability under the low-illumination environments, the occlusion problem still effects the detection accuracy.

In recent years, the great progress of object detection have been made by combining the region proposal algorithms and CNNs. One of the most notable works is the Region CNN (R-CNN) framework [5], which consists of generating the object region proposals using the selective search scheme [6], extracting CNN features of the candidate regions, and classify them with class-specific support vector machines (SVMs). Then, the Fast R-CNN method [7] improves the R-CNN speed using feature sharing and region of interest (ROI) pooling. In [8], the faster R-CNN method further increases the computational efficiency by integrating the external region proposal modules into the CNN for boosting both training and detection. Moreover, the detection accuracy in [8] can be also enhanced by joint learning of region proposal and classification modules.

II. PROPOSED METHOD AND THE PRELIMINARY RESULTS

Recently, researchers have realized that the global appearance of the entire object area may not be sufficient to accurately classify and locate objects. In [9], Ouyang and Wang proposed a joint deep learning pedestrian detection method, which considers the mutual relationship among individual components to decrease the average miss rate of the detection results. In [10], it is suggested that the detection of the occluded object can be missed due to the lack of details of some parts of the occluded object. CNN features discriminability area to demonstrate the benefits of learning features from the variable deformation characteristics of different body parts. To address the limitations of the previous methods, in this paper, we want to utilize the regional part-model detector using CNN and enhance the relationship between the object regions to increase the detection accuracy.

Figure 1 shows the overall framework of the proposed method. The proposed framework is implemented based on the Fast R-CNN method that the acquired image is resized and goes through a CNN-based extractor such as ResNet. The RPN and the ROI align stages are used to extract the features of ROI box proposals and solve the problem of mis-alignment. In the final stages, we replace the conventional fully connected layers by two separate branches: a full body branch and a region decomposition branch (including a head layer, a body trunk layer, and a foot layer). Therefore, the proposed region decomposition branch can integrate the features from different regions step-by-step and infer the essential cues as the regions with strong feature response.

Research was supported in part by the National Science Council under Grants MOST 108-2221-E-027 -095 -MY2.

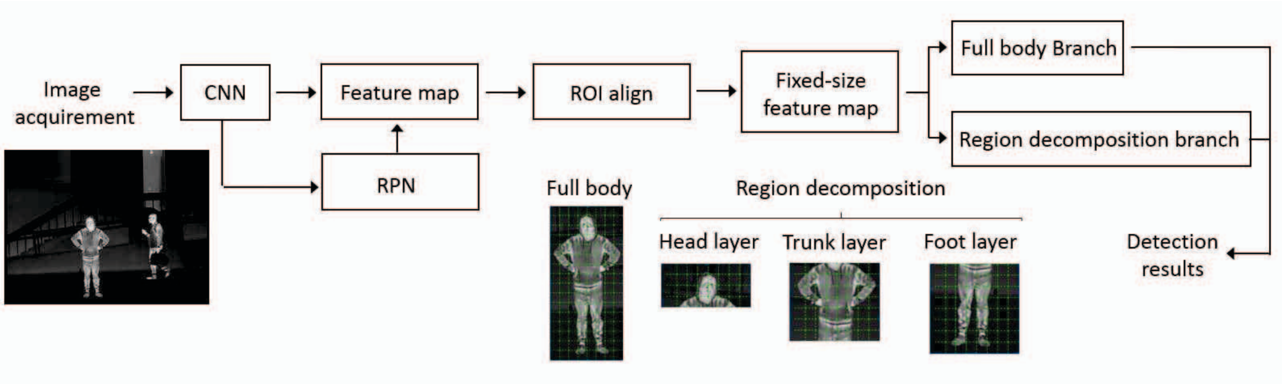
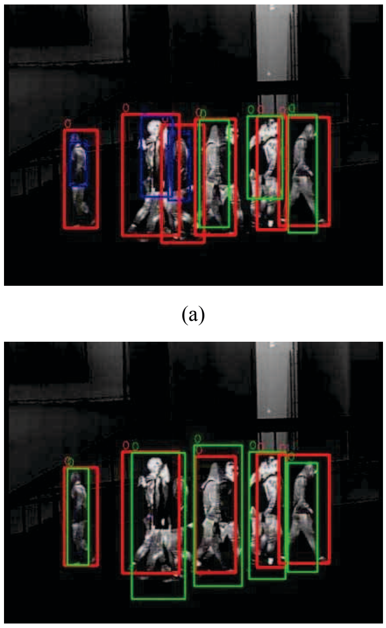


Fig. 1. Overall framework of the proposed method.

Typically, the results of RPN is supposed to contain the entire body of a pedestrian. Whereas, in the region decomposition branch, the region proposal is decomposed into three layers: a head layer, a body trunk layer, and a foot layer. It is based on our observation that in the infrared thermal images, these three regions are most obvious for a pedestrian. For example, the head of a pedestrian in a thermal image is usually lighter than other regions because of the body temperature (Normally, the average body temperature of a human being is ranging from 36.1 Celsius degree to 37.2 Celsius degree, which is higher than most background at nighttime). However, when it is raining, the umbrella usually covers the head region, and therefore, the feature of the head would be lost. Without region decomposition branch, a detector which only contains full body branch might lead to a lower detection rate in a rainy day.

The underneath motivation of the proposed method is to mainly detect the body trunk and generate the semantic relations between trunk and other parts such as head and legs by a region decomposition branch. In this study, we propose a novel thermal-based system for nighttime pedestrian detection, which exploits not only the global appearance features of an entire object region, but also the individual region decomposition appearance features. The proposed method takes the multi-region based appearance model into account and jointly describe the global and partial appearance of the object at the same time. From the preliminary experiments shown in Fig. 2, it proves that the strategy of the proposed method is effective to handle the occlusion problem.



(a)

(b)

Fig. 2. Preliminary experimental results. (a) Detection result of the faster R-CNN method [8]. (b) Detection result of the proposed method. The red, green and blue rectangles indicate ground truth, correct and false result, respectively.

III. CONCLUSIONS

In this work, we present a CNN-based detector with a region decomposition branch using thermal imaging to solve the occlusion problem of nighttime pedestrian detection. We utilize a whole object region and the multiple small regions, and learn high-level semantic features by combining a holistic and part model features stage by stage. The experimental results validate the effectiveness of this work.

REFERENCES

[1] C. Qi, W. Liu, C. Wu, H. Su, and L. Guibas, "Frustum pointnets for 3D object detection from RGB-D data," *in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 918–927, June 2018.

[2] Z. Cai and N. Vasconcelos, "Cascade R-CNN: delving into high quality object detection," *in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 6154–6162, June 2018.

[3] J. Li, F. Zhang, L. Wei, T. Yang, and Z. Lu, "Nighttime foreground pedestrian detection based on three-dimensional voxel surface model," *Sensors*, Vol. 17, pp. 2354, Oct. 2017.

[4] J. Baek, S. Hong, J. Kim, and E. Kim, "Efficient pedestrian detection at nighttime using a thermal camera," *Sensors*, Vol. 17, pp. 1850, Aug. 2017.

[5] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 580–587, Sep. 2014.

[6] K. Sande, J. Uijlings, T. Gevers, and A. Smeulders, "Selective search for object recognition," *in Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, pp. 1879–1886, Jan. 2012.

[7] R. Girshick, "Fast R-CNN," *in Proc. IEEE Int. Conf. Computer Vision (ICCV)*, pp. 1440-1448, Feb. 2016.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 1137–1149, June 2016.

[9] W. Ouyang and X. Wang, "Joint deep learning for pedestrian detection," *in Proc. IEEE Int. Conf. Computer Vision*, pp. 2056–2063, Dec. 2013.

[10] W. Ouyang, H. Zhou, H. Li, Q. Li, J. Yan, and X. Wang, "Jointly learning deep features, deformable parts, occlusion and classification for pedestrian detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 40, pp. 1874–1887, Aug. 2017.

H-BTC Database: A Brief Review on Halftone based Block Truncation Coding (H-BTC) Images

Jing-Ming Guo, *Senior Member, IEEE*, and Sankarasrinivasan S
Dept. of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan
jmguo@seed.net.tw, D10507805@mail.ntust.edu.tw

***Abstract*— Halftone based block truncation coding images (H-BTC) are the improved version of BTC images which can offer superior representation and enhanced image quality. The application of the H-BTC images range from image compression and retrieval, indexing, reconstruction, classification and so on. In this paper, a brief introduction to various H-BTC images is presented and their evaluation is carried out in terms of image quality and computational demand. Further, the application prospective of the H-BTC images are also provided. The developed H-BTC database comprises of 30K images comprising of the five types of H-BTC images in which three categories are constructed based on digital halftoning and two types are based on digital multitoning. The database would be very useful to carry out deep learning research and various image processing tasks. The database along with its source code will be made open source for the research and academic purpose.**

Keywords: Digital halftoning, block truncation coding image, image database, image compression

I. Introduction

Block truncation coding (BTC) [1] is a very popular choice of image compression method for its simplistic approach and less computational demand. The technique performs the compression in the block-wise manner and the original image block is replaced in terms of its high and low mean value of the block. The image quality of the BTC usually reduce with increasing block size and suffer from serious blocking artifacts. On the other hand, the halftone based BTC achieves significantly better image quality and suffer from less blocking artifacts. The digital halftoning [5] is performed predominantly using four methods such as ordered dithering (OD), error diffusion (ED) [2], dot diffusion (DD) [3] and direct binary search (DBS) algorithm. Among them, the first three methods are extended to BTC to obtain ODBTC, EDBTC and DDBTC images. In latest, the digital multitoning is getting prominence which is the upgraded version of halftone. With the previous relevance, the digital multitoning version is also extended to BTC to obtain the multitone BTC images (MTBTC) [4]. The MTBTC images have better image quality than the halftone BTC but suffers from low compression ratio.

II. H-BTC Database

The database consists BTC images along with its five halftone BTC versions such as ODBTC, EDBTC, DDBTC, MTBTC-DD and MTBTC-DD. For each class, 5K images are provided and the database comprise of 30K images of size 512x512. The database images are constructed for the block size of 16x16. The brief explanation of various halftone based BTC is provided below.

A. ODBTC

The ODTC image is constructed based on the ordered dithering method of halftones. To begin with, the original image is processed in a block wise manner and the maximum and minimum term corresponding to a block is computed. Based on the difference, a specific dithering array is picked from the look up table and the block is thresholded using that dithering screen as shown in Fig. 1.

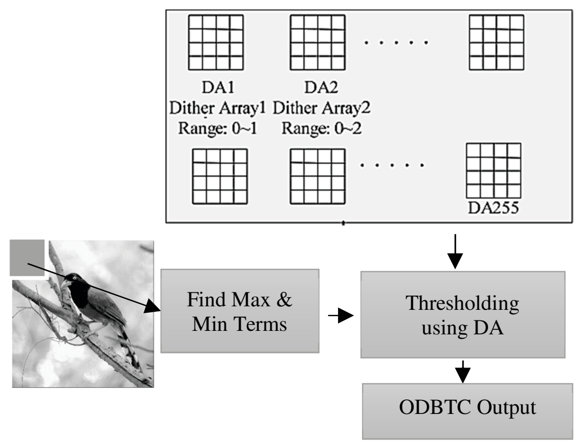


Fig. 1. ODBTC Scheme

Usually, the dithering screen pertaining to the blue noise spectrum are used as it results in dispersed halftone patterns. Finally, the 1's and 0's in the halftone binary image is assigned with maximum and minimum term.

B. EDBTC

The error diffusion method is an improved version of ordered dithering and it involves processing the image in a sequence manner. Initially the gray scale image of range 0-255 is thresholded using its mid value 128. The difference between the output and the actual image is computed and it is further distributed to the neighborhood pixels. The error kernels for distributing the difference is shown in Fig. 2.

	(i, j)	7
3	5	1

Fig. 2. Floyd-Steinberg

Once the halftone image is generated, it is replaced with the maximum and minimum term of the corresponding block.

C. DDBTC

The dot diffusion is a combination of ordered dithering and error diffusion technique. In contrast to error diffusion, the dot diffusion comprises of two matrices such as class and diffusion matrix. The class matrix dictates the order in which the pixels are processed, and the difference is distributed to the neighborhood pixels. Among different class matrix configurations, the Knuth dot diffusion kernel is used in this work. The class matrix adopted in this work is shown in Fig. 3, the class matrix is of size 8x8 with diffuse matrix of size 3x3.

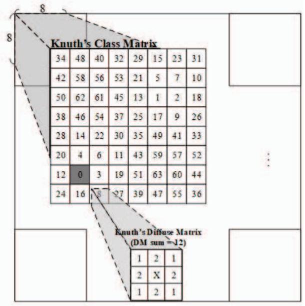


Fig. 3. Class Matrix of Dot-Diffusion

D. MTBTC

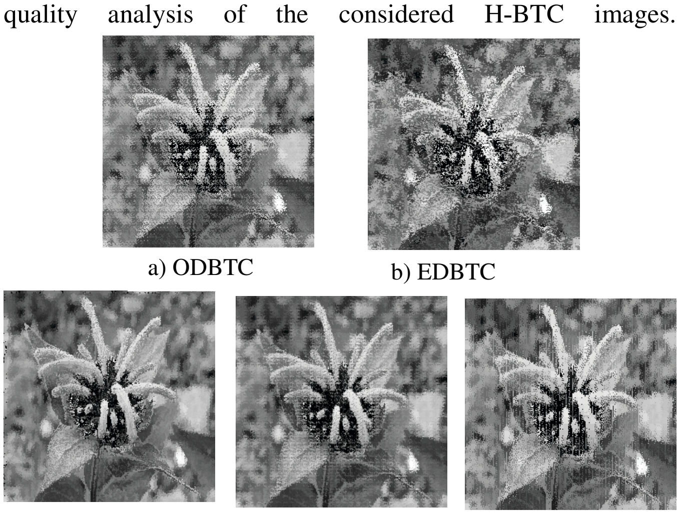
The MTBTC is an upgradation of the ODBTC technique and offers a better image quality than the former one. In comparison to ODBTC, the MTBTC use more dithering screens and their corresponding output is assigned with more number of tones other than maximum and minimum. The generation of 4-tone MTBTC image is shown in Eq. 1.

$$MT_4 = \begin{cases} \max & if\ im\{m,n\} > DA^{k,1} + \min; \\ \max - 0.33 * k & if\ DA^{k,1} + \min \geq im[m,n] > DA^{k,2} + \min; \\ \max - 0.66 * k & if\ DA^{k,2} + \min \geq im[m,n] > DA^{k,3} + \min; \\ \min & otherwise. \end{cases}$$

Where MT_4 refers to the multitone BTC image comprise of four tones such as maximum, minimum and its intermediate values. $DA^{k,1}$ ... $DA^{k,3}$ refers to the dithering screens.

III. H-BTC Analysis

In the section, the H-BTC images are compared in terms of the image quality, computational demand and their application prospective is also presented. Table I shows the result of image quality analysis of the considered H-BTC images.



a) ODBTC b) EDBTC

c) DDBTC d) MTBTC-DD e) MTBTC-CD

Fig. 4. H-BTC Image from Database

Table 1 shows the image quality assessment of the various halftone BTC images. The MTBTC-DD image offer high image quality as it has more number of tones than the other halftone BTC versions. Among the other BTC versions, DDBTC is found to obtain the better image quality.

Table I. Comparison of Image Quality for Various H-BTC Versions

Block Size	MTBTC -DD	ODBTC	EDBTC	DDBTC	BTC
8	**0.970**	0.9267	0.9347	0.936	0.941
16	**0.966**	0.8674	0.8730	0.8815	0.851
32	**0.934**	0.784	0.7880	0.7916	0.731
64	**0.889**	0.717	0.7020	0.7232	0.683

Table II. Computational Demand of the Different Methods

Type	Addition/ Subtraction/ Comparison	Multiplication/ Division	Square Root
BTC	2 x (M x N)+3	M x N + 9	2
ODBTC	M x N	0	0
EDBTC	M x N + 1	(M x N) x 6	0
DDBTC	M x N + 1	(M x N) x 6	
MTBTC	(T-1) M x N	0	0

Table II shows the computational demand of the various H-BTC techniques. In general, the ordered dithering based ODBTC and MTBTC are found to have lowest computation as it involves only a comparison, and moreover, ED, DD and BTC have slightly more computations.

Some application prospects of this database is provided below,

- **Image Compression:** The primary intention of H-BTC images is to obtain the compressed image format. The H-BTC can achieve a compression ratio up to 8.
- **Image Retrieval**: ODBTC, EDBTC and DDBTC images features are found to be very effective to perform image retrieval tasks. In recent works, this image features are combined with deep learning features to perform image retrieval with improved accuracy.
- **Image Forensics:** The color and bit features of the EDBTC images is used to detect copy-move forgery image regions.
- **Others:** Image Indexing, Watermarking, Reconstruction and classification tasks.

IV. Conclusion

In this work, a H-BTC database is generated and is made open-source for research and academic purpose. The database comprises of 30K images and would be an ideal to perform deep learning and image processing tasks. A brief summary of the H-BTC generation, its analysis and application prospects are also provided for better understanding.

Code: https://github.com/SankarSrin

References

[1] J. M. Guo, and M. F. Wu, "Improved block truncation coding based on the void-and-cluster dithering approach," *IEEE Transactions on Image Processing,* vol. 18, no. 1, pp. 211-3, Jan. 2009.

[2] J. M. Guo, "Improved block truncation coding using modified error diffusion," *Electronics Letters*, vol. 44, no. 7, pp. 462-464, Mar. 2008.

[3] J. M. Guo, and Y. F. Liu, "Improved block truncation coding using optimized dot diffusion," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1269-1275, Mar. 2014

[4] J. M. Guo and S. Sankarasrinivasan, "Enhanced block truncation coding image using digital multitone screen," *APSIPA ASC*, pp. 672-676, 2017.

[5] Ulichney, R. *Digital halftoning*. MIT press, 1987.

Function Block-Based Robust Firmware Update Technique for Additional Flash-Area/Energy-Consumption Overhead Reduction

Jisu Kwon, Jeonghun Cho, and Daejin Park*
School of Electronics Engineering, Kyungpook National University *boltanut@knu.ac.kr

***Abstract*—Energy consumption and flash-memory usage are very limited in microcontrollers that make up the sensor network, accordingly, the process of updating the embedded firmware should also be low cost and energy efficient. This work proposes a technique that overcomes limitations due to increased costs of configuring the sensor network by additional memory usages and increased energy consumption resulting from firmware updates. Instead of dealing with the whole firmware, we split the firmware into function blocks and managed them with a function map that indicate each function block address. Further, by only updating function blocks where differences exist, we successfully reduced flash memory usage and energy consumption that occurred during the firmware update process. We implemented the proposed technique with the target measurement environment, and the result shows that maximum flash memory usage reduced by 91.4% and that 71.4% reduction in execution time resulted in a 69% reduction in energy consumption over the conventional method.**

Keywords–firmware update, microcontroller, function block, flash memory, energy consumption

I. Introduction

With the recent explosive growth of Internet-of-Thing (IoT), demand for a microcontroller unit (MCU), which plays a key role in the edge nodes that make up the sensor network, is also growing rapidly. Due to the characteristics of MCU that repeat only a certain action once the code is stored in a flash memory, users must update embedded software, or firmware, to change MCUs' hardware behavior. Since available resources in the embedded systems are constrained compared to the PC or server, energy consumption or memory usage that occurs during firmware updates is becoming more important [1]. Therefore, diverse studies have been conducted from various perspectives to reduce the energy consumption [2] and overhead generated by the firmware update process [3]. Unlike previous studies focused on the differences between the two firmware [4], this approach concentrated on the energy consumption and flash memory usage that occurred during firmware updates. The key idea of the proposed technique is to first divide firmware into function blocks (FBs) and to only replace FBs differences exist with a function map that stores the FB addresses.

II. Proposed Architecture and Implementation

Fig, 1 summarized the conventional firmware update process, which is represented with a flash memory map for each step. The purpose of the update is to exchange FW_old to FW_new. To update old firmware, we need to store new firmware on flash memory space, because the update may not be able to finish for several reasons. When the new firmware is stored, we can erase the old firmware from the flash memory. Assuming that the size of the new firmware does not increase compared to the existing firmware, we can copy the pre-stored new firmware into the space created after erasing the existing firmware. Finally, erasing the pre-stored new firmware will complete the whole update process. This update process requires additional flash memory space for storing firmware and consumes much more energy for frequent flash writing and erasing.

The firmware structure proposed in this paper consists of a partitioned FB and function map pointing to each FB start address. Fig. 2 illustrates the firmware difference-only update processing when upgrading from FW_old to FW_new. Before the update, we need to analyze the difference between old and new firmware and recognize that it belongs to which FB. Only FBs with a difference will be exchanged from the previous firmware. Among the $f1$, $f2$, and $f3$ that make up the FW_old, we detected the difference between FW_old and the latest version of the firmware in FB $f1$; thus, exchange FB $f1$ to revised FB $f4$. After storing FB $f4$ on the flash memory space, we reassigned the function map FB $f1$ pointer to the updated FB $f4$. We finished the whole update process by erasing FB $f1$. Different from the conventional method, the proposed technique required additional flash memory space only for a specific FB and also reduced energy consumption from the flash memory writing and erasing.

III. Evaluation Results

Fig. 3 shows the evaluation environment of the proposed technique. We used STM32F4DISCOVERY which embeds STM32F407VGT6 MCU as a target board and measures energy consumption and execution time with an Atmel energy debugger. We updated the firmware and sent the command from the computer to the bootloader via USB-CDC communication. Bootloader analyzes commands, and determines what operations to complete, such as programming firmware on flash memory. In the case of the conventional firmware update process, we used 140KB firmware and updated it. As shown in Fig. 4 (a), this process spent 8.821s of execution time and consumed 3.315J amount of energy. To implement the technique proposed in this paper, we assumed that the changed part of the firmware belongs to the FB of 12KB, and we made a section for a function map that points to each FB. We changed the address of the FB pointed out in the function map after flashing the new FB to the flash memory. In contrast with the conventional method, Fig. 4 (b) shows the result of the proposed technique, which spent 2.521s of execution time and consumed 1.025J of energy. As a result, not only was the execution time reduced by 71.4% but there was also a 69% reduction in energy consumption and memory usage, as the additional required memory space only needed the size of the FB being changed.

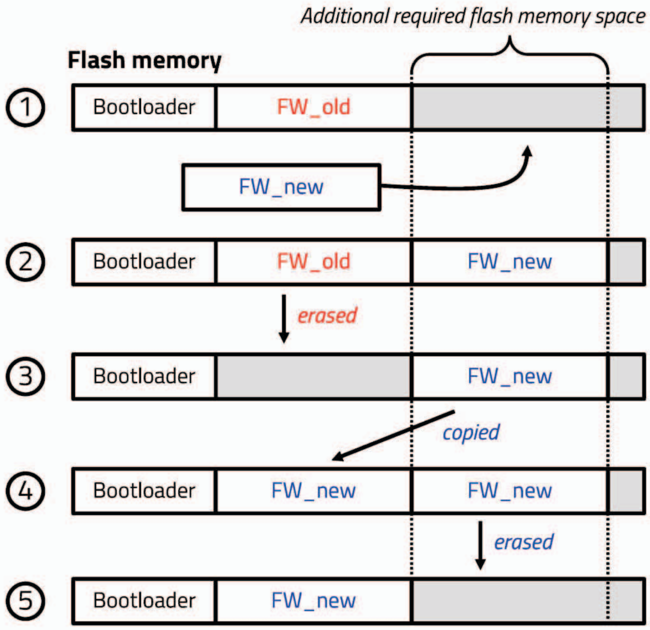


Fig. 1: Conventional firmware update process

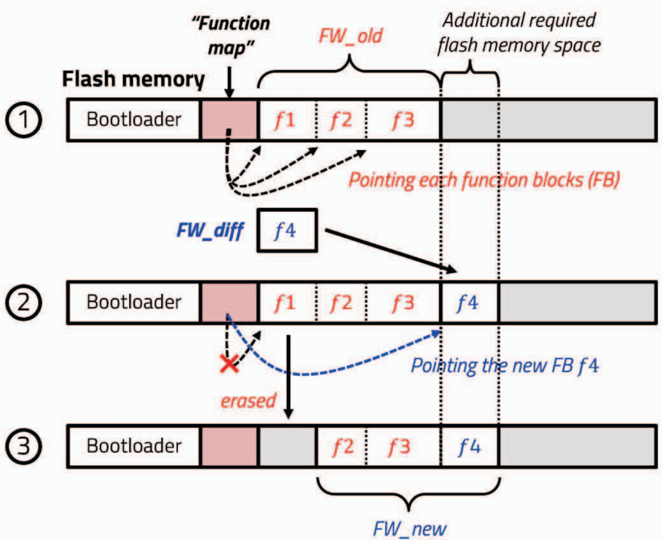


Fig. 2: Proposed FB-based robust firmware update process

IV. Conclusion

In this paper, we proposed a firmware update technique that improves energy consumption and flash memory usage during the update process by separating the firmware into individual FBs and by using the function map to manage their addresses. Because the proposed firmware update technique requires only maximum FB-sized memory space and takes less execution time, it is space and energy efficient. The proposed technique was successfully implemented on the target board based on MCU, and the result showed that the amount of flash memory required and energy consumption were reduced compared to the conventional method.

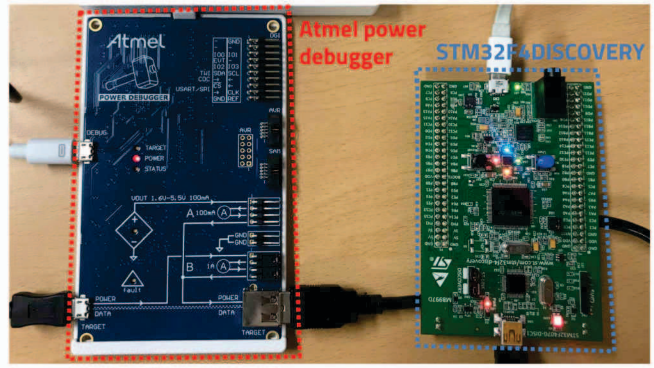
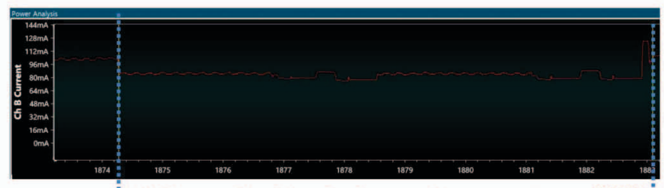
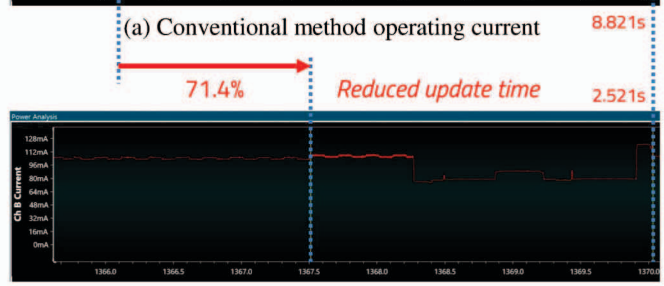


Fig. 3: Measurement environment for energy consumption and flash memory area



(a) Conventional method operating current



(b) Proposed method operating current

Fig. 4: Evaluation results (operating current and execution time)

References

[1] S. Rein and M. Reisslein, "Low-memory wavelet transforms for wireless sensor networks: A tutorial," IEEE Communications Surveys Tutorials, vol. 13, no. 2, Second 2011, pp. 291–307.

[2] H. Jayakumar, K. Lee, W. S. Lee, A. Raha, Y. Kim, and V. Raghunathan, "Powering the internet of things," in Proceedings of the 2014 International Symposium on Low Power Electronics and Design, ser. ISLPED '14. New York, NY, USA: ACM, 2014, pp. 375–380. [Online]. Available: http://doi.acm.org/10.1145/2627369.2631644

[3] W. Dong, Y. Liu, C. Chen, J. Bu, C. Huang, and Z. Zhao, "R2: Incremental reprogramming using relocatable code in networked embedded systems," IEEE Transactions on Computers, vol. 62, no. 9, Sep. 2013, pp. 1837–1849.

[4] O. Kachman and M. Balaz, "Optimized differencing algorithm for firmware updates of low-power devices," in 2016 IEEE 19th International Symposium on Design and Diagnostics of Electronic Circuits Systems (DDECS), April 2016, pp. 1–4.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2019R1A2C2005099), and Ministry of Education (NRF-2018R1A6A1A03025109).

Adaptive Step-Size Recursive Least Biphase Errors Algorithm

Shin'ichi Koike
Consultant
Tokyo, Japan

***Abstract*—In this paper, Adaptive Step-Size Recursive Least Biphase Errors Algorithm (ARBEA) is proposed, employing a new "biphase function." Results of analysis and experiments demonstrate its effectiveness in making adaptive filters fast convergent and robust against impulsive observation noise.**

***Index Terms*—biphase function; recursive least estimation; adaptive step size; impulse noise.**

I. Introduction

In adaptive filtering systems, we often find impulse noise in the observation noise contained in the error signal [1]. To make adaptive filters robust against impulsive observation noise, we propose use of "biphase function" which takes two phase values [2]. To accelerate the filter convergence, we introduce a recursive least estimate of the *inverse* covariance matrix of the regressor, and an adaptive step-size control method, yielding Adaptive Step-Size Recursive Least Biphase Errors Algorithm (ARBEA) for adaptive filters in the complex-number domain.

In this paper, we develop statistical analysis of the ARBEA for calculating theoretical filter convergence. Numerical experiments with some examples are carried out to compare simulation results and theory to evaluate the filter performance and to examine the accuracy of the analysis.

II. Impulse Noise Model, Biphase Function and Biphase Error Algorithm

A. Impulse Noise Model

Impulsive observation noise is modeled as Contaminated Gaussian Noise (CGN) [3] which is a stochastic combination of background noise with variance σ^2_{vb}, and impulse noise with variance σ^2_{vi} and probability of occurrence p_{vi}.

B. Biphase Function and Biphase Error Algorithm

We propose a "biphase function" $\mathrm{biph}(z)$ having two phase points $z = \exp(j\pi/4)$ and $z = \exp(j5\pi/4)$. For $z = x + j\,y$, we can define $\mathrm{biph}(z) = \mathrm{sgn}(x+y)\exp(j\pi/4)$ where $\mathrm{sgn}(\cdot)$ is the signum function [2]. See Fig. 1. Using the biphase function, we derive an update equation for the tap weights $\mathbf{c}(n)$ (N taps) for Biphase Error Algorithm (BiPhEA) as given by

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \alpha_c\,\mathrm{biph}[e^*(n)]\,\mathbf{a}(n), \tag{1}$$

where $e(n)$ is the error, $\mathbf{a}(n)$ is the regressor vector, and α_c is the step size. $\mathrm{biph}[e^*(n)]$ successfully suppresses impulse noise.

III. Adaptive Step-Size Recursive Least Biphase Errors Algorithm (ARBEA)

A. Formulation

For the proposed ARBEA, the tap weight update equation is given by

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \alpha_c(n)\,\mathbf{P}(n)\,\mathrm{biph}[e^*(n)]\,\mathbf{a}(n), \tag{2}$$

where $\mathbf{P}(n)$ is a recursive least estimate of the inverse covariance matrix of the regressor calculated by

$$\mathbf{P}(n+1) = \lambda^{-1}\,\mathbf{P}(n)\,\{\,\mathbf{I} - \mathbf{a}(n)\,\mathbf{a}^H(n)\,\mathbf{P}(n) \,/\,[\lambda\,|\,e_R(n) - e_I(n)\,|^{1/2} + \mathbf{a}^H(n)\,\mathbf{P}(n)\,\mathbf{a}(n)]\,\}, \tag{3}$$

with $e(n) = e_R(n) + j\,e_I(n)$ and the forgetting factor λ.

The adaptive step size $\alpha_c(n)$ is calculated by

$$\alpha_c(n) = \beta_c\,s_e(n) \tag{4}$$

and $s_e(n+1) = \mu^{-1}s_e(n)\{1 - s_e(n)/[\mu\,|\,e_R(n) - e_I(n)\,|^{1/2} + s_e(n)]\}$, (5)

with β_c being a scaling factor and μ a leakage factor.

B. Statistical Analysis - Summary

Assume that the resressor $\mathbf{a}(n)$ is a stationary Gaussian process with covariance matrix $\mathbf{R_a} = E[\mathbf{a}(n)\mathbf{a}^H(n)]/2$. For the mean $\mathbf{m}(n) = E[\boldsymbol{\theta}(n)]$ and the second-order moment $\mathbf{K}(n) = E[\boldsymbol{\theta}(n)\,\boldsymbol{\theta}^H(n)]$ of tap weight misalignment vector $\boldsymbol{\theta}(n) = \mathbf{h} - \mathbf{c}(n)$ ($\mathbf{h}$ is unknown system response), we derive from (2) a set of difference equations: $\mathbf{m}(n+1) = \mathbf{m}(n) - E[\alpha_c(n)]\,\mathbf{p}(n)$ and $\mathbf{K}(n+1) = \mathbf{K}(n) - E[\alpha_c(n)]\,[\mathbf{V}(n) + \mathbf{V}^H(n)] + E[\alpha_c(n)]^2\,\mathbf{T}(n)$, where $\mathbf{p}(n) = \boldsymbol{\Pi}(n)\,\mathbf{W}(n)\,\mathbf{m}(n)$, $\mathbf{V}(n) = \boldsymbol{\Pi}(n)\,\mathbf{W}(n)\,\mathbf{K}(n)$, $\mathbf{T}(n) \cong \boldsymbol{\Pi}(n)\,\mathbf{S}\,\boldsymbol{\Pi}(n)$, $\boldsymbol{\Pi}(n) = E[\mathbf{P}(n)]$, $\mathbf{W}(n) = (2/\pi)^{1/2}/\sigma_e(n)\cdot\mathbf{R_a}$, $\mathbf{S} = 2\,\mathbf{R_a}$, $\sigma^2_e(n) = \varepsilon(n) + \sigma^2_v$ is the error variance, σ^2_v is the noise variance, and $\varepsilon(n) = \mathrm{tr}[\mathbf{R_a}\mathbf{K}(n)]$ is Excess Mean Square Error (EMSE). For the CGN, we use σ^2_{vb} and σ^2_{vi}, and averaging.

The expectation of recursive least estimate $\boldsymbol{\Pi}(n)$ is updated as $\boldsymbol{\Pi}(n+1) = \lambda^{-1}\,\boldsymbol{\Pi}(n)\,[\mathbf{I} - \boldsymbol{\Phi}_\mathbf{P}(n)\,\boldsymbol{\Pi}(n)]$, where $\boldsymbol{\Phi}_\mathbf{P}(n) = \int_0^\infty J[u^2/2\cdot\lambda\,\sigma^{1/2}_{DP}(u,n);\,1/2]\cdot 2\,\mathbf{D_P}(u,n)\,|\,\mathbf{A_P}(u,n)\,|^{-1}\,u\,du$, $\mathbf{A_P}(u,n) = \mathbf{I} + u^2\,\mathbf{R_a}\,\boldsymbol{\Pi}(n)$, $\mathbf{D_P}(u,n) = \mathbf{A_P}^{-1}(u,n)\,\mathbf{R_a}$, $\sigma^2_{DP}(u,n) = \mathrm{tr}[\mathbf{D_P}(u,n)\mathbf{K}(n)] + \sigma^2_v$, and a function $J(x;\,1/2) = (2/\pi)^{1/2}\int_0^\infty \exp[-(x\,t^{1/2} + t^2/2)]\,dt$ [4].

For the adaptive step size, we find $E[\alpha_c(n)] = \beta_c\,E[s_e(n)]$, $E[s_e(n+1)] = \mu^{-1}E[s_e(n)]\{1 - \phi_s(n)\,E[s_e(n)]\}$, and $\phi_s(n) = \int_0^\infty\{\mu\,\sigma^{1/2}_e(n)\,t^{1/2} + E[s_e(n)]\}^{-1}\exp(-t^2/2)\,dt$.

IV. Experiments - Numerical Results

In this section, we present results of experiments for the ARBEA, where theoretically calculated filter convergence is compared with simulation results for the following examples. In Monte Carlo simulations, we compute an ensemble average of squared excess error $<|\epsilon(n)|^2>/2$, where we run filter convergence 1000 times using the tap weight update equation.

<u>*Example #1*</u> number of taps: $N = 4$; regressor: AR1 Gaussian process with variance $\sigma^2_a = 1$ (0dB) & regression coefficient $\eta = 0.5$; for ARBEA: $\lambda_c = 1 - \lambda = 2^{-7}$, $\beta_c = 2^8$ & $\mu_c = 1 - \mu = 2^{-8}$; for BiPhEA: $\alpha_c = 2^{-12}$; Case 1: background noise only $\sigma^2_{vb} = 0.01$ (-20 dB); Case 2: CGN $\sigma^2_{vb} = 0.01$, $\sigma^2_{vi} = 10$ (+10 dB) & $p_{vi} = 0.1$.

<u>*Example #2*</u> $N = 32$; regressor: AR1 Gaussian process with $\sigma^2_a = 1$ (0dB) & $\eta = 0.9$; for ARBEA: $\lambda_c = 2^{-10}$, $\beta_c = 2^5$ & $\mu_c = 2^{-5}$; for BiPhEA: $\alpha_c = 2^{-12}$; Case 1: background noise only $\sigma^2_{vb} = 1$ (0 dB); Case 2: CGN $\sigma^2_{vb} = 1$, $\sigma^2_{vi} = 100$ (+20 dB) & $p_{vi} = 0.1$.

In *Example #1*, the number of taps N is small, the regressor is moderately correlated, the noise variance is small in Case 1, and the step size is selected small. Fig. 2 provides results of simulations and theoretical calculations of filter learning curves for Cases 1 and 2 of *Example #1* for the ARBEA. In the figure, we also plot and draw simulated and theoretical convergence for the BiPhEA. It is observed that the ARBEA makes the adaptive filter convergence much faster than the BiPhEA, while it preserves the robustness against the impulsive observation noise.

In Fig. 3, results are shown for *Example #2*, where N is large, the regressor is highly correlated, the noise variance is large, and the step size is small. We again observe that the ARBEA is effective in making the adaptive filter fast convergent and robust in the presence of impulse noise.

In Figs. 2 and 3, we see good match between simulated and theoretical convergence curves that shows the validity of the analysis summarized in Subsection III.*B*.

V. Conclusion

In this paper, we have first reviewed the BiPhEA that employs the proposed biphase function for adaptive filters in the complex-number domain. Then, the ARBEA has been proposed, introducing recursive least estimation of the inverse covariance matrix of the regressor and a stable adaptive step-size control method.

Through statistical analysis and numerical experiments with some examples, we have demonstrated that the ARBEA successfully accelerates the filter convergence speed, and at the same time protects the filter from impulse noise. Comparison of the theoretically calculated results with those of simulations exhibits good agreement that shows the accuracy of the analysis.

Analysis for a non-stationary system is left for further study. Performance comparison with other fast convergent and robust filtering algorithms including the one in [2] is also left as a future work.

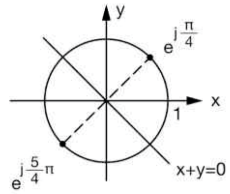


Fig. 1. Sketch of biph(z).

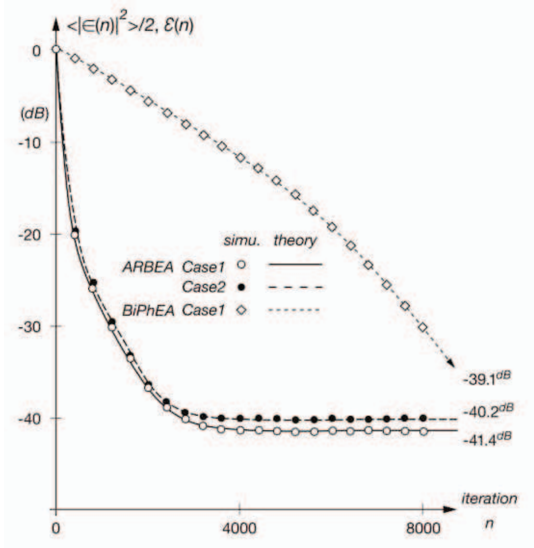


Fig. 2. Filter learning curve (*Example #1*, $N = 4$, ARBEA & BiPhEA).

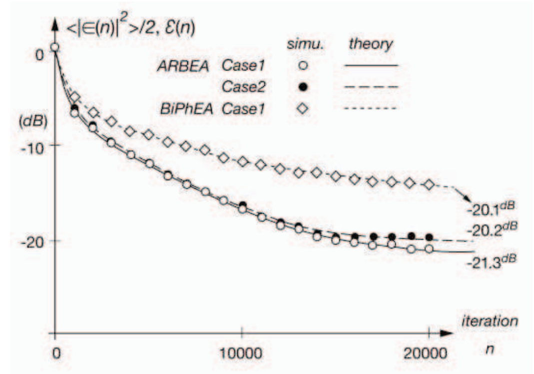


Fig. 3. Filter learning curve (*Example #2*, $N = 32$, ARBEA & BiPhEA)

References

[1] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Select. Area Commun.*, vol. 11, pp. 991-1001, Sept. 1993.

[2] S. Koike, "Adaptive step-size biphase error Newton algorithm," *in Proc. IEEE ISPACS 2018*, Ishigaki Isl., Okinawa, Japan, Nov. 2018, pp. 341-345.

[3] S. C. Bang and S. Ann, "A robust adaptive algorithm and its performance analysis with contaminated-Gaussian noise," *in Proc. IEEE ISPACS 94*, Seoul, Korea, Oct. 1994, pp. 295-300.

[4] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. ASSP*, vol. 34, no. 8, pp. 793-806, Aug. 1986.

Proposal of BSS method to separate the respiratory sound and the heart sound

Yuki Kubota, Minoru Komatsu, Hiroki Matsumoto
$m1671019,\ komatsu,\ matsumoto@maebashi-it.ac.jp$
Maebashi Institute of Technology 460–1 Kamisadori, Maebashi, Gumma, 371–0816 Japan

I. INTRODUCTION

Conventionally, there is a method to use auscultation when medical doctors diagnose a disease from heartbeats with heart noise. However, the respiratory sound and the heart sound are mixed. Further, it is difficult for the desired heart sound to be heard. For solving this problem, we propose a method to separate respiratory sound and heart sound based on blind source separation (BSS). In BSS method, we use a non-linear function corresponding to the probability distribution of source signals in order to separate. A distribution of the respiratory sound is a high-order sub-Gaussian distribution close to uniform distribution, a distribution of the heart sound is a super Gaussian distribution. In the conventional method, there is no non-linear function corresponding to both distribution. Therefore, in this paper, in order to separate the respiratory sound and the heart sound, we propose new non-linear function accurately separated from super Gaussian distribution through sub-Gaussian distribution to uniform distribution. Then, we build the BSS method based on the proposed non-linear function. We evaluate the proposed method by computer simulation.

II. PROPOSED METHOD

The method of Ref. [1] and the method of Ref. [2] are known as methods with a relatively wide range of corresponding distribution in the conventional methods. The method of Ref. [1] corresponds from low-order sub-Gaussian distribution through high-order sub-Gaussian distribution to uniform distribution. The method of Ref. [2] corresponds from super Gaussian distribution to low-order sub-Gaussian distribution. However, both methods do not correspond to everything from super Gaussian distribution through sub-Gaussian distribution to uniform distribution. From that reasons above, these methods are insufficient separation precision for respiratory sound and heart sound. Therefore, in this section, we propose a method to appropriately correspond from super Gaussian distribution to uniform distribution by giving an optimal non-linear function. Fig.1 shows a model of the BSS.

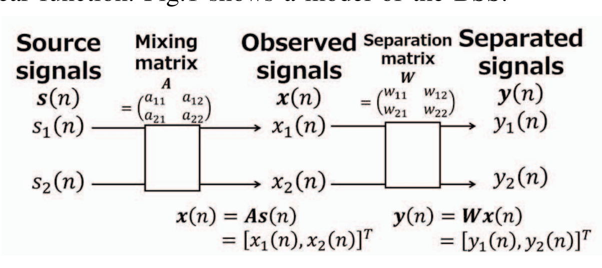


Fig. 1. A model of the BSS

For the model of Fig.1, the following shows the procedure of the proposed algorithm.
1.Estimate the kurtosis $\boldsymbol{K_y}(n)$ of the separated signal.
2.From the kurtosis, determine non-linear function $\varphi(\boldsymbol{y}(n))$.
3.Update separation matrix $\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \eta[\boldsymbol{I} - \varphi(\boldsymbol{y}(n))(\boldsymbol{y}(n))^T]\boldsymbol{W}(n)$.
4.Go to 1.
First, we describe the kurtosis. The kurtosis $\boldsymbol{K_y}(n)$ is calculated by

$$K_{y_i}(n) = \frac{\sigma^4(n)}{(\sigma^2(n))^2} = \frac{(1-\lambda)\sigma^4(n-1)+\lambda(y_i(n)-\overline{y_i}(n))^4}{((1-\lambda)\sigma^2(n-1)+\lambda(y_i(n)-\overline{y_i}(n))^2)^2} \quad (1)$$

,where $\overline{\boldsymbol{y}}(n)$ is the average of $\boldsymbol{y}(n)$. $\sigma^4(n)$ is the central moment of order 4th, $\sigma^2(n)$ is variance.λ $(0 < \lambda \ll 1)$ is a forgetting factor, and i=1 or 2. Next, the proposed non-linear function vector $\varphi(\boldsymbol{y}(n))$ is given by

$$\varphi(\boldsymbol{y}(n)) = \varphi_1(\boldsymbol{y}(n)) + \varphi_2(\boldsymbol{y}(n)) \quad . \quad (2)$$

In equation(2), based on both conventional methods, the non-linear functions are as follows:

$$\varphi_1(\boldsymbol{y}(n)) = \quad [a_1 \tanh(y_1(n)), a_2 \tanh(y_2(n))]^T \quad (3)$$

$$\varphi_2(\boldsymbol{y}(n)) = \quad \Big[(1-a_1)\frac{q_1 y_1(n)|y_1(n)|^{q_1-2}}{4^2}, \quad (1-a_2)\frac{q_2 y_2(n)|y_2(n)|^{q_2-2}}{4^2}\Big]^T . \quad (4)$$

Fig.2 shows the relationship between parameters a_i and q_i in the non-linear function of the proposed method. Fig.3 shows graph of proposed and conventional non-linear function. From Fig.2 and Fig.3, the proposed non-linear function can correspond to everything from super Gaussian distribution through sub-Gaussian distribution to uniform distribution.

III. COMPUTER SIMULATION AND CONSIDERATION

First, we show a example in the case of separation according to the proposed method. The simulation conditions are as follows. Sampling frequency is 8[kHz](The number of samples m is 16000), loop is 10 times. Fig.4 shows source signals. Fig.5 shows observed signals. Fig.6 shows separated signals. Next, we compare the proposed method with the conventional method. The comparison is performed by the similarity sim of the separated signal to the source signal. sim is given by

$$sim = \frac{1}{1+d} \quad , \quad (5)$$

where, d is the euclidean distance,

$$d = \sqrt{\sum_{n=0}^{m-1}(s_i(n) - y_i(n))^2} \quad (i = 1, 2). \quad (6)$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

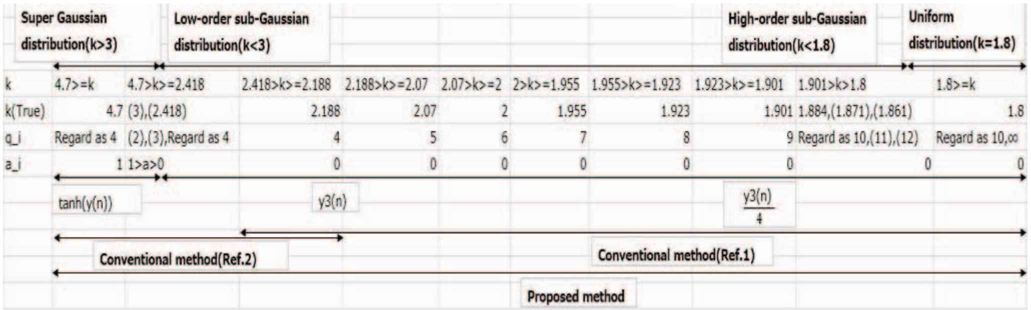


Fig. 2. The relationship between parameters a_i and q_i

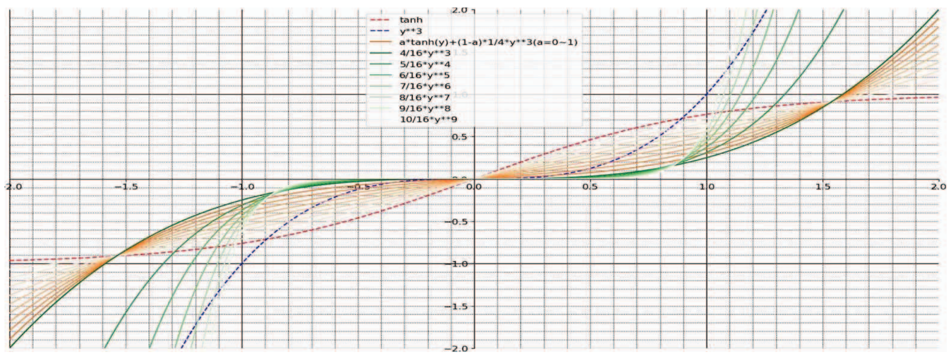


Fig. 3. Graph of proposed and conventional non-linear function

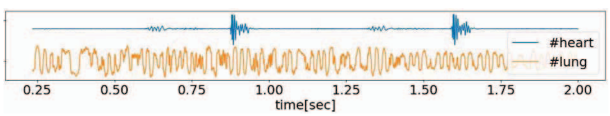


Fig. 4. Source signals

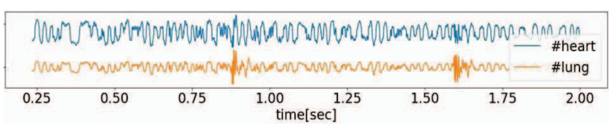


Fig. 5. Observed signals

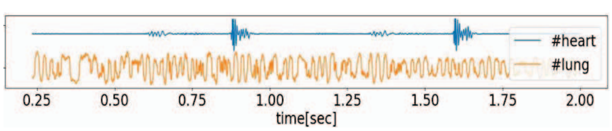


Fig. 6. Separated signals

TABLE I
THE RESULTS OF SIMILARITY sim

$non-linear\ function$	heart	lung
$Proposed\ method$	0.522	0.823
$Ref.$[1]	0.502	0.454
$Ref.$[2]	0.475	0.467
$y(n)^3$	0.313	0.235
$\tanh(y(n))$	0.059	0.072

Table I shows the results of similarity sim. From Table I, the separation precision is higher than the conventional method. Because it is thought that proposed non-linear function is given appropriately correspond to the distributions of both source signals.

IV. Summary

In this paper, we proposed the BSS method that appropriately corresponds to the probability density distributions of source signals. Since the separation precision is higher than the conventional method, it is suitable for separation etc of respiratory sound and heart sound in auscultation sound. The future work is to consider for practical use.

References

[1] Shinya KUSAKARI, Hiroki MATSUMOTO, Nobuo OTANI, Toshihiro FURUKAWA, "A Study on Blind Source Separation Method Using Kurtosis" *Journal of Signal Processing*, Vol. 9, No. 5, pp.397-407, Sept., 2005

[2] Kenji NAKAYAMA, Akihiro HIRANO, Takayuki SAKAI, "An Adaptive Nonlinear Function Contorolled by Kurtosis for Blind Source Separation" *IJCNN'2002, Honolulu, Hawaii*, pp.1234 - 1239, May., 2002

Hammerstein Spline Adaptive Filtering based on Normalised Least Mean Square Algorithm

Sethakarn Prongnuch* Suchada Sitjongsataporn† Theerayod Wiangtong‡
*The Electrical Engineering Graduate School, †Dept. of Electronic Engineering, Faculty of Engineering
Mahanakorn University of Technology, 140 Cheumsamphan Rd., Nongchok, Bangkok 10530 Thailand
E-mail: sethakarn.pr@ieee.org*and ssuchada@mut.ac.th†
‡Department of Electrical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang
1 Chalongkrung Rd., Ladkrabang, Bangkok 10520 Thailand, Email: theerayod.wi@kmitl.ac.th‡

***Abstract*—This paper proposes a normalised least mean square algorithm based on Hammerstein spline adaptive filtering. A nonlinear Hammerstein adaptive filters consists of memory-less function modified during learning and the spline control point is automatically controlled by gradient-based method. Simulation results demonstrate that the proposed algorithm exhibits more robust performance compared with the conventional spline adaptive filtering algorithms.**

I. Hammerstein Spline Adaptive Filtering

Following [1], the Hammerstein spline adaptive filter is called *linear-nonlinear network*, as shown in Fig. 1. This network is consisting of two parts as linear and nonlinear network. An adaptive function is applied to the input of linear adaptive filter. Part of nonlinear network is used for the identification of Hammerstein-type nonlinear systems [2]. Other part of non-linear network is combined with an adaptive look-up table [3] and spline interpolation network. Nonlinear SAF has been used practically to model and identify nonlinear systems [4]-[6]. Consider an error $e(n)$ as

$$e_n = d_n - y_n = d_n - \mathbf{w}_n^T\,\mathbf{s}_n \ , \tag{1}$$

where y_n is the output of Hammerstein spline adaptive filtering (HSAF), d_n is the desired signal and $\mathbf{w}_n$ is the adaptive tap-weight vector. The output of nonlinearity $\mathbf{s}(n)$ is given [5]

$$\mathbf{s}_n = \varphi(u_n) = \mathbf{u}_n^T\,\mathbf{C}\,\mathbf{q}_{i,n} \ , \tag{2}$$

$$\mathbf{u}_n = [u_n^3,\ u_n^2,\ u_n,\ 1]^T \ , \tag{3}$$

where $\mathbf{q}_{i_n}$ is the control point vector as

$$\mathbf{q}_{i,n} = [\ q_{i,n}\ q_{i+1,n}\ q_{i+2,n}\ q_{i+3,n}\ ]^T \ .$$

Following [5], the local parameter u_n and index i can be evaluated as

$$u_n = \frac{x_n}{\Delta x} - \left\lfloor \frac{x_n}{\Delta x} \right\rfloor \ , \tag{4}$$

$$i = \left\lfloor \frac{x_n}{\Delta x} \right\rfloor + \frac{Q-1}{2} \ , \tag{5}$$

where Δx is the uniform space between two-adjacent control points, x_n is the input vector with the length of tap delay N, Q is the number of control point, and $\lfloor\cdot\rfloor$ is floor operator.

The output of nonlinearity s_n is related with a nonlinear activation function using the span index i and the local parameter u, where $u \in [0,1]$. Since, spline basis matrix $\mathbf{C}_B$ is called *B-spline* [5].

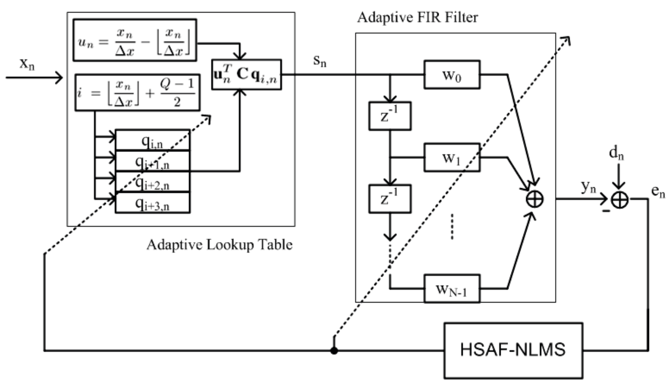


Fig. 1. Linear-Nonlinear network of Hammerstein SAF-NLMS structure.

II. Proposed Hammerstein SAF-NLMS

Following [1], the cost function of normalised least mean square for Hammerstein spline adaptive filter (HSAF-NLMS) can be minimised as

$$J(\mathbf{w}_n, \mathbf{q}_{i,n}) = \min_{\mathbf{w}_n} \left\{ \frac{1}{2} (\mathbf{u}_n^T\,\mathbf{u}_n)^{-1} \mid e_n \mid^2 \right\} \ , \tag{6}$$

where e_n is given in Eq.(1).

Considering the chain rule by differentiating the cost function in Eq.(6) with respect to $\mathbf{w}_n$, we arrive at

$$\frac{\partial J(\mathbf{w}_n, \mathbf{q}_{i,n})}{\partial \mathbf{w}_n} = -(\mathbf{u}_n^T\,\mathbf{u}_n)^{-1} \left\{ e_n\,\mathbf{s}_n \right\} \ , \tag{7}$$

Let the derivative of cost function in Eq.(6) with respect to $\mathbf{q}_{i,n}$ with the chain rule, we arrive at

$$\frac{\partial J(\mathbf{w}_n, \mathbf{q}_{i,n})}{\partial \mathbf{q}_{i,n}} = -(\mathbf{u}_n^T\,\mathbf{u}_n)^{-1} \left\{ \frac{\partial \mathbf{w}_n\,\mathbf{u}_n^T\,\mathbf{C}\,\mathbf{q}_{i,n}\,e_n}{\partial \mathbf{q}_{i,n}} \right\} \ . \tag{8}$$

So, we can get

$$\frac{\partial J(\mathbf{w}_n, \mathbf{q}_{i,n})}{\partial \mathbf{q}_{i,n}} = -(\mathbf{u}_n^T\,\mathbf{u}_n)^{-1} \left\{ \mathbf{C}^T\,\mathbf{u}_n\,\mathbf{w}_n\,e_n \right\} \ . \tag{9}$$

Further, the proposed tap-weight vector $\mathbf{w}_n$ of HSAF-NLMS algorithm is obtained by

$$\therefore \mathbf{w}_{n+1} = \mathbf{w}_n + \frac{\mu_w \, \mathbf{s}_n \, e_n}{\mathbf{u}_n^T \, \mathbf{u}_n} \; , \tag{10}$$

where μ_w is the step-size parameter for learning rate of linear network part of HSAF structure, $\mathbf{s}_n$ and $\mathbf{u}_n$ are defined in Eqs. (2) and (3), respectively.

Therefore, the proposed control points vector $\mathbf{q}_{i,n}$ of HSAF-NLMS can be expressed as

$$\therefore \mathbf{q}_{i,n+1} = \mathbf{q}_{i,n} + \mu_q \frac{\mathbf{C}^T \, \mathbf{u}_n \, \mathbf{w}_n \, e_n}{\mathbf{u}_n^T \, \mathbf{u}_n} \; , \tag{11}$$

where μ_q is the step-size parameter for learning rate of nonlinear network part of HSAF structure.

III. Simulation Results

We simulate the random processes for the computer simulations. Performance of proposed HSAF-NLMS algorithm is compared with the SAF-LMS [5] in the Wiener system over 100 Monte-Carlo trials and 3,000 samples used. The input signal can be generated by

$$x_n = \alpha \cdot x_{n-1} + \sqrt{1-\alpha^2}\zeta_n \; , \tag{12}$$

where ζ_n is a zero mean white Gaussian noise with unitary variance and α is set to $[0.01, 0.99]$.

In the identification, the experiment consists of an unknown Wiener system composed by a linear component as [5].

$$\mathbf{w}_0 = [0.6, \; -0.4, \; 0.25, \; -0.15, \; 0.1] \; .$$

A nonlinear memoryless target function implemented by a 23-point length LUT $\mathbf{q}_0$ which is interpolated by a uniform third degree spline with an interval sampling $\Delta x = 0.2$ as [7]

$$\begin{aligned}\mathbf{q}_0 =& \{-2.2, -2, -1.8, \ldots, -1.0, -0.8, -0.91, 0.42, \\ & -0.01, -0.1, 0.1, -0.15, 0.58, 1.2, 1.0, 1.2, \ldots, 2.0, 2.2\}\end{aligned}$$

Initial parameters of SAF model for adaptive FIR filter are as $\delta_w = 0.001$, $\mu_w = \mu_q = 7.5 \times 10^{-3}$, a signal to noise ratio at $SNR = 10, 20$ dB. Length of tap (M) coefficients of filter is of 5 for both algorithms and C_B is used by following in [5].

Comparison of mean square error (MSE) for the experiments with the different of $\alpha = 0.1, 0.5$ in Fig. 2 and Fig. 3 that show the MSE curves of proposed HSAF-NLMS and SAF-LMS [5] with the two different choices of the parameter α shown in Eq.(12) and the different of SNR is used. While the MSE is calculated as $10\log(e_n^2)$ in dB. It is found that the MSE curves of proposed HSAF-NLMS algorithm can attain the fast convergence compared with the SAF-LMS algorithm at the steady state.

IV. Conclusions

In this paper, we have introduced the proposed normalised least mean square algorithm for Hammerstein spline adaptive filtering (HSAF-NLMS) model. In addition, a constraint on the choice of learning rate is to assure the algorithm convergence. The proposed HSAF-NLMS algorithm can achieve the good performance compared with the conventional SAF-LMS algorithm in the area of system identification.

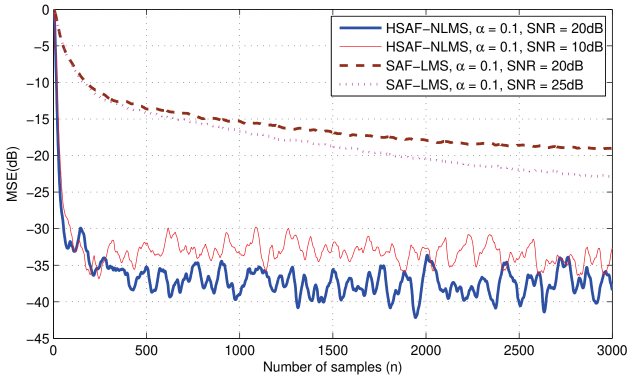


Fig. 2. MSE curves of proposed HSAF-NLMS and SAF-LMS [5] with the different SNR, where $\alpha = 0.1$.

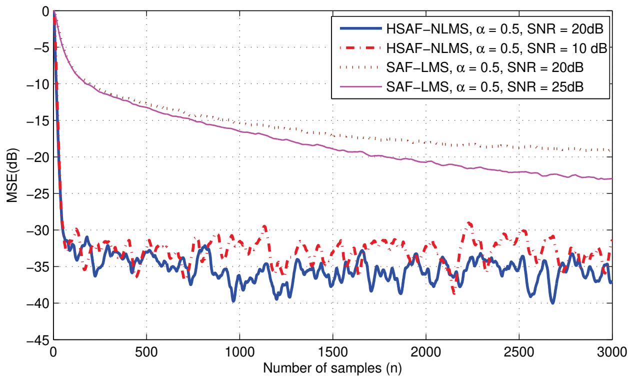


Fig. 3. MSE curves of proposed HSAF-NLMS and SAF-LMS [5] with the different SNR, where $\alpha = 0.5$.

References

[1] M. Scarpiniti, D. Comminiello, R. Parisi and A. Uncini, "Hammerstein uniform cubic spline adaptive filtering : learning and convergence properties", *Signal Processing*, vol. 100, pp. 112-123, 2014.
[2] F. Ding, X.P. Liu, G. Liu, "Identification methods fo rHammerstein nonlinear systems", *Digital Signal Processing*, vol. 21, no. 2, pp. 215-238, 2011.
[3] S. Prongnuch, "The Study of Hardware Co-processing Element Developments based on a HSA", *Electrical Engineering Conference*, Phetchaburi, pp. 745-748, 2016.
[4] M. Scarpiniti, D. Comminiello, R. Parisi and A. Uncini, "Novel cascade spline architectures for the identification of nonlinear systems", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, issue. 7, pp. 1825-1835, 2015.
[5] M. Scarpiniti, D. Comminiello, R. Parisi and A. Uncini, "Nonlinear spline adaptive filtering", *Signal Processing*, vol. 93, issue. 4, pp. 772-783, 2013.
[6] S. Guan, Z. Li, "Normalised apline adaptive filtering algorithm for nonlinear system identification", *Neural Processing Letter*, vol. 5, pp. 1-13, 2017.
[7] C. Liu and Z. Zhang, "Set-membership normalised least M-estimate spline adaptive filtering algorithm in impulsive noise", *Electronics Letters*, vol. 54, no. 6, pp. 393-395, March 2018.

Large-Cell Wireless Train Radio Communications Employing Narrowband Multiple Single Carrier Modulation Schemes for High-Speed Railways

Takuya Takahashi[1] , Hiroshi Kubo[2]
[1]Graduate School of Science and Engineering, Ritsumeikan University,
[2]Department of Electronic and Computer Engineering, Faculty of Science and Engineering, Ritsumeikan University,
[1,2]1-1-1 Noji-Higashi, Kusatsu-shi, Shiga 525-8577, Japan
E-mail :[1]ri0051se@ed.ritsumei.ac.jp , [2]kubohiro@fc.ritsumei.ac.jp

***Abstract*—This paper proposes train radio communications for high-speed railways to achieve the following: 1) higher frequency efficiency; 2) faster mobile speed; 3) larger cell radius. Compared with Global System for Mobile Communications-Railway (GSM-R) and Terrestrial Trunked Radio (TETRA) in European train radio communications, although the conventional Japanese train radio communications have excellent frequency efficiency and communication quality, they have lower mobile speed and smaller cell radius. This paper proposes train radio communications that can achieve the same levels as GSM-R and TETRA for the items 2) and 3) while maintaining the advantages of the conventional Japanese systems for the item 1). Finally, computer simulation results show that the proposed systems can extend the cell radius of the base station to be comparable to those of GSM-R and TETRA at a mobile speed of 300km/h.**

***Index Terms*—Train Radio Communications, High-Speed Railways, Large-Cell, Double Selectivity, Multiple Single Carrier**

I. Introduction

Currently, European mobile communications such as Global System for Mobile Communications-Railway (GSM-R) and Terrestrial Trunked Radio (TETRA) [1], [2] are widely used in the world for train radio communications. On the other hand, the train radio communications currently used in Japan can be roughly divided into the space wave method [3] and the leaky coaxial cable (LCX) method [2]. Since the LCX method requires a higher cost than the space wave method, there are strong demands for realizing wireless high-speed train radio communications. The conventional Japanese train radio communications are superior in terms of frequency efficiency and communication quality; however, they have problems in terms of mobile speed and cell radius compared with GSM-R and TETRA. Thus, the problems of wireless high-speed train radio communications are as follows: a) to achieve higher frequency efficiency; b) to cope with faster mobile speed; c) to achieve longer distance between adjacent base stations (BSs).

To cope with these problems, this paper proposes train radio communications employing point to multi-point (P-MP) communications where multiple BSs simultaneously transmit the same information at the same frequency band, i.e., P-MP communications with a single frequency network (SFN), per-survivor processing maximum-likelihood sequence estimation (PSP-MLSE) with channel prediction [4], and narrowband multiple single carrier (MSC) modulation scheme for problems a), b), and c), respectively. Finally, computer simulation results show that the proposed system can extend the distance between adjacent BSs to approximately 46km at a mobile speed of 300km/h while maintaining the same communication quality and frequency efficiency as the conventional Japanese systems.

II. Problems of Train Radio Communications

Table I shows a comparison of GSM-R and TETRA with Japanese train radio communications.

TABLE I
Comparison of European and Japanese Systems

	GSM-R	TETRA	Space wave	LCX
cell radius	5-10km	10-25km	1.5km	-
bit error rate (BER)	several%	several%	0.01%	0.01%
frequency repetition number	multiple	multiple	1	1
carrier frequency band	900MHz	400MHz	400MHz	400MHz
train speed	300-500km/h		150km/h	300km/h

Table I shows that GSM-R and TETRA are more effective than the Japanese systems in terms of cell radius and train speed, and the Japanese systems are more effective than GSM-R and TETRA in terms of communication quality and frequency efficiency. This paper proposes train radio communications migrating the conventional Japanese space wave method that can cope with large-cell and high-speed movement equal to or greater than those of GSM-R and TETRA. However, high-speed train radio P-MP communications with an SFN have the following problems:

i) larger Doppler shift due to the high-speed of the train;
ii) larger delay spreads of signals from adjacent BSs due to larger cell radius R_{cell}.

The problems i) and ii) are related to time selectivity and frequency selectivity, respectively, and the simultaneous occurrence of the two sensitivities is called double selectivity. In general, the time selectivity is often defined as the maximum Doppler frequency (f_D) normalized by the symbol rate ($f_s = 1/T$), f_DT, and the frequency selectivity is often defined as the maximum delay interval (τ_D) normalized by the symbol period (T), τ_D/T.

Table II shows f_DT in high-speed train radio communications. This paper assumes that the multi-carrier (MC) modulation schemes achieve 4.8ksps by parallel transmission of multiple subcarriers, v is the mobile speed, f_c is the carrier frequency, and N_{SC} is the number of subcarriers.

TABLE II
$f_D T$ IN HIGH-SPEED RAILWAYS

Items	SC	MC	
N_{SC}	1	2	4
$f_s(= 1/T)$	4.8ksps	2.4ksps×2	1.2ksps×4
f_c	400MHz		
v	300km/h = 83.3m/s		
f_D	111Hz		
$f_D T$	2.3%	4.6%	9.3%

Table II shows that the MC modulation schemes must cope with $f_D T$ more than N_{SC} times the normalized Doppler frequency for the SC modulation scheme, assuming a constant occupied bandwidth of the modulation schemes.

III. PROPOSED TRAIN RADIO COMMUNICATIONS

A. P-MP Communications for High Frequency Efficiency

This paper employs P-MP communications where BSs in the zone transmit the same information at the same frequency band to realize high frequency efficiency.

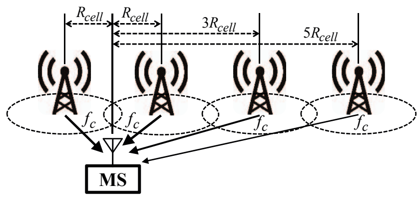


Fig. 1. P-MP communications model.

In the system shown in Fig. 1, only a single frequency (f_c) is allocated to all BSs. Thus, the frequency efficiency is high, but the mobile station (MS) needs to cope with the delay difference corresponding to more than $4R_{cell}$.

B. PSP-MLSE with Channel Prediction for Severe $f_D T$

This paper employs PSP-MLSE for differential encoding of joint detection to cope with severe time selectivity, where $f_D T$ is more than several percent [4]. PSP-MLSE estimates information sequence based on the Viterbi algorithm and can control the trade-off between receiver sensitivity and tracking performance by selecting appropriate channel estimation weights.

C. MSC Modulation Schemes for Large-Cell

To cope with severe frequency selectivity associated with the expansion of cell size, this paper employs narrowband MSC modulation scheme, which is a migrated version of SC modulation scheme employed by the conventional Japanese train radio communications. In addition, MSC modulation scheme can improve the resistance to intersymbol interference because of the small bandwidth occupied by the subcarriers.

IV. COMPUTER SIMULATION

Let us discuss the simulation parameters. The modulation scheme is $\pi/4$-shifted differential QPSK, the number of transmit antennas is 1, the number of receive antennas is 4, and the demodulation scheme is PSP-MLSE. The case of differential detection (DD) is also plotted for reference. Channels are assumed to be independent Rayleigh fading channels, where $f_D T_0$ is 2.3% and T_0 is the symbol period in SC modulation scheme shown in Table II. This simulation assumes P-MP communications with an SFN with 6 BSs. The distances between the MT and BSs are R_{cell}, $3R_{cell}$, $\cdots$, $11R_{cell}$, where the distance attenuation is based on the 2.5th-law.

Fig. 2 shows BER performance as a function of $\Delta\tau/T_0$, where $\Delta\tau$ is the delay difference between adjacent BSs.

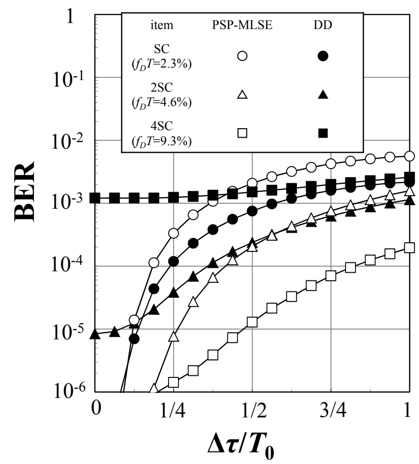


Fig. 2. BER performance as a function of $\Delta\tau/T_0$ at average E_b/N_0 of 25dB.

Fig. 2 shows that the proposed narrowband 4SC modulation scheme achieves BER of less than 10^{-4} on doubly-selective fading channels with $f_D T_0$ of 2.3% and $\Delta\tau/T_0$ is less than 3/4. In addition, that the distance d is 46km between BSs can be obtained from $1/T_0$ of 4.8ksps and $\Delta\tau$ of $3T_0/4$.

V. CONCLUSION

This paper has proposed train radio communications employing P-MP communications with an SFN, PSP-MLSE with channel prediction, and narrowband MSC modulation scheme to cope with the problems in wireless high-speed train radio communications. Finally, computer simulation results have confirmed that the proposed systems can extend the cell radius of the BS to a value comparable to those of GSM-R and TETRA at a mobile speed of 300km/h and BER of 10^{-4}.

Acknowledgments This work was partly supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Number 19K04382.

REFERENCES

[1] Sepura "A comparison of TETRA and GSM-R for railway communications," Available:http://fplreflib.findlay.co.uk/images/pdf/tetratoday/A-comparison-of-TETRA.pdf [Sep. 25, 2019].

[2] Report ITU-R M.2418-0 "Description of Railway Radiocommunication Systems between Trainand Trackside (RSTT)," Nov. 2017.

[3] ARIB STD-T61: "Narrow band digital telecommunication system (SCPC/FDMA)," 2005.

[4] H. Kubo, A. Okazaki, K. Tanada, P. Betrand and K. Murakami, "A multiple-symbol differential detection based on channel prediction for fast time-varying fading," IEICE Trans. Commun., vol. E88-B, no. 8, pp. 3393–3400, Aug. 2005.

A Preprocessing for Sound Source Separation Using Complex Weighted Sum Circuits

Shun Nishimaki and Kenji Suyama
School of Engineering, Tokyo Denki University
5 Senju Asahi-cho, Adachi-ku, Tokyo 120-8551, Japan

***Abstract*—In this paper, a preprocessing method for sound source separation by forming the directivity is studied. A complex weighted sum circuit is used to form the directivity. Furthermore, multiple complex weighted sum circuits are used to spread the suppression interval. However, even if the suppression interval has been spread, the interference sound may be remained due to the reverberation. In the proposed method, as a preprocessing for separation, the frames including just the interference sound is suppressed selectively. To detect the frames having just only the interference sound, the DOA (Direction-Of-Arrival) of the source is used. Applying this preprocessing, the interference sound that can not be suppressed by forming the directivity can be suppressed, and an improvement of the separation performance is expected.**

***Index Terms*—signal processing, sound source separation, microphone array**

I. Introduction

Sound source separation is an important technique for various applications such as a speech recognition and a teleconferencing system.

When the sound source direction is known in advance, the sound source separation can be carried out by forming the spatial directivity. The directivity is adjusted by forming a unity gain toward the target sound direction while keeping forming a null toward the interference sound direction. Furthermore, multiple complex weighted sum circuits are used and the suppression interval is spread. However, the sound sources are expected to come from a wide direction due to the reverberation. Therefore, even if the suppression interval has been expanded, the interference sound may be remained.

In the proposed method, the frames including just the interference sound are detected and suppressed as a preprocessing for separation. In [1], the frequency distribution is created from the DOA estimation results. Then, the frames dominated by single sound source energy are detected by evaluating the frequency distribution with the Hoyer's criterion[2]. In the proposed method, Root-MUSIC (MUltiple SIgnal Classification) is used for the DOA estimation[3]. Then, the histogram is estimated from the DOA estimation results. Finally, the histogram is evaluated by Hoyer's criterion to detect and suppress the frames having just the interference sound. Applying this preprocessing, it is expected to improve separation performance by suppressing the interference sound that can not be suppressed by complex weighted sum circuits.

Several experimental results in a real environment are shown to present the effectiveness of the proposed method.

II. Problem Formulation

Two sound sources, $s_i(n),\ i = 1, 2,$ where i is the sound source index and n is discrete time, are received by two microphones placed at a width d. In a frequency domain, the received signal at the m-th microphone can be expressed as

$$X_m(t,k) = \sum_{i=1}^{2} S_i(t,k) e^{-j\omega_k (m-1)\tau_i(t)} + \Gamma_m(t,k), \quad (1)$$

where t is the frame index, k is the frequency index, $S_i(t,k)$ is a complex amplitude of $s_i(n)$, ω_k is the angular frequency at the k-th frequency index and $\Gamma_m(t,k)$ is the noise component observed at the m-th microphone. $\tau_i(t)$ is the TDOA (Time-Difference-Of-Arrival) between two microphones of $s_i(n)$, and is expressed as $\tau_i(n) = d\sin\theta_i(t)/c$, where $\theta_i(t)$ is a sound source direction of $s_i(n)$, c is the speed of sound. The aim of the sound source separation is to separate the sound source $s_i(n)$ from the received signal $x_m(n)$.

III. Forming the Directivity

To form the directivity, the received signals $X_m(t,k)$ are weighted by the complex weights $W_{i,m}(t,k)$ and summed. The complex weights $W_{i,m}(t,k)$ are calculated by following equations,

$$\begin{cases} W_{i,1}(t,k) + W_{i,2}(t,k) e^{-j\omega_k \tau_{si}(t)} = 1 \\ W_{i,1}(t,k) + W_{i,2}(t,k) e^{-j\omega_k \tau_{ni}(t)} = 0 \end{cases}, \quad (2)$$

where $\tau_{si}(t)$, $\tau_{ni}(t)$ are the TDOA of the target sound direction $\theta_{si}(t)$ and the interference sound direction $\theta_{ni}(t)$. In the equation (2), the first equation means the gain to the i-th target sound equals to 1 and the second equation means the gain to the i-th interference sound equals to 0. Thus, the output signal $Y_i(t,k)$ can be calculated as

$$Y_i(t,k) = W_{i,1}(t,k) X_1(t,k) + W_{i,2}(t,k) X_2(t,k). \quad (3)$$

To spread the suppression interval, multiple complex weighted sum circuits $[W^{[q]}], q = 0, 1, \cdots, Q-1$, are used and the amplitude of the output signal of each circuit are multiplied. Then, the output signals of the multiple weighted sum circuits $\|Y_i(t,k)\|$ are calculated as

$$\begin{aligned} \|Y_i(t,k)\| &= \|Y_i^{[0]}(t,k)\| \times \|Y_i^{[1]}(t,k)\| \times \cdots \times \|Y_i^{[Q-1]}(t,k)\| \\ &= \prod_{q=0}^{Q-1} \|W_{i,1}^{[q]} X_1(t,k) + W_{i,2}^{[q]} X_2(t,k)\|. \end{aligned} \quad (4)$$

Because Q output signals of the multiple weighted sum circuits are multiplied, a correction of the amplitude and phase characteristic are required. The amplitude characteristics is corrected as

$$\langle Y_i(t,k)\rangle = \frac{\prod_{q=0}^{Q-1}\|Y_i^{[q]}(t,k)\|}{\left(\frac{1}{Q}\sum_{q=0}^{Q-1}\|Y_i^{[q]}(t,k)\|\right)^{Q-1}}. \quad (5)$$

To correct the phase characteristics, $\angle Y_i^{[0]}(t,k)$ is used. The separated signals $y_i(n)$ are obtained by IDFT (Inverse Discrete Fourier Transform) of $Y_i(t,k)$. The circuit configuration of the complex weighted sum circuits is shown in Fig.1.

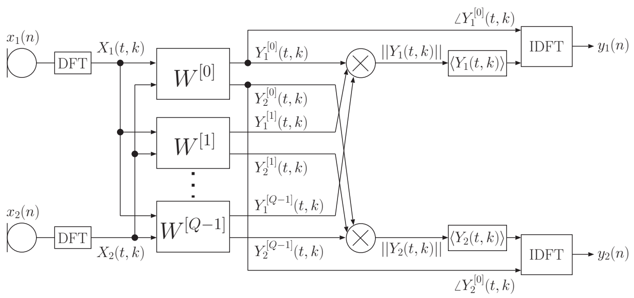


Fig. 1. The circuit configuration of complex weighted sum circuits

IV. Proposed Method

In the proposed method, as a preprocessing for separation, the frames having just the interference sound are suppressed. The process of the proposed method is shown as follows.

step.1 The received signals $x_m(n)$ are transformed to the frequency domain as $X_m(t,k)$ by DFT (Discrete Fourier Transform).

step.2 The DOA is estimated using Root-MUSIC and the histogram is created from the estimation results of the DOA.

step.3 Evaluation of Hoyer's criterion H from the histogram. H is defined as

$$H = \left(\sqrt{N} - \frac{\sum_{n=1}^{N} P_n}{\sqrt{\sum_{n=1}^{N} P_n^2}}\right)\left(\sqrt{N}-1\right)^{-1}, \quad (6)$$

where N is division number of the histograms and P_n is the signal power ratio at the n-th direction. H takes a value from 0 to 1. The larger the value of H, the larger the one sound source component is indicated.

step.4 If $H>\alpha$, where α is the threshold, the frame is judged as one sound source frame and go to step.5. Otherwise, it is judged as two sound source frame or the silent frame and go to step.7.

step.5 Determine the direction θ_{peak} in which the signal power ratio takes maximum in the histogram.

step.6 If $\theta_{interf} - 30 < \theta_{peak} < \theta_{interf} + 30$, where θ_{interf} is the interference sound direction, it is judged as the frame having just the interference sound and multiply $X_m(t,k)$ by the attenuation factor β. Otherwise, it is judged as the frame having just the target sound and go to step.7.

step.7 $X_m(t,k)$ is input to the complex weighted sum circuits and the output signals $Y_i(t,k)$ is output.

step.8 The separated signals $y_i(n)$ are obtained by IDFT (Inverse Discrete Fourier Transform) of $Y_i(t,k)$.

V. Experiments

Several experiments were conducted in the real environment to present the efficiency of the proposed method. The experimental conditions are listed in Table 1. The reverberation time was 1.09[s], and the noise level was 39.8[dB]. The direction of each sound were set to 30° and -30°. The distance from the sound sources to the microphones was set to 1[m].

TABLE I
Experimental conditions

sampling frequency	8000[Hz]
signal length	10[s]
frame size	2048
overlap size	50[%]
microphone width	0.04[m]
frequency band for separation	80–4000[Hz]
number of nulls	31
threshold	0.5
attenuation factor	0.15

The accuracy of the separation was measured by the SIR (Source to Interference Ratio)[4]. The larger the value of SIR, the higher separation performance is indicated. In the experiments, 5 patterns of sound sources were used.

From the experimental results, the average value of SIR was 8.21[dB] in the previous method. On the other hand, the average value of SIR was 12.3[dB] in the proposed method. Furthermore, we confirmed that SIR was improved in all sound source patterns. This indicates that the proposed method suppresses interference sound that can not be suppressed by the complex weighted sum circuits.

VI. Conclusion

In this paper, a preprocessing for sound source separation using complex weighted sum circuits was studied. The DOA was used to detect the frames with only the interference sound. Several experimental results in real environment were shown to present the effectiveness of the proposed method.

References

[1] N. Iwasaki, M. Tamaki, S. Fukase, K. Inoue and H. Gotanda, "A Study on Sound Source Tracking Based on a Frame DOA Estimation," Proceedings of the 47th ISCIE International Symposium on Stochastic Theory and Its Applications, pp.74–80, 2015.

[2] Patrik O. Hoyer, "Non-negative Matrix Factorization with Sparseness Constraints," Journal of Machine Learning Research, Vol.5, pp.1457–1469, 2004.

[3] T. Suzuki and K. Suyama, "Sound source tracking based on updating histogram," Proc. of IEEE ISCP, pp. 140–143, October 2012.

[4] E.Vincent, R.Gribonval, and C.Fevotte, "Preformance measurement in blind audio source separation," IEEE Transactions on Audio, Speech, and Language Processing, Vol.14, no.4, pp. 1462–1469, July 2006.

Feedback Active Noise Control using Linear Prediction Filter for Colored Wide-band Background Noise Environment

Riku Takasugi
Saitama University, Saitama, Japan
E-mail: takasugi@sie.ics.saitama-u.ac.jp

Yosuke Sugiura
Saitama University, Saitama, Japan
E-mail: ysugiura@mail.saitama-u.ac.jp

Tetsuya Shimamura
Saitama University, Saitama, Japan
E-mail: shima@sie.ics.saitama-u.ac.jp

***Abstract*—In this paper, we propose a new Feedback Active Noise Control (FB-ANC) system to improve the suppression performance for narrow-band noise buried in colored wide-band background noise. The proposed system adopts a linear prediction filter to the input signal in order to separate narrow-band noise and wide-band noise. The simulation result shows the proposed method can perform robust noise suppression.**

I. Introduction

Acoustic noise problems become more and more serious as increasing number of industrial equipment such as engines, blowers, fans, transformers, compressors, and so on. As the conventional approach to suppress the acoustic noise, we often use the passive techniques such as enclosures, barriers, and silencers. Such the passive silencers require extremely high spatially costs for the satisfactory attenuation over broad frequency range.

Active Noise Control (ANC) is an electroacoustic system that cancels the primary undesired noise by the antinoise sound emitted from the second source speaker. ANC system can be categorized into 2 types, which are the feedforward ANC and the feedback ANC. The feedback ANC system efficiently attenuates narrow-band noise but has a problem that the noise suppression effect is degraded when the narrow-band noise is disturbed by colored wide-band noise. This is because the whitening effect of the feedback ANC strongly affects the colored wide-band noise when the wide-band noise level is high. To solve this, this paper describes a method for separating narrow-band noise and colored wide-band noise using a linear predictor to efficiently attenuate narrow-band noise.

II. Active noise control using linear prediction

We describes the principle of the proposed method. Figure 1 shows the structure of the proposed method. Comparing with the original feedback ANC system, three linear predictors are added to the error signal $e(n)$, the filtered restoration noise $r(n)$, and presudo-output signal with secondary path model $y'(n)$. These linear predictors have common filter coefficients. The detail of the signals in Fig. 1 are shown as below.

$d(n)$:Noise
$y(n)$:Output signal
$e(n)$:Error signal

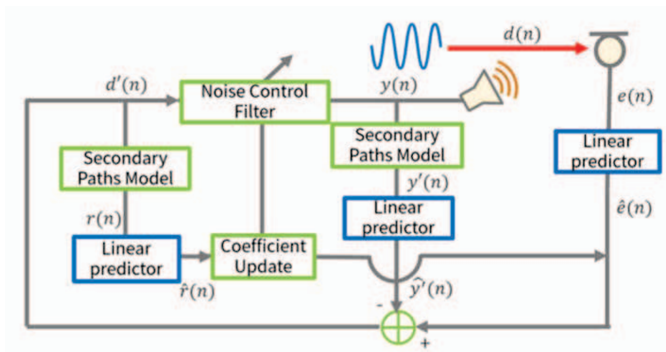


Fig. 1. The structure of linear prediction feedback ANC.

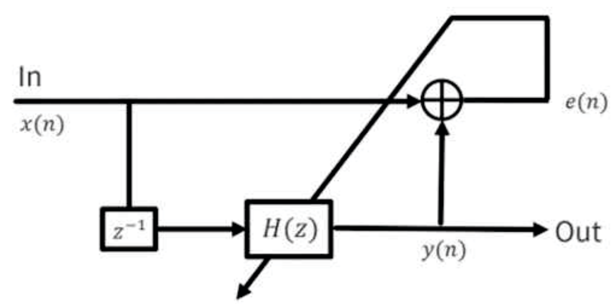


Fig. 2. The structure of linear prediction.

$\hat{e}(n)$:Predicted error signal
$r(n)$:Filtered-restoration noise
$\hat{r}(n)$:Predicted filtered-restoration noise
$d'(n)$:Restoration noise
$y'(n)$:Pseudo output signal with secondary path model
$\hat{y}'(n)$:Predicted pseudo output signal with secondary path model

To achieve the linear predictor reducing the wide-band noise, we utilize an adaptive filter. The structure of the linear predictor is shown in Figure 2. The output signal $p(n)$ is

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

obtained by

$$p(n) = \sum_{i}^{M} h_i(n)x(n-i), \tag{1}$$

where $x(n)$ is the input, $h_i(n)$ is the i-th filter coefficient corresponding to $x(n-i)$, and M is the filter order. prediction error used for updating the filter coefficients is calculated by

$$p_e(n) = x(n) - p(n), \tag{2}$$

The filter coefficient $h_i(n)$ is updated using Normalized Least Mean-Square (NLMS) algorithm. The updation is defined by

$$h_i(n+1) = h_i(n) - \mu \frac{p_e(n)x(n-i)}{E[x^2(n-i)]},\ i = 1, \cdots, M, \tag{3}$$

where μ is the step-size parameter and $E[\cdot]$ is a mean operator.

The linear predictor can predict only the periodic components included in the input. After convergence, ideally the output of the linear predictor includes only a narrow-band noise signal since the background wide-band signal is weakly periodic.

III. Noise suppression experiment using harmonic noise

In order to verify the effectiveness of the proposed method, the noise suppression simulation was performed. As the narrow-band noise, the recorded sound of driving vacuum cleaner is used. In the proposed method, we set $M = 40$ and $\mu = 0.5$. For comparison, we used the normal feedback ANC which is the same system as shown in Fig. 1 without the linear predictors. In the feedback ANC, the order of the noise control filter is 150, the order of the secondary path model is 100. The secondary path model is appropriately estimated in advance. For updating the noise control filter, we used NLMS algorithm with $\mu = 0.01$.

Figure 3 shows the output waveforms for three methods, which are no ANC, the conventional ANC, and the proposed ANC. From this figure, ANC works well for suppressing the noise, but there is a little difference between the result of the conventional ANC and the one of the proposed ANC.

Figure 4 shows the power spectrum of the error signal after convergence. From this figure, we can see that the input narrow-band noise whose fundamental frequency is 500Hz is buried in the wide-band colored noise. The conventional ANC, which corresponds to the green line, provides the highly suppression noise performance for low frequency band, since the conventional ANC works so as to whiten the input signal, that is, to suppress the high level narrow-band noise such as 500Hz frequency component and 1000Hz one. On the other hand, the low level narrow-band noise such as 5000Hz frequency component or 6000Hz one tends to be ignored. The proposed method, which is blue line, can suppress not only the high level narrow-band noise but also the low level one. In Fig. 4, it is clear that the proposed method can effectively suppress the narrow-band noise and output only the wide-band noise. From the above results, we can verify the effectiveness of the proposed method.

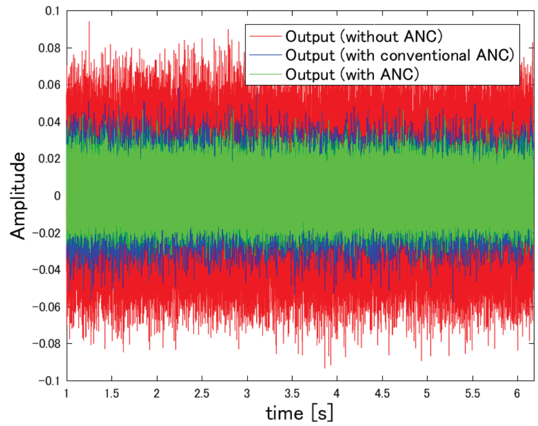


Fig. 3. Waveform obtained from error microphone of noise suppression experiment.

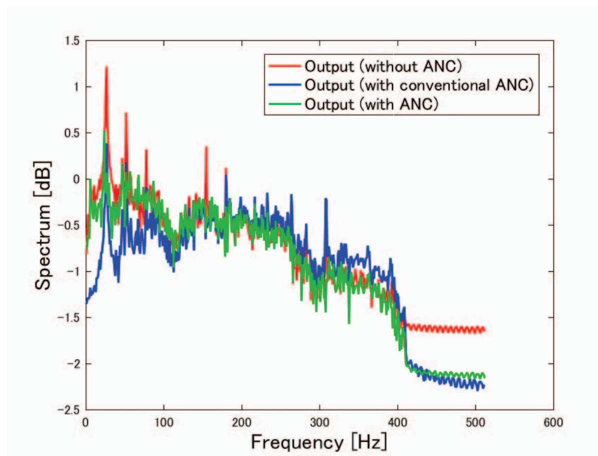


Fig. 4. Spectrum of the error signal after convergence (red: only noise, green : the conventional method, blue : the proposed method.

IV. Conclusion

In this paper, we proposed a new feedback ANC system using three linear predictors. Since the linear predictor allows only narrowband noise to pass, we can separate the narrow-band noise from the background wide-band noise which degrades the noise suppression effect. From the experiment, the effectiveness of the method was confirmed.

References

[1] L. J. Eriksson, " Recursive algorithms for active noise control, " Proc. Int. Symp. Active Control of Sound Vibration, pp. 137–146, 1991.
[2] S. M. Kuo and D. R. Morgan, " Active Noise Control: A Tutorial Review, " Proceedings of the IEEE, vol.87, no.6, pp.943-973, 1999
[3] A. Kawamura, K. Fujii, Y. Itoh and Y. Fukui " A new noise reduction method using linear predictor and adaptive filter, " IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002

Chord Label Estimation from Acoustic Signal Considering Difference in Electric Guitars

Nozomiko Yasui*, Masanobu Miura†and Tetsuya Shimamura*
**Graduate School of Science and Engineering, Saitama University*, Saitama, Japan
†*Faculty of Music Studies, Kunitachi College of Music*, Tokyo, Japan
*{nyasui, shima}@sie.ics.saitama-u.ac.jp
†miura.masanobu@kunitachi.ac.jp

***Abstract*—Our previous study has developed a chord estimation system. Depending on the guitars, however, there are problem which it would be erroneously estimated as different chords. To solve this problem, this study proposes an improved method by focusing on the performance sound estimation in signal processing. As a result, it is found that the differences between the kinds of guitars are reduced using the method to be proposed in this study.**

***Index Terms*—chord label estimation, electric guitar, signal processing, comb filter, pitch class profiles**

I. Introduction

Tonal music is constructed by melodies and accompanying harmonies. Therefore, the chord constructing harmony is an important element for musical excerpts in tonal music. For this reason, many studies, which aim at realizing chord recognition, have been intensively conducted [1-4]. In chord label estimation from acoustic signals, pitches estimation included in a chord has been the primal problem. A solution to this problem has been attempted using several methods such as the pitch class profiles (PCP)[5] or the chroma vectors[6], so a chord label is obtained according to these estimated pitches. At this point, many studies assume that the chord label is uniquely determined by a combination of pitches or notes. Chord is defined as the combination of pitches or notes, but it is difficult to determine an appropriate chord label by observing the combination of notes because of the following reasons. In actual musical performances or musical compositions, some techniques concerning allocations of chord notes such as "omitting", "inversions" and "tension voicing" are often used. By using these techniques, the allocation of notes are not the same as the definition of chords, that is, the allocated notes are sometimes reduced, increased, and changed to other notes. Therefore, to determine a chord label by observing such "altered" notes, it is hard to determine an appropriate chord label even if we know the definition of the chord as a combination of notes.

We tried to solve these problems by constructing and applying a "Searching tree for chord labels" and "Chord progression database"[8]. However, the estimation accuracy depended on the played guitar. Here, this study attempt to solve the dependence by investigating the difference in electric guitars. All of the chord types employed in this study are the sixteen patterns frequently used in chord guitar performances.

II. Chord Label Estimation Method [8]

First, sounds from an electric guitar are converted into the chroma vectors, which represent the power of all pitch classes on a twelve-tone scale. Then, performed notes, which represent a combination of notes that are recognized by the system as actually performed pitch classes using electric guitar, are obtained by calculating the ratio of powers in the chroma vectors. Next, possible chord labels, which have a combination of notes in the performed notes are listed by using the "Searching tree". Then, the performed chord progression (it is also regarded as a sequence of chords which are recognized by the system for past performances) is compared with chord progression patterns included in the "Chord progression database" to estimate chord labels with a high possibility that should be performed at the time. Finally, the chord label for performance is estimated by calculating a logical AND operation between results from the search tree and the chord progression database.

A. Chroma Vectors

The chroma vectors show the total number of powers in the same pitch class of each octave. Thus, the powers in each frequency domain included in input sounds are calculated using fast Fourier transform (FFT). Input sounds sampled at a rate of 22 kHz are divided into 8182 points and quantized at 16 bits. The range of pitches dealt with is approximately from 82 to 5000 Hz (E2 – D#8) and the chroma vectors are calculated in six octaves. Therefore, the chroma vectors v_c(c: pitch class) are calculated with the following equations.

$c = \{$E, F, F#, G, G#, A, A#, B, C, C#, D,D#$\}$
E(E2 – E7) $\leq c \leq$ B(B2 – B7)

$$v_c = \sum_{o=2}^{7} P_{c,o} \tag{1}$$

C(C3 – C8) $\leq c \leq$ D#(D#3 – D#8)

$$v_c = \sum_{o=3}^{8} P_{c,o} \tag{2}$$

where o is the number of octave (e.g. 2 of E2) and $P_{c,o}$ is the powers of the oth octave in pitch class c. Each power in the $P_{c,o}$ is calculated using a comb filter shown in Fig. 1 because there is a possibility that the performed pitches are different

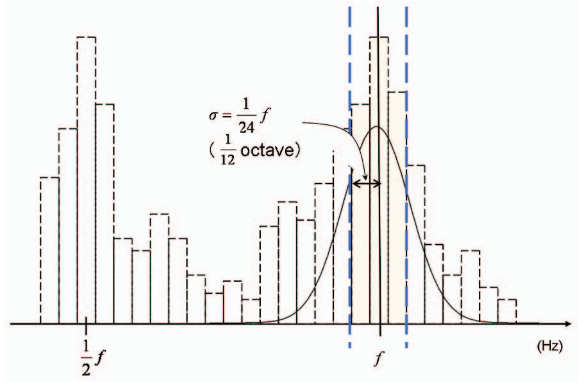


Fig. 1. Shape of individual comb filter[8]

from the theoretical frequency f. The comb filter introduced here corresponds to the normal distribution, and the standard deviation is $\frac{1}{24}f$ ($\frac{1}{12}$octave).

B. Estimating Performed Notes

First, averages of each power in the chroma vectors within a set time is calculated. Next, stronger four pitch-classes are extracted in order from the top. If the pitch class with the strongest power is over the threshold, the system regards that the chord performance has been done and the pitch class is decided as one of the performed notes. The remaining pitch classes are dealt with as performed notes if each power is over 65% in the power of the strongest pitch class. Thus, the maximum number of performed notes is four.

III. Investigation of difference in electric guitars

To investigate the influence of the difference in electric guitars on performed notes estimation, we compared the spectrum of each open notes on two guitars. The used guitars were Godin's xtSA (Gt 1) and Gibson's LesPaul Classic (Gt 2). As a result, the harmonic structure was almost the same, but the power around each harmonic overtone was different.

To investigate characteristics that influence the estimation accuracy in single chord performance played by those guitars, estimated performed notes, chord label, and temporal change of power on each performed notes were recorded when single played chord was input into the chord label estimation system developed in our previous study [8]. Played chords were a total of 48 chords of 12 tones in 4 types. As a result, the power on the third and fourth performed notes influenced chord label estimation.

IV. Evaluation Experiment

A. Method

From the result described in III, it is shown that calculating chroma vectors and estimating performed note have some problems. So, this study employed cutting high frequency components and another comb filter corresponding to Cauchy distribution in calculating chroma vectors, another threshold value in estimating performed notes. In this experiment, a player with five years of experience was asked to perform twenty conjunctive chords extracted from ten Japanese popular songs constructed from the chords employed in this study. The used guitars were Gt 1, Gt 2, Ibanez's RG1520GK 5(Gt 3), Greco's SE-600(Gt 4). This experiment was conducted five times with those guitars.

B. Result and Discussion

The average chord label estimation accuracy on each guitar(Gt 1 - Gt 4) was calculated. The estimation accuracy is defined as the chord label estimation rate for chord performances in the musical excerpt. In the case of only Gt 3, the average estimation accuracy of our previous method [8] is the highest in all methods. In the case of the others, the average estimation accuracy of the method employed cutting high frequency components is higher than the accuracy of the other methods. Thus, it is thought that cutting high frequency components is effective for reducing the influence of the difference in electric guitars. However, the average estimation accuracy using cutting high frequency is 40 %, the average accuracies in Gt 1(42 %), Gt 2(38 %) and Gt 4(39 %) are lower than 50 %. Improving estimation accuracy will be discussed in the near future.

V. Conclusion

This study proposes an improved method by focusing on the performance sound estimation in signal processing. As a result, it is found that the differences between the kinds of guitars are reduced using the method proposed in this study. Future works are to improve the influence of difference in electric guitars, to enlarge patterns of chord progression and prioritize chord labels in a database, and to deal with many more chord types.

References

[1] T. Yoshioka, T. Kitahara, K. Komatani, T. Ogata and H. G. Okuno, "Automatic chord transcription with concurrent recognition of chord symbols and boundaries," Proceedings of International Conference on Music Information Retrieval, pp. 100–105, 2004.

[2] A. Sheh and D. P. W. Ellis, "Chord Segmentation and Recognition using EM-Trained Hidden Markov Models," Proceedings of the International Conference on Music Information Retrieval, 2003.

[3] K. Lee and M. Slaney, "Automatic Chord Recognition from Audio Using an HMM with Supervised," Proceedings of International Conference in Music Information Retrieval, 2006.

[4] G. Cabral, F. Pachet and J. P. Briot, "Automatic X ttaditional descriptor extraction: the case of chord recognition," Proceedings of International Conference on Music Information Retrieval, 2005.

[5] T. Fujishima, "Realtime chord recognition of musical sound: A system using Common Lisp Music," Proceedings of International Computer Music Conference, 1999.

[6] M. Goto, "SmartMusicKIOSK: Music-playback interface based on chorus-section detection method," Journal of the Acoustical Society of America, vol.115, no.5, Pt.2, p.2494, 2004.

[7] C. Harte, M. Sandler, S. Abdallah and E. Gómez, "Symbolic representation of musical chords: a proposed syntax for text annotations" Proceedings of International Conference on Music Information Retrieval, 2005.

[8] Y. Konoki, N. Emura, M. Miura, "Chord Estimation Using Chromatic Profiles of Sounds Played by an Electric Guitar," Proceeding of International Conference on Music Perception and Cognition, pp. 734–737, 2008.

A Benchmark for Homework Tidiness Assessment

Hanxiao Wu[1], Zhenyu Zhang[1], Zhichao Zheng[1], Fei Shen[1], Weiwei Zhang[1], Jianqing Zhu[1*] and Huanqiang Zeng[2]
[1]College of Engineering, Huaqiao University, 269 Chenghua North Road, Quanzhou, Fujian, China
Email:{1695131046, jqzhu}@hqu.edu.cn
[2]College of Information Science and Engineering, Huaqiao University, 668 Jiemei Avenue, Xiamen, Fujian, China
Email: zeng0043@hqu.edu.cn

***Abstract*—The homework tidiness assessment aims to auto evaluate the writing tidiness of homework, playing an important role in daily teaching.However, there is still no comprehensive basis for homework tidiness assessment. For this, a benchmark for homework tidiness assessment (HTA) is proposed. Firstly, a database named HTA 1.0 containing 1000 homework images is collected. Each image is manually annotated by multiple volunteers. Secondly, a comprehensive evaluation protocol is designed, using *mean absolute error* (MAE), *root mean square error* (RMSE), *mean absolute percentage error* (MAPE) and *accuracy* (Acc) as performance indicators. Finally, three deep learning models (i.e., LeNet, AlexNet and VGGNet) are applied as baseline methods and the results are reported and analyzed.**

***Index Terms*—Homework Tidiness Assessment, Deep Learning, Benchmark**

I. Introduction

Homework plays an important role in daily teaching for each teacher and also is an indispensable part of everyone's study life. Therefore, it is interesting to design an automatic grading method for homework tidiness assessment.

The most relevant works in this paper include aesthetic image analysis [1]–[3]. For example, Datta et al. [1] models the aesthetic image analysis task as a regression problem. Jin et al. [3] proposed a normalized regression model for aesthetic quality assessment. Kao et al. [2] achieved the most advanced aesthetic quality assessment results by using a convolutional neural network based regression model. Homework images usually do not contain complex backgrounds or intricate aesthetic meanings, however homework tidiness assessment is still underestimated since there is no public database dedicated to it so far. For this, a homework tidiness assessment benchmark is proposed in this paper.

The main contributions can be summarized into three aspects. 1) HTA 1.0 database, which contains 1000 homework images annotated with multiple volunteers. 2)A comprehensive evaluation protocol, which uses *root mean square error* (RMSE) [4], *mean absolute percentage error* (MAPE) [5], *mean absolute error* (MAE) [6] and *accuracy* (Acc) as performance indicators. 3) Baseline results, three deep learning based baseline methods are evaluated and analyzed.

* Jianqing Zhu is the corresponding author.

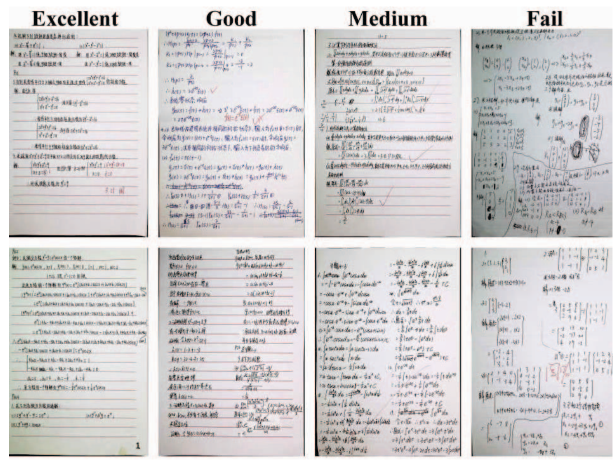


Fig. 1. Classical homework samples selected from HTA 1.0 dataset. The homework images of the same grade are arranged in the same column.

II. Method

A. Database Collection

We collect the HTA 1.0 database including 1000 homework written on B5 sized papers and transcribe them into digital images by using a high definition camera equipment (Kemi GP-3000). Fig. 1 shows the same classical samples of the HTA 1.0 database. All these homework are from the Thing of Internet major and the Information Engineering major.

The annotation of each sample in the HTA 1.0 database is finished by at least 10 volunteers in different majors(i.e., Business Administration, Law, Pharmacy, etc.) For each homework, each volunteer is told to give a score ranged from 0 to 10. The larger score means the better tidiness. When all volunteers finish annotations, the highest score and the lowest score of a homework are firstly removed to avoid possible prejudices. Then, the average of each homework image is taken to get the final tidiness score. Table I shows the statistics of the homework tidiness assessment (HTA) 1.0 database. As shown in Table I, there are fail, medium, good and excellent grades resulted from discretizing scores. We release the HTA 1.0 database on BaiduYun, which can be downloaded from https://pan.baidu.com/s/1WDRco3sQiTNnEVC9rMZmQA, and the enter code is s51u.

B. Evaluation Protocol

To test the performance of a homework tidiness assessment method, RMSE [4], MAPE [5] and MAE [6] are applied as

TABLE I
THE STATISTICS OF THE HOMEWORK TIDINESS ASSESSMENT (HTA) 1.0 DATABASE.

Grade	Fail	Medium	Good	Excellent
Score	$s<6$	$6\leqslant s<7$	$7\leqslant s<8$	$s\geqslant 8$
Number	58	229	415	298

TABLE II
PERFORMANCE OF DIFFERENT METRICS ON HTA 1.0 DATA SETS

	RMSE [4]	MAE [6]	MAPE [5]	Acc
LeNet [7]	8.28	8.07	27.55	75.10
AlexNet [8]	2.77	1.92	**1.65**	**87.15**
VGGNet [9]	**1.44**	**1.09**	2.37	78.69

indicators. In addition, we further design an accuracy (Acc) indicator to evaluate the performance intuitively, which is formulated as follows:

$$\text{Acc} = \frac{R + Q \times 0.5 + W \times 0}{N}, \tag{1}$$

where R represents the number of samples whose predicted grades are equal to ground truths (i.e., manual annotated grades); Q represents the number of samples whose predicted grades are one grade different from those ground truths; W denotes the count of samples whose predicted scores are > 1 scores different from those ground truths; N represents the total number of samples. Based on Eq. (1), a 5-fold cross-validation is implemented to report the baseline performance.

C. Baseline Methods

LeNet [7], ALexNet [8] and VGGNet [9] are adopted as baseline methods. For both these three deep models, the mean square error function is applied as the loss function. The input data are uniformly resized into 256×256 sized images, and all max pooling layers use 3×3 sized pooling window. Other configurations are the same with those in [7]–[9].

III. EXPERIMENT

As can be seen from Table II, AlexNet [8] performs best in MAPE [5], while VGGNet [9] performs best in RMSE [4] and MAE [6]. In addition, AlexNet [8] performs best in homework tidiness accuracy. Figure. 2 shows the prediction of tidiness results of 20 test samples selected randomly in three different depth networks. It can be seen that in test samples, the prediction results in AlexNet [8] are close to the real value, but LeNet [7] and VGGNet [9] fluctuated greatly.

IV. CONCLUSION

In this paper, a homework tidiness assessment benchmark is proposed. A database named HTA 1.0 is firstly collected, which is composed of 1000 manually annotated images. Moreover, a comprehensive evaluation protocol using RMSE, MAPE, MAE, and accuracy as performance indicators are designed. In addition, three deep learning models (i.e., LeNet, ALexNet and VGGNet) are applied as baseline methods and the corresponding results are reported and analyzed. The baseline results show that the homework tidiness assessment is still challenging. We hope that the proposed benchmark will facilitate the study of homework tidiness assessment.

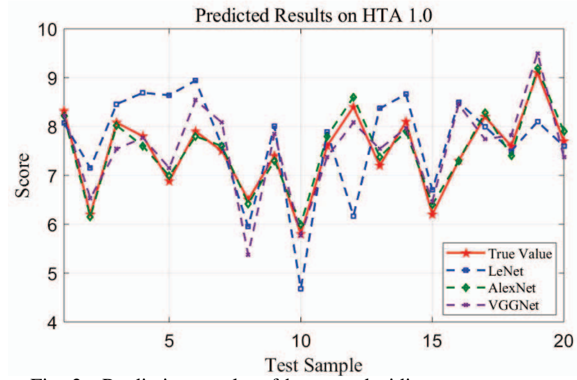


Fig. 2. Prediction results of homework tidiness assessment.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under the Grants 61976098, 61602191, 61871434, 61802136 and 61876178, in part by the Natural Science Foundation of Fujian Province under the Grant 2018J01090, in part by the Natural Science Foundation for Outstanding Young Scholars of Fujian Province under the Grant 2019J06017, in part by the Open Foundation of Key Laboratory of Security Prevention Technology and Risk Assessment, People's Public Security University of China under the Grant 18AFKF11, in part by the Science and Technology Bureau of Quanzhou under the Grants 2018C115R, 2017G027 and 2017G036, in part by the Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University under the Grants ZQN-PY418 and ZQN-YX403, and in part by the Scientific Research Funds of Huaqiao University under the Grant 16BS108, in part by the Teaching Development and Reform Program of Huaqiao University under the Grant 18JF-JXGZ25.

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in *European conference on computer vision (ECCV)*. Springer, 2006, pp. 288–301.
[2] Y. Kao, C. Wang, and K. Huang, "Visual aesthetic quality assessment with a regression model," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 1583–1587.
[3] B. Jin, M. V. O. Segovia, and S. Süsstrunk, "Image aesthetic predictors based on weighted cnns," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 2291–2295.
[4] W. Wang and Y. Lu, "Analysis of the mean absolute error (mae) and the root mean square error (rmse) in assessing rounding model," in *IOP Conference Series: Materials Science and Engineering (MSE)*, vol. 324, no. 1. IOP Publishing, 2018, p. 012049.
[5] S. Mohanty, S. K. Verma, and S. K. Nayak, "Dynamic mechanical and thermal properties of mape treated jute/hdpe composites," *Composites Science and Technology*, vol. 66, no. 3-4, pp. 538–547, 2006.
[6] P. Demorest, T. Pennucci, S. Ransom, M. Roberts, and J. Hessels, "A two-solar-mass neutron star measured using shapiro delay," *nature*, vol. 467, no. 7319, p. 1081, 2010.
[7] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

Super-Resolution via Wavelet Transform and Advanced Learning Techniques

Yi-Wen Chen
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
r06942046@ntu.edu.tw

Jian-Jiun Ding
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
jjding@ntu.edu.tw

***Abstract*—Image super-resolution aims to generate a high-resolution (HR) image from a low-resolution (LR) input image. In this paper, we propose a deep learning-based approach for image super-resolution. We use the wavelet transform to separate the input image into four frequency bands, and train a model for each sub-band. By processing information from different frequency bands via different CNN models, we can extract features more efficiently and learn better LR-to-HR mapping. In addition, we add dense connection to the model to make better use of the internal features in the CNN model. Furthermore, geometric self-ensemble is applied in the testing stage to maximize the potential performance. Extensive experiments on four benchmark datasets show the efficiency of the proposed method.**

Keywords—super-resolution, interpolation, convolutional neural network, wavelet transform

I. INTRODUCTION

As the thriving development of deep learning in recent years, deep convolutional neural networks (CNNs) have been widely used in many computer vision tasks, including recognition and segmentation. Recent work on image super-resolution also resorts to CNNs and achieves good performance.

Dong et al. [1] first introduce the super-resolution CNN (SRCNN) method to learn a nonlinear mapping from low resolution (LR) to high resolution (HR) images via a CNN model. After that, numerous approaches are proposed to improve the performance by increasing the network depth [2], exploring different loss functions [3] or adopting more efficient upsampling steps [4]. However, existing methods usually require several days of training on large scale of datasets to achieve favorable performance. Furthermore, most of the methods need to train a specific model for each scaling factor, which is not practical for usage. To deal with these problems, we propose a learning-based approach for image super-resolution that can be efficiently trained and is applicable to several scaling factors. To facilitate the training process, we propose to use the 2D discrete wavelet transform to separate the information into four frequency bands. We train a CNN model for each frequency band and use the 2D inverse discrete wavelet transform to reconstruct the HR image from the four outputs of CNN models. In addition, we apply dense connection to encourage feature reuse. To further augment the results, geometric self-ensemble strategy is adopted to make use of different outputs from the geometrically transformed inputs.

In Section II, we describe the network architecture of the proposed method and the details of each component. Then we provide the experimental results in Section III and conclude the paper in Section IV.

The main contributions of the proposed method are summarized as follows:

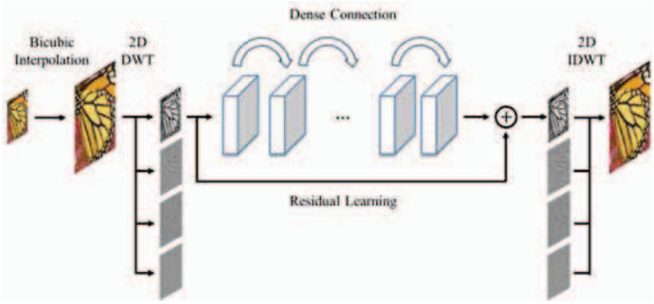


Fig. 1. Network architecture of the proposed algorithm.

1. We adopt the wavelet transform to separate the image into low and high frequency bands, and facilitate the training process of the network.

2. We utilize dense connection to make better use of the features in each layer of the CNN model.

3. We employ geometric self-ensemble to improve the final results without additional training.

II. PROPOSED METHOD

A. Network Architecture

We construct our network based on the very deep super resolution (VDSR) [2] method, for its efficiency in training and ability to deal with multiple scaling factors. The network architecture is depicted in Fig. 1.

Since the information of an image can be separated into low and high frequency bands, we propose to use the 2D discrete wavelet transform to extract features in low and high frequency bands in horizontal and vertical directions, and reconstruct the images of these frequency bands separately. By conducting the 2D discrete wavelet transform, we obtain four images that preserve either low or high frequency information in horizontal and vertical directions. For each of the frequency bands, we train a convolutional neural network to extract features and reconstruct high-resolution information in that frequency band. We use the 2D inverse discrete wavelet transform to obtain the HR image from each of the four outputs.

B. 2D Discrete Wavelet Transform

Wavelet transform is often used in numerical analysis and functional analysis. One of its advantages over Fourier transform is that it captures both frequency and location information.

Discrete wavelet transform is any wavelet transform where the wavelets are discretely sampled. To perform the 1D discrete wavelet transform (DWT), a signal $x[n] \in R^N$ is passed through a half band high-pass filter $G_H[n]$ and a low-pass filter $G_L[n]$ simultaneously. We adopt the Haar wavelet in the proposed method, where the high-pass and low-pass filters are defined as:

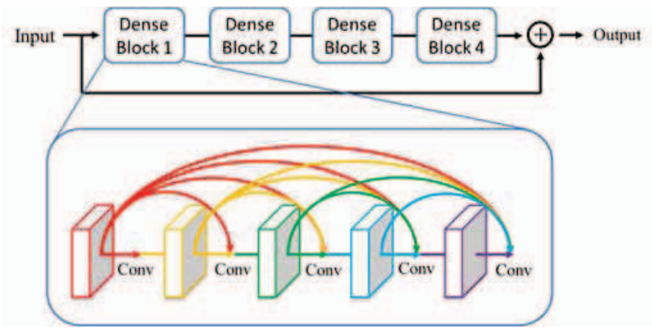


Fig. 2. Illustration of dense connection in the proposed network.

$$G_H[n] = \begin{cases} 1, & n = 0 \\ -1, & n = 1 \\ 0, & otherwise \end{cases}, \quad (1)$$

$$G_L[n] = \begin{cases} 1, & n = 0, 1 \\ 0, & otherwise \end{cases}. \quad (2)$$

The output will be downscaled by the factor of 2 after each filter. To deal with 2D signals such as images, we adopt the 1D discrete wavelet transform first along the horizontal direction and then the vertical direction. Then we obtain four feature maps with half size of the original image, retaining either low or high frequency information in horizontal and vertical directions.

C. Dense Connection

In order to make full use of the features in different layers, we adopt dense connection to the original network of VDSR [2]. As shown in Fig. 2, the network is separated into four dense blocks. In each dense block, every two layers are connected in a feed-forward fashion. Features from different layers are combined by concatenation.

D. Geometric Self-Ensemble Augmentation

In order to explore the potential performance of the proposed model, we adopt the geometric self-ensemble strategy similar to that in [3]. In the testing stage, we flip and rotate the input LR image I^{LR} to generate seven additional inputs $I_{n,i}^{LR} = T_i(I_n^{LR})$, where each T_i stands for one of the 8 geometric transformations including the original one. With the 8 augmented low-resolution images, we generate the corresponding SR images $\{I_{n,1}^{SR}, \dots, I_{n,8}^{SR}\}$ respectively using the proposed network. Inverse transform is then applied to those output images to produce the original geometry $\tilde{I}_{n,i}^{SR} = T_i^{-1}(I_{n,i}^{SR})$. Finally, we average all the inverse transformed outputs to generate the final result $I_n^{SR} = \frac{1}{8}\sum_{i=1}^{8} \tilde{I}_{n,i}^{SR}$.

III. Experimental Results

For training our network, we use 91 images from Yang *et al.* [5] and 200 images from BSD [6], following [2]. We evaluate the performance on four benchmark datasets, including the Set5 [7], Set14 [8], BSD100 [6] and Urban100 [9] datasets. The results are evaluated with PSNR. Following the setting in [2], we only consider PSNR on the Y channel of the transformed YCbCr color space.

In TABLE I, we compare the proposed method with state-of-the-art image super-resolution algorithms for scaling factor 2. We show that our algorithm outperforms other methods on the Set5, Set14 and BSD100 datasets, and performs competitively to DRRN on the Urban100 dataset.

TABLE I. PSNR (in dB) of reconstructed HR images on four datasets by different image super-resolution algorithms.

Dataset	Set5	Set14	BSD 100	Urban 100
SRCNN [1]	36.34	32.18	31.11	28.65
A+ [10]	36.55	32.28	30.78	29.20
SRCNN-Ex [1]	36.66	32.45	31.36	29.50
SelfExSR [9]	36.49	32.44	31.18	29.54
VDSR [2]	37.53	33.03	31.90	30.76
DRCN [11]	37.63	33.04	31.85	30.75
FSRCNN [4]	37.00	32.63	31.43	30.70
DRRN [12]	37.74	33.23	32.05	**31.23**
SRGAN [13]	29.40	26.02	25.16	26.49
Ours	**37.89**	**33.54**	**32.25**	31.09

IV. Conclusions

In this paper, we propose an algorithm for image super-resolution based on deep convolutional neural networks. We build our framework upon the VDSR [2] method, and enhance the performance by adding dense connection to the model. We further propose to use wavelet transform to deal with different frequency bands separately and improve the performance. The proposed network is applicable to several scaling factors by augmenting the training data, and can be learned efficiently. Extensive experiments demonstrate the superiority of the proposed algorithm over state-of-the-art methods.

References

[1] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, issue 2, pp. 295–307, 2015.

[2] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *CVPR*, pp. 1646-1654, 2016.

[3] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *CVPR Workshops*, pp. 136-144, 2017.

[4] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *ECCV*, pp. 391-407, 2016.

[5] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Processing*, vol. 19, issue 11, pp. 2861–2873, 2010.

[6] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV*, pp. 1-10, 2001.

[7] C. G. M. Bevilacqua, A. Roumy, and M. L. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *BMVC*, artcile 135, pp. 1-10, 2012.

[8] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, pp. 711-730, 2012.

[9] J. B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution using transformed self-exemplars," in *CVPR*, pp. 5197-5206, 2015.

[10] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *ACCV*, pp. 111-126, 2014.

[11] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *CVPR*, pp. 1637-1645, 2016.

[12] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in *CVPR*, pp. 3147-3155, 2017.

[13] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. S. Twitter, "Photo-realistic single image super-resolution using a generative adversarial network," in *CVPR*, pp. 4681-4690, 2017.

Learning Based SLIC Superpixel Generation and Image Segmentation

Chieh-Sheng Chang
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
r06942042@ntu.edu.tw

Jian-Jiun Ding
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
jjding@ntu.edu.tw

Heng-Sheng Lin
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
r07942039@ntu.edu.tw

***Abstract*—Superpixel generation is to cluster the pixels with similar features and plays an important role for image segmentation. Conventional superpixel generation methods are more meaningful, however, the learning based method can generate the superpixels directly from the segments in the ground truth and achieve even better performance. In this work, an advanced superpixel generation algorithm that combines the advantages of conventional methods and modern neural network techniques is proposed. In addition to colors and locations, we find that the feature generated by neural networks also provide useful information for superpixel assignment. Simulations show that, with the proposed superpixels, a much more precise segmentation result can be achieved.**

Keywords—Superpixel, image segmentation, deep learning, feature map, neural network

I. Introduction

Superpixel generation is important for modern image segmentation techniques. It combines the pixels with similar colors and locations. There are several ways to generate superpixels, such as the simple linear iterative clustering (SLIC) method [1]. When using superpixels instead of pixels, the effect of tiny detail and noise can be much reduced and accurate image segmentation results can be achieved [2].

In recent years, deep learning is widely used in many fields about image processing. There are also some works that combine superpixels and the deep neural network together for saliency detection [3] and depth estimation [4]. However, directly creating superpixels by the neural network is more difficult due to the lack of differentiability.

A deep learning based superpixel algorithm was proposed in [5]. It transforms the into affinity maps. Then, a the pixel affinity net is built and trained on those maps to produce pixel affinities. Then, in [6], the superpixel sampling network (SSN) was proposed. In the SSN, a deep convolutional neural network is adopted to extract k-dimensional features. Then, a loss function is applied to avoid a superpixel overlapping two segments in the ground truth as possible.

Although features adopted by conventional superpixel generation methods are more meaningful, the SSN can learn the features from the segments in the ground truth directly and achieve a more accurate segmentation result.

In this paper, we proposed an advanced way to generate superpixels. It combines the modern SSN technique and the conventional superpixel generation method. In addition to the color and location, the output of the SSN also provides an important information for superpixel assignment. Simulations show that, with the proposed superpixel generation method, even better image segmentation results can be achieved.

II. Proposed Superpixel Generation Algorithm

The proposed algorithm is an improvement of the SNN [6]. Its architecture is shown in Figs. 1 and 2.

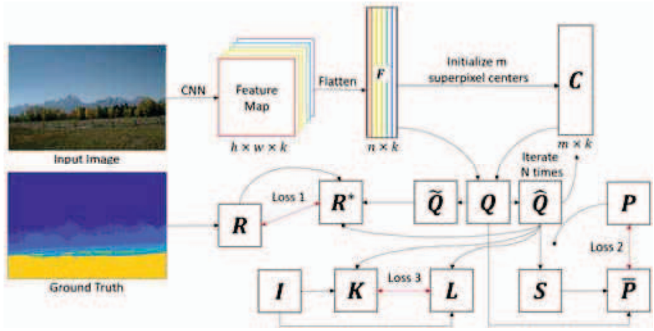


Fig. 1. Flowchart of the proposed method in training stage.

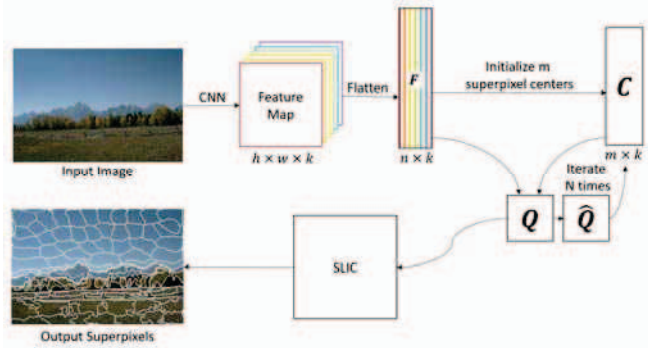


Fig. 2. Flowchart of the proposed method in testing stage.

Images are first input into a convolution neural network (CNN) and flattened into a feature matrix $\boldsymbol{F}$. In Fig. 1, n is the number of pixels, k is the number of features, m is the number of superpixels to be generated, and the $(j, k)^{\text{th}}$ entry of $\boldsymbol{C}$ is the mean value of the k^{th} feature in the j^{th} superpixel. Then, a soft pixel-cluster association matrix $\boldsymbol{Q}$ is assigned where

$$Q_{ij} = e^{-D_{ij}} \quad \text{where} \quad D_{ij} = \left\|\boldsymbol{F_i} - \boldsymbol{C_j}\right\|_2^2, \tag{1}$$

$\boldsymbol{F_i}$ is the i^{th} column of $\boldsymbol{F}$, and $\boldsymbol{C_j}$ is the j^{th} column of $\boldsymbol{C}$. If

$$j_1 = \arg\max_j Q_{ij}, \tag{2}$$

then the pixel i is predicted to be within the j_1^{th} superpixel. In Fig. 1, $\tilde{Q}$ and $\hat{\boldsymbol{Q}}$ are the row normalization and the column normalization versions of Q, respectively.

In the training stage, three loss functions are applied to back propagate through the network: (i) the reconstruction loss (the difference between the segmentation result R^* and the ground truth R), (ii) the compactness loss (determined according to whether the positions P of the pixels within a superpixel has smaller differences), and (iii) the proposed variance loss, which is described as follows.

We design the variance loss to keep the color variance within a superpixel low. We first construct a color feature matrix $\boldsymbol{I}_{(n\times 3)}$ which represents the CIELAB color space value of the n image pixels. Then, we set

$$\boldsymbol{K} = \hat{\boldsymbol{Q}}^T\boldsymbol{I}. \tag{3}$$

Note that K contains the mean color of each superpixel. Then

$$\boldsymbol{L} = \widehat{\boldsymbol{Q}}^{\boldsymbol{T}}(\boldsymbol{I} \circ \boldsymbol{I}) \tag{4}$$

where $\circ$ denotes the pixel-wise product and $\boldsymbol{L}$ can be viewed as the mean of the square color value of each superpixel. Since the variance is equal to the mean of the square minus the square of the mean, the proposed variance loss is given by

$$\boldsymbol{Var}(\boldsymbol{I}) = \boldsymbol{L} - (\boldsymbol{K} \circ \boldsymbol{K}). \tag{5}$$

Then, we propose a post-processing method that combines SLIC and the SSN together to improve the performance of superpixels, as in Fig. 2. Recall that the distance used for assigning pixels to their nearest clusters in SLIC is defined as

$$D = \sqrt{d_C^2 + (d_S / U)^2 m^2} \tag{6}$$

where d_C and d_S are color and spatial domain distances, respectively, U increases with the number of superpixels, and m is an adjustable weight. We modify the distance into

$$D' = \frac{1}{(Q_{ij})^{\alpha}} \sqrt{d_C^2 + (d_S / U)^2 m^2} \tag{7}$$

where Q_{ij} is determined from (1) and α controls the effect of the SSN. Specially, if $\alpha = 0$, the generated superpixels are all the same as those of SLIC. If $\alpha \rightarrow \infty$, the generated superpixels are the same as those of the SSN. Then, we assign the superpixel based on minimizing the distance D'.

Compared to the original SSN method, the proposed superpixel generation method has the following advantages:

(1) A new loss function is appended to reduce the convergence time during the training session.

(2) With the proposed variance loss and the modified distance D', the feature values within a superpixel are more uniform. It lead to even better segmentation results, as shown in Section III.

III. Simulations

In Table I, we evaluate the image segmentation performance on the Berkeley segmentation database [7]. The algorithm for merging superpixels is fixed to the SSN approach proposed in [8] but different superpixels are adopted. For the proposed superpixels, we also test the performance using different values of α in (7). The evaluation metrics of segmentation results include the probabilistic random index (PRI), the variation of information (VI), the global consistency error (GCE), and the boundary displacement error (BDE). An accurate segmentation result should have a *higher* PRI and *lower* values of the VI, the GCE, and the BDE.

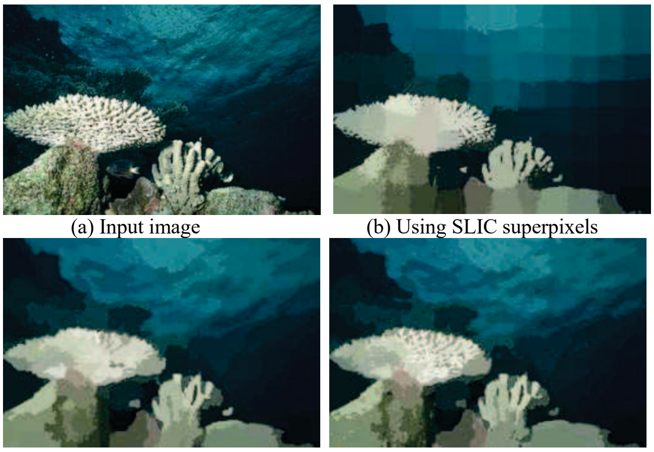
Moreover, some visual comparisons are given in Figs. 3 and 4. The results in Table 1 and Figs. 3 and 4 show that the proposed superpixel generation algorithm can lead to even better image segmentation results than using the SLIC method and the SSN. This is due to that, with the modified distance D' in (7), the feature values within a superpixel are more uniform.

IV. Conclusiom

In this paper, an advanced way to generate the superpixels is proposed. It combines the advantage of the conventional and deep learning based superpixel generation method and take both the variance within a superpixel and the features generated by the neural network into account. With it, an even more accurate segmentation result can be achieved.

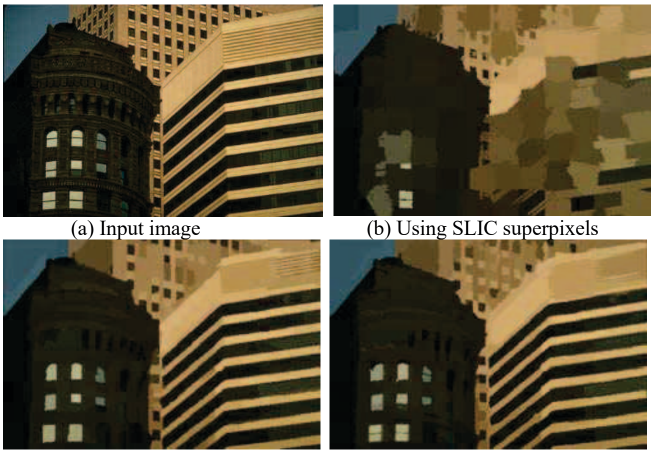
TABLE I. Image Segmentation Performance Using Different Pixels. A Better Segmentation Results Should Have a Higher PRI and Lower Values of the VI, the GCE, and the BDE.

	PRI	VI	GCE	BDE
SAS [8]	0.8319	1.6849	0.1779	11.29
SLIC [1] ($\alpha = 0$)	0.8157	1.6952	0.1784	11.77
Proposed ($\alpha = 0$)	0.8278	1.6744	0.1764	11.16
Proposed ($\alpha = 5$)	**0.8392**	**1.6465**	**0.1723**	**11.04**
Proposed ($\alpha = 10$)	0.8371	1.6563	0.1792	11.34
SSN [6] ($\alpha \rightarrow \infty$)	0.8266	1.6991	0.1811	11.46



(a) Input image (b) Using SLIC superpixels
(c) Using SSN superpixels (d) Using the proposed superpixels

Fig. 3 Visual comparison of image segmentation results.



(a) Input image (b) Using SLIC superpixels
(c) Using SSN superpixels (d) Using the proposed superpixels

Fig. 4 Visual comparison of image segmentation results.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “SLIC superpixels compared to state-of-the-art superpixel methods,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, pp. 2274-2282, 2012.

[2] C. Y. Hsu and J. J. Ding, “Efficient image segmentation algorithm using SLIC superpixels and boundary-focused region merging,” in Int. Conf. Information, Communications and Signal Processing, Tainan, Taiwan, pp. 1-5, Dec. 2013.

[3] R. Zhao, W. Ouyang, H. Li, and X. Wang, "Saliency detection by multi-context deep learning", in IEEE Conf. Computer Vision and Pattern Recognition, pp. 1265-1274, 2015.

[4] F. Liu, C. Shen, and G. Lin, "Deep convolutional neural fields for depth estimation from a single image,” in IEEE Conf. Computer Vision and Pattern Recognition, pp. 5162-5170, 2015.

[5] W. C. Tu, M. Y. Liu, V. Jampani, D. Sun, S. Y. Chien, M. H. Yang, and J. Kautz, "Learning superpixels with segmentation-aware affinity loss", in IEEE Conf. Computer Vision and Pattern Recognition, pp. 568-576, 2018.

[6] V. Jampani, D. Sun, M. Y. Liu, M. H. Yang, J. Kautz, “Superpixel Sampling Networks,” in European Conf. Computer Vision (ECCV), pp. 352-368, 2018.

[7] D. Martin, C. Fowlkes, D. Tal, and J. Malik. “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *IEEE Int. Conf. Computer Vision*, pp. 416–423, 2001.

[8] Z. Li, X. M. Wu, S. F. Chang, "Segmentation using superpixels: A bipartite graph partitioning approach,” in IEEE Conf. Computer Vision and Pattern Recognition, pp. 789-796, 2012.

Automatic Chinese Handwriting Verification Algorithm Using Deep Neural Networks

Chi-Chang Lee
Department of Engineering Science and Ocean Engineering
National Taiwan University
Taipei, Taiwan
b03505052@ntu.edu.tw

Jian-Jiun Ding
Graduate Institute of Communication Engineering
National Taiwan University
Taipei, Taiwan
jjding@ntu.edu.tw

***Abstract*—Handwriting verification is to identify whether a script was written by a person himself or forged. Conventional handwriting verification algorithms are based on feature extraction. However, the features of scripts are highly affected by the writing instrument, the posture, and the force of writing, even if the scripts were written by the same person, the extracted features will be quite different. Moreover, since some writers might not write some strokes clearly or ignore some strokes, not all features can be well extracted in every script. Therefore, in this paper, we apply a deep neural network based algorithm for handwriting verification. With the proposed algorithm, the parts that are really powerful and robust for handwriting verification can be highlighted by the auto-encoder. Then, a very high accurate handwriting verification result can be achieved.**

Keywords*—*handwriting verification, pattern recognition, convolution neural network, forensics, deep learning

I. Introduction

Automatic handwriting verification is to check whether a script was written by a person himself or forged by a computer program. It is a classification problem. Many existing handwriting verification algorithms are based on extracting the features of scripts, including Gabor features [1], the Hu & affine & Tsirikolias moment, the gray-level co-occurrence matrix [2], or the feature points using the scaled invariant feature transform (SIFT) [3]. Then, a feature-based classifier, such as the weighted Euclidean distance [4], the *k*-nearest neighbor method (*k*-NN) [5], unsupervised K means [6], and the support vector machine (SVM) [7], is applied to determine whether the script is genuine or forged.

However, there are some problems when applying a feature based algorithm on handwriting verification. First, most features are highly affected by the writing instrument, the writing gesture, the force, and the object beneath the paper. Therefore, it often happens that the features of two scripts are quite different even if they were written by the same person. Second, when writing a script, it usually happens that some strokes were skipped or not written in standard ways. Therefore, the number of features for two scripts may not be equal, which makes feature matching very difficult.

In this paper, we propose an automatic handwriting verification algorithm based on the deep neural network (DNN) [8] is proposed. With the DNN, no human-defined features should be extracted in prior. It uses an end-to-end method to learn the best filters that are helpful for distinguish forged scripts from the genuine ones. Then, a very high accurate handwriting verification result can be achieved.

II. Proposed Algorithm

Before performing handwriting verification, first, we remove the background by thresholding. Then, instead of binarization, we normalize the intensity of the script part to the range of [0, 1]. Since the variation of intensity can reflect how a person changed the force to write different part of the script, applying intensity normalization instead of binarization can preserve more information for handwriting verification.

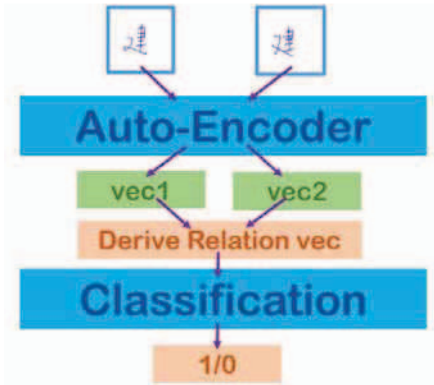


Fig. 1 Architecture of the auto-encoder for handwriting verification.

Then, we find the centroid and the size of the script part to perform alignment and rescaling before applying the network.

Then, the convolutional neural network (CNN) is applied. It can consider the relations among pixels and extract the patterns that are helpful for handwriting verification. Although the scripts written by the same person may have slight differences in stroke location, stroke orientation, and the relative size of the stroke, due to the pooling layer in the CNN architecture, the extracted patters will be robust to the small amount of translation, rotation, and scaling. Moreover, the pooling layer is also helpful for filtering the redundancy information and dimensionality reduction. Furthermore, the activation function adopted in the CNN is helpful for generating a nonlinear transformation and makes the outputs match the desired ones.

In the optimization process, the CNN determines the patterns for handwriting verification according to the distribution of labels. Moreover, since in practice there is usually not enough amount of training data for handwriting verification, we also adopt an unsupervised dimensionality reduction method, the auto-encoder [9], to avoid over-fitting and extract the features that are really helpful for handwriting verification, as in Fig. 1.

The auto-encoder consists of the encoder and the decoder parts. The encoder is applied to feature extraction. The selection of features is according to the loss function such as the mean square error. Compared to principle component analysis and linear discriminant analysis [10], the deep-based dimensionality reduction method applies nonlinear transformation. It can construct a more complicated model together with the activation function and increase the dependency among different elements.

The auto-encoder can extract some geometric patterns and refine them into the patterns that are meaningful for handwriting verification after several times of training. Compared to human-defined feature extraction methods, the auto-encoder can extract the patterns that are hard to define but really usually for classification.

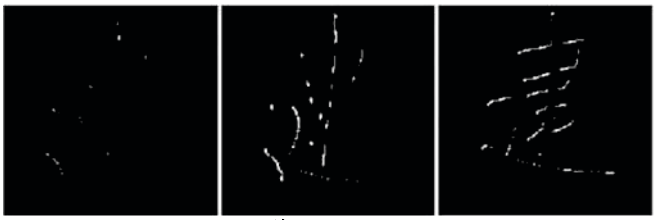


Fig. 2 The parts of the script "建" that are highlighted by the auto-encoder. They are considered to be helpful for distinguishing the genuine scripts from the forged ones.

In Fig. 2, we show some of the patterns highlighted by the auto-encoder. Note that these three subfigures are near to the cross points of stokes, the start / end / bent parts of stokes, and other parts of stokes, respectively. From the advisory from the forensic expert in Investigation Bureau, Taiwan, the variations of intensities these three parts are indeed the critical clues to identify whether the script was forged, since different persons may apply different amounts of force or writing styles on the start / end / bend part, the cross part, and other part of a stroke.

III. Simulations

In this section, some simulations are performed. We apply the test data in [4]. We apply the classifiers of the *k*-NN [5], the binary decision tree (BDT) [11], the K-means method [6], and the proposed DNN based architecture for handwriting verification. We apply the auto-encoder to the test data. Moreover, to avoid the overfitting problem caused from the limited amount of training data, we perform data augmentation to convert an encoder vector into a pair of vectors with the same square difference to the original one and balance the distribution of data in the training phase.

We test the accuracy for every Chinese words and show the means, maximums, minimums, and medians of accuracies in Table I. In Fig. 3, we show the distributions of accuracies for the four methods. The results show that, with the proposed DNN-based classifier, the accuracy of handwriting verification can be much improved.

IV. Conclusion and Future Work

In this paper, a DNN-based automatic handwriting verification algorithm was proposed. Conventionally, there are two problems that may affect the performance of handwriting verification. First, different scripts may have different styles, even if they were written by the same persons. It leads to that the features may not have smaller difference for the scripts written by the same person. Second, some strokes may be ignored or connected when people write a word. It makes that some of the features cannot be well extracted in every script and the number of extracted features may not be equivalent for every script. These limited the performance of feature-based handwriting verification methods.

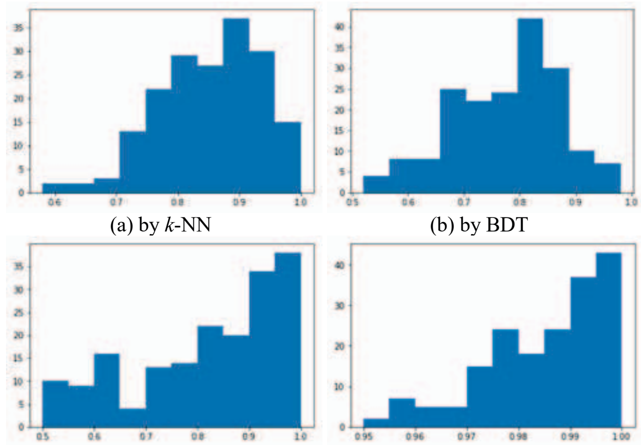
By contrast, the proposed DNN-based algorithm does not require any human-defined features and the patterns that are really helpful for classification are automatically learned by the auto-decode. Simulations show that, with the proposed method, a high accurate handwriting verification result can be achieved. In the future, we will try to apply the generative adversarial network (GAN) [12] to achieve even better performance for handwriting verification.

Acknowledge

The authors thank for Investigation Bureau, Taiwan, who provides advisory about manual handwriting verification methods.

TABLE I. Accuracies of Using Different Classifiers for Handwriting Verification. The Means, Maximums, Minimums, and Medians of Accuracies are Shown. (There are 180 Different Words).

	mean	max	min	median
by *k*-NN	0.849	0.98	0.58	0.9
by BDT	0.86	0.98	0.48	0.88
by K-means	0.816	1	0.865	0.816
by DNN (Proposed)	0.985	1	0.95	0.986



(a) by *k*-NN (b) by BDT

(c) by K-means (d) by DNN (Proposed)

Fig. 3 The distributions of accuracies using different classifiers for handwriting verification for 180 different Chinese words.

References

[1] C. Shen, X. G.Ruan, and T. L.Mao,"Writer identification using Gabor wavelet,"in IEEE World Congress on Intelligent Control and Automation, vol. 3, pp. 2061-2064, 2002.

[2] J. F. Vargas, M. A. Ferrer, C. M. Travieso, and J. B. Alonso, "Off-line signature verification based on grey level information using texture features," *Pattern Recognition*, vol. 44, issue 2, pp. 375-385, 2011.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, issue 2, pp. 91-110, 2004.

[4] P. X. Lee, J. J. Ding, T. C. Wang, and Y. C. Lee, "Automatic writer verification algorithm for Chinese characters using semi-global features and adaptive classifier," *IEEE Int. Conf. Multimedia and Expo, Multimedia & Expo Workshops*, San Diego, USA, pp. 1-6, July 2018.

[5] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in *OTM Confederated Int. Conf. on the Move to Meaningful Internet Systems,* Springer, Berlin, Heidelberg, pp. 986-996, Nov. 2003.

[6] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A.Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," in IEEE Trans. Pattern Anal. Mach. Intell., vol. 7, pp. 881-892, 2002.

[7] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intelligent Systems and Technology, vol. 2, issue 3, article 27, 2011.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, vol. 25, pp. 1097-1105, 2012.

[9] I. Guyon, G. Dror, V. Lemaire, G. Taylor, and D. Silver, "Autoencoders, unsupervised learning, and deep architectures," in Proc. ICML Workshop on Unsupervised and Transfer Learning, vol. 27, pp. 37–49, 2012.

[10] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. R Mullers, "Fisher discriminant analysis with kernels," in Neural Networks for Signal Processing IX: Proc. IEEE Signal Processing Society Workshop, pp. 41-48, Aug. 1999.

[11] J. R. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, issue 1, pp. 81-106, 1986.

[12] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks" in IEEE Int. Conf. Computer Vision, pp. 2223-2232, 2017.

Learning Based Noise Identification Techniques Using Time-Frequency Analysis and the U-Net

Chih-Hao Wang [1], Jian-Jiun Ding [2], Chieh-Sheng Chang [3]
Graduate Inst. Communication Engineering,
National Taiwan University
Taipei, Taiwan
r06942116@ntu.edu.tw [1], jjding@ntu.edu.tw [2],
r06942042@ntu.edu.tw [3]

Liang-Yu Ouyang
National Chung-Shan Institute of Science and Technology
Taipei, Taiwan
d98942009@ntu.edu.tw

***Abstract*—In wireless communication, it is inevitable that the signal is highly affected by the noise. For example, for the radar located in the sea shore, due to the effects of sea clutter and remote detection range, the signal to noise ratio (SNR) is only about 0~15 dB. In this manuscript, we develop an advanced noise determination and removal algorithm based on the deep learning method of the U-net. The U-net is a pixel-wise classification network and widely used in image segmentation. In this work, we find that it is also an effective way to determine whether a pixel in the time-frequency domain is the signal part or the noise part, even in the low SNR case. It is very helpful for reducing the noise effect and improving the accuracy of fundamental frequency analysis for radar signal processing.**

Keywords—noise removal; U-net; time-frequency analysis; radar signal; deep learning

I. Introduction

In wireless communication, it only happens that the signal is highly affected by noise. In some conditions, the energy of the noise is even near to that of the signal. For example , if the radar is located near the seashore, then, due to the effects of the sea clutter and remote detection range, the SNR is usually in the range of 0~15dB [1, 2]. Thus, it is important to develop an effective noise identification and removal algorithm.

Conventionally, the noise is removed by the filters that contain some parameters. For example, the lowpass filter is able to remove the noise in the high frequency part. However, one should determine the cutoff frequency in prior. When designing the bilateral filter, the standard deviations of two Gaussian filters should be chosen in prior [3]. When using the principal component analysis based method, the parameters for reference patch selection should also be assigned.

In this paper, we propose a deep learning based method that applies the U-net [4] in the time-frequency plane for noise removal. When using the U-net, no parameter assignment is required and good noise removal result can be achieved with enough augmentation training data. Simulations show that the U-net based method can well remove the noise and estimate the fundamental frequency of the input periodic signal accurately even in the low SNR case.

II. Preliminary

Since the proposed algorithm is based on time-frequency analysis and the U-net, we first give a brief description about them. Time-frequency analysis is to analyze how the instantaneous frequency of a signal varies with time. There are several time-frequency analysis methods. One of the most famous methods is the short-time Fourier transform (STFT):

$$X(t,f)=\int_{-\infty}^{\infty}x(\tau)w(t-\tau)e^{-j2\pi f\tau}d\tau \qquad (1)$$

where the window $w(t)$ is a time-limited function. Specially, if $w(t)$ is a Gaussian function $\exp(-\pi\sigma t^2)$, then the STFT is called the Gabor transform [5].

Fig. 1. Flowchart of the proposed U-net based noise removal architecture.

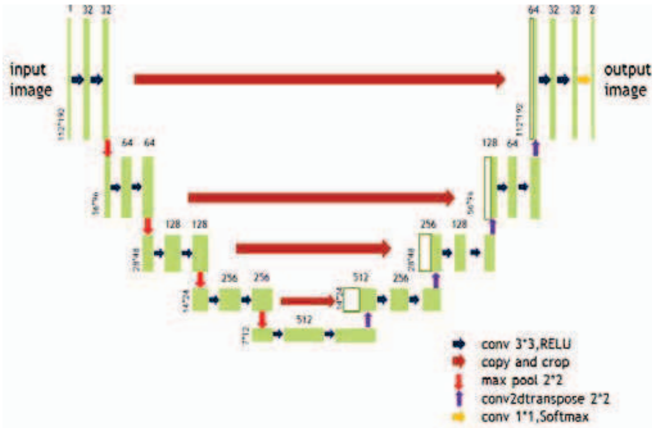


Fig. 2. The adopted modified U-net.

The U-net is an advanced deep learning method that consists of several pooling and up-sampling processes [4]. Since its architecture is similar to the letter "U", it is called the U-net. Since the output of the U-net is pixel-wise (each pixel is assigned a label) and the features with different scales are applied to determine the output, the U-net is usually applied for pixel-wise classification problems, such as image segmentation and medical image processing.

III. Proposed Algorithm

In this paper, we apply the U-net in the time-frequency plane to distinguish the noise part from the signal part and determine the fundamental frequency of a periodic signal. The flowchart of the proposed algorithm is as in Fig. 1.

First, since the accuracy of a deep-learning based architecture is highly dependent on the amount of training data, it is important to acquire enough signal+noise data in the training phase. We try to simulate the signal as follows:

$$y(t)=x_1(t)+x_2(t)+noise \qquad (2)$$

where $$x_1(t)=\sin(2\pi ft)\sum_n \Lambda(t-n/(2f)), \qquad (3)$$

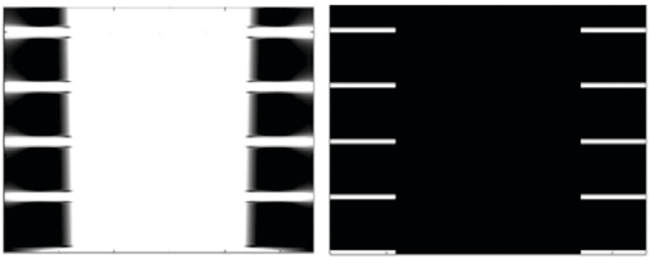
and $\Lambda(t)$ is a pulse that contains a lot of high frequency component. Note that that $x_1(t)$ is periodic and satisfies

$$x_1(t)=x_1(t+1/f). \qquad (4)$$

$x_2(t)$ is a middle frequency signal:

$$X_2(f)\cong 0 \quad \text{when } f<f_1 \text{ and } f>f_2 \qquad (5)$$

where $X_2(f)$ is the Fourier transform of $x_2(t)$ and f_1 and f_2 are the two cutoff frequencies of the middle frequency signal $x_2(t)$. We apply the signal with the form as in (2) for simulations because it is similar to a radar signal reflected from a periodically rotated target in a noisy scenario [1].



(a) (b)

Fig. 3 (a) The Gabor transform of $y(t)$ in (2) when there is no noise. (b) The ground truth of the periodic signal part

When there is no noise, the Gabor transform of $y(t)$ is as in Fig. 3.

We then generate a huge amount of training data according to (2) with different fundamental frequencies f and different styles, magnitudes, and variances of noise. Then, we fed them into the U-net to train a powerful model to distinguish the noise part from the signal part in the time-frequency domain.

However, to perform training, we modify the original U-net architecture a little. Considering that the overfitting problem may occur during the training process, we have to adjust the feature channels of each layer of the U-Net to half. In addition, we also change the original up-sampling layers to the deconvolution layers. Moreover, in the deconvolution layer, the transposed convolution is applied. The transposed convolution and deconvolution can restore the image to its original size. The transposed convolution layer still perform the normal convolution operation. However, it also performs different kinds of padding to make the size of the output image same as that the input image.

Based on the above changes, we can get the architecture diagram of the modified U-net as in Fig. 2. Note that the number on the top of the green rectangles is the number of feature channels and the lower left number is the size of the input image. After pre-processing and adjusting the U-Net architecture, we can start training based on the large amount of training data. After identifying the signal part from the modified U-net architecture, we calculate

$$y_p(t) = \int Y_s(t,f)df \, , \tag{6}$$

$$y_c(t) = y_p(t) * y_p^*(t) = \int y_p(\tau) y_p(t+\tau) d\tau \tag{7}$$

where $Y_s(t, f)$ is the time-frequency distribution of $y(t)$ after noise removal and * means convolution. Then, the fundamental frequency of the periodic part (i.e., $x_1(t)$ in (2)) can be estimated from the lowest local maximum of $|Y_c(f)|$ where $Y_c(f)$ is the Fourier transform of $y_c(t)$.

IV. Simulations

In this section, some simulations are performed. After training the modified U-net architecture, we apply 4 test signals with different SNRs. For test signals 5 and 6, the SNR is lower than 15dB. Moreover, we also compare the proposed architecture and the noise removal results using the support vector machine (SVM) [6], and the k nearest neighborhood (kNN) method [7]. The fundamental frequency estimation results are shown in Table 1. In Figs. 4-6, we also show the Gabor transforms of the signals before and after noise removal. The results show that the proposed modified U-net based algorithm can well distinguish the noise part from the signal part and well determine the fundamental frequency of a periodic signal, even if it is severely interfered by the noise.

TABLE I. The Results of Fundamental Frequency Prediction Using the Proposed Noise Removal Method Based on the Modified U-Net and Other Methods

	SNR	Ground Truth	SVM	kNN	Modified U-Net
Test Signal 1	26.29dB	10Hz	9.35Hz	9.35Hz	9.43Hz
Test Signal 2	24.70dB	16Hz	13.09Hz	13.09Hz	13.21Hz
Test Signal 3	24.35dB	10Hz	16.83Hz	16.83Hz	9.43Hz
Test Signal 4	23.37dB	12.5Hz	11.22Hz	11.22Hz	11.32Hz
Test Signal 5	10.42dB	16Hz	5.61Hz	5.61Hz	13.21Hz
Test Signal 6	9.26dB	16Hz	5.61Hz	13.09Hz	13.21Hz

(a) T-F plot of the input (b) Noise removal by the U-net

Fig. 4 Using the proposed algorithm to remove the noise for Signal 2.

(a) T-F plot of the input (b) Noise removal by the U-net

Fig. 5 Using the proposed algorithm to remove the noise for Signal 3.

(a) T-F plot of the input (b) Noise removal by the U-net

Fig. 6 Using the proposed algorithm to remove the noise for Signal 4.

Acknowledge

The authors thank for the support of National Chung-Shan Institute of Science and Technology, Taiwan.

V. Conclusion

In this paper, a noise removal algorithm that applies the deep learning method of the U-net and time-frequency analysis is proposed. With the proposed algorithm, the noise part of a signal can be well estimated and the fundamental frequencies of periodic signals can be accurately estimated even in the very low SNR case.

References

[1] P. Tait, Introduction to Radar Target Recognition, The Institution of Engineering and Technology, Jan 2006.

[2] L. Ye, D. Xia, W. Guo, “Comparison and analysis of radar sea clutter k distribution sequence model simulation based on zmnl and sirp,” Modeling and Simulation, vol., 7, issue 1, article 8, 2018.

[3] M. Elad, "On the origin of the bilateral filter and ways to improve it," IEEE Trans. Image Processing, vol. 11, issue 10, pp. 1141-1151, 2002.

[4] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Int. Conf. Medical Image Computing and Computer-Assisted Intervention, pp. 234-241, Oct. 2015.

[5] S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way, Academic Press, 3rd ed., 2009.

[6] C. C. Chang and C. J. Lin. "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Tech., vol. 2, article 27, 2011.

[7] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, “KNN model-based approach in classification,” in *OTM Confederated Int. Conf. on the Move to Meaningful Internet Systems,* Springer, Berlin, Heidelberg, pp. 986-996, Nov. 2003.

Autonomous Vehicle Trajectory Planning and Control Based on Traffic Motion Prediction

Yuho Song
Department of Automotive Engineering
Hanyang University
Seoul, Korea
song4405@hanyang.ac.kr

Dongchan Kim
Department of Automotive Engineering
Hanyang University
Seoul, Korea
jookker@hanyang.ac.kr

Kunsoo Huh
Department of Automotive Engineering
Hanyang University
Seoul, Korea
khuh2@hanyang.ac.kr

***Abstract*—The autonomous vehicle should be properly operated according to the road geometry and behavior of the surrounding vehicles. For this purpose, we propose a supervisor that predicts the motion of the surrounding vehicle using the artificial potential field method, determines the collision risk of the vehicle with the collision check algorithm and generates the velocity distribution map considering the velocity of the surrounding vehicles. In addition, the trajectory planner is designed to minimize the jerk for passenger comfort, considering the road shape and the velocity distribution determined by the supervisor. Finally, the nonlinear model predictive controller is formulated considering the limit of the actuator and collision avoidance.**

I. INTRODUCTION

Currently, the autonomous vehicle technology covers the extended level of driving workload. But, there is an opinion that it is premature to move on to higher self-driving level [1]. The main argument is that the self-driving system has not been fully developed. Since the roads are still occupied by human-driven vehicles, the autonomous vehicle should be developed considering the human-driven vehicles. Autonomous vehicle system usually consists of the localization, perception, supervisor, trajectory planner, and a tracking controller [2]. Each component is interconnected.

In this paper, an autonomous vehicle control system is proposed including a supervisor, motion planning, and tracking controller. The supervisor includes predicting traffic motion with artificial dynamics, assessing collision risk with geometric collision check, and determining the desired cruise speed using velocity field. The trajectory planner is designed to determine the desired trajectory and velocity profile using risk and velocity information from the supervisor. Finally, the tracking controller is formulated based on nonlinear vehicle dynamics and surrounding vehicle constraints.

II. VEHICLE MOTION CONTROL SYSTEM

The proposed self-driving system architecture is shown in Fig. 1. It is assumed that the host vehicle is equipped with a built-in speed controller and active steering. The threatening

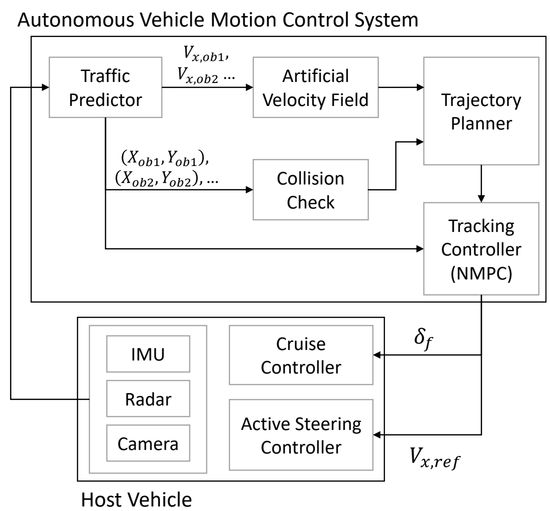


Fig. 1. System architecture

level of the surrounding vehicle is evaluated with the geometric collision detection algorithm. Then, utilizing the information from the supervisor, the trajectory planner generates the desired trajectory considering the comport and collision risk. Finally, the tracking controller drives the host vehicle with steering and in-vehicle speed controller.

III. SUPERVISOR

In order to determine the collision-free trajectory and speed of the autonomous vehicle, traffic motion prediction is considered. Dynamics of a surrounding vehicle is described utilizing the artificial potential field (APF) of the road. The APF of the road consists of the repulsive road boundary potential and attractive lane center potential.

The collision risk needs to be evaluated not only by considering the future position of the surrounding vehicles but also by considering its shape. Because the vehicle shape can be described as a rectangle, the simplified version of the Gilbert-Johnson-Keerthi (GJK) algorithm is utilized as shown in Fig. 2 [3].

The velocity field is defined with a reference velocity map near the surrounding vehicle and speed limit. The safe speed

This work was supported by the Industrial Strategic Technology Development Program(10079730, Development and Evaluation of Automated Driving Systems for Motorway and City Road) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea)

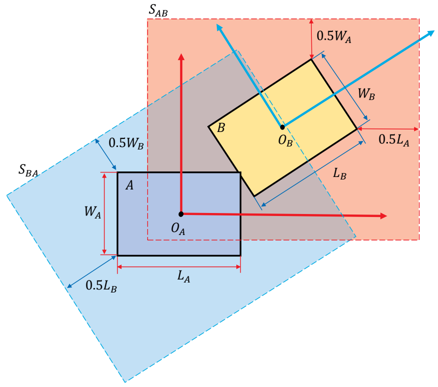


Fig. 2. Collision checked by the simplified GJK algorithm

of the ego vehicle and velocities of surrounding vehicles are determined depending on the road geometry. Because the road shape is not consistent, it is difficult to define the velocity field in the Cartesian coordinate system. Instead, the velocity field is described on the curvilinear coordinate.

IV. Trajectory Planner

Longitudinal movement of the ego vehicle is obtained using a constant acceleration set. Here, longitudinal velocity is assumed equal to the s-axis velocity and has the same direction as the road. When trajectory planning is carried out, there are several restrictions to exclude some invalid trajectory candidates. Firstly, collision check with surrounding vehicles is performed using the simplified GJK algorithm explained in Section III.Next, longitudinal and lateral trajectories are generated in the curvilinear coordinates. The acceleration candidate set is used to define s-axis velocity. Also, the lateral trajectory candidates are generated by utilizing a cubic polynomial function.

To select optimal trajectory, a cost function is constructed with respect to the longitudinal jerk, lateral jerk, road APF, velocity field, and consistency of the consecutive paths. Collision-free trajectory against the cut-in vehicle and overtaking trajectory results are demonstrated. The selected optimal trajectory through the cost function is transformed back to the Cartesian coordinates for the control part in Section V.

V. Tracking controller

The optimal control input for the nonlinear model predictive control (NMPC) is calculated considering the nonlinear vehicle dynamics and actuator limit. In order to predict the trajectory of the vehicle, a model is needed to represent the position of the ego vehicle with respect to the fixed coordinate defined at every control period. The lateral dynamics is described as a bicycle model [4], and the speed-controlled longitudinal dynamics is included with the 1^{st} order delay system. The vehicle planar motion dynamics is discretized utilizing the Euler method.

NMPC is designed considering nonlinearity of the vehicle model and non-convexity of the collision risk. At first, the

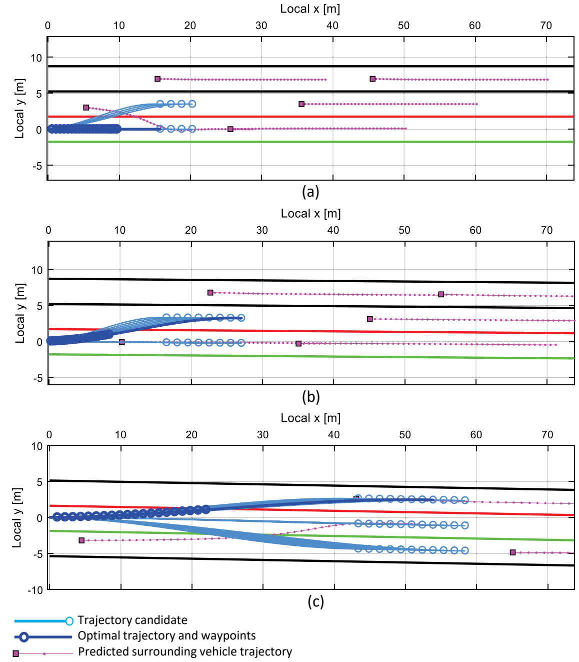


Fig. 3. Trajectory planning results: (a) cut-in vehicle collision avoidance with brake (b) overtaking slow vehicle (c) predicting front vehicle cut-in and collision avoidance with steering maneuver

objective function is defined to track the reference position and target velocity profile from the trajectory planner. The defined NMPC problem is solved by Primal-dual interior point algorithm [5]. The optimal path planning and control results are shown in Fig. 3.

VI. Conclusions

In this paper, an autonomous vehicle control system is constructed utilizing the risk predictive supervisor, trajectory planner, and NMPC. The trajectory planner determines the desired position and velocity profile using the evaluated risk and the velocity field from the supervisor. Finally, NMPC calculates the optimal steering and acceleration considering the vehicle dynamics, actuator limit, and surrounding vehicles. In the future project, more sophisticated supervisor and trajectory planner will be implemented, and NMPC will be reorganized considering the real-time performance.

References

[1] T. Litman, *Autonomous vehicle implementation predictions*. Victoria Transport Policy Institute Victoria, Canada, 2017.

[2] J. Van Brummelen, M. OBrien, D. Gruyer, and H. Najjaran, “Autonomous vehicle perception: The technology of today and tomorrow,” *Transportation research part C: emerging technologies*, 2018.

[3] E. G. Gilbert, D. W. Johnson, and S. S. Keerthi, “A fast procedure for computing the distance between complex objects in three-dimensional space,” *IEEE Journal on Robotics and Automation*, vol. 4, no. 2, pp. 193–203, 1988.

[4] R. N. Jazar, *Vehicle dynamics: theory and application*. Springer, 2017.

[5] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

Implementation and Evaluation of CNN Based Traffic Sign Detection with Different Resolutions

Ying-Chi Chiu and Huei-Yung Lin
National Chung Cheng University
Chiayi 621, Taiwan
p692584k@gmail.com, lin@ee.ccu.edu.tw

Wen-Lung Tai
Create Electronic Optical Co., LTD
New Taipei 23557, Taiwan
dennydai@create-eop.com.tw

***Abstract*—In this paper, we develop a traffic sign detection and recognition technique based on the YOLOv3 framework. The public traffic dataset TT100K and our own Taiwan road scene image dataset are used for network training. In the performance evaluation, we have extensively compared different input image resolution and ROI for network training and testing. In the experiments, the proposed method is able to achieve the mAP of 92.99% with 21 most common traffic signs on the road.**

***Index Terms*—traffic sign detection, recognition, CNN**

I. Introduction

In the advanced driver assistance system (ADAS), the detection and identification of traffic signs is very important. It provides the necessary information for a vehicle to follow the traffic regulations. There have been many traditional methods in the past using image features such as color, shape and gradients for traffic sign detection and recognition [1]. Recently, machine learning approaches have successfully achieved significant progress on object recognition [2]. This paper presents a road sign detection method using convolutional neural networks. We adopt YOLOv3 [3] as the detection framework and evaluate with different image resolutions.

Due to the difference of traffic signs in different countries and the number of available samples, in addition to the use of public datasets such as TT100K [4], we also collect our own traffic sign images for training and testing. In our dataset, the images are used with the resolution of 512×512 and 2048×2048. The features of high resolution 2048×2048 images might be suppressed after resizing for YOLOv3 input, while the features of 512×512 images can be preserved. We set 23 and 21 classes for traffic sign recognition for the high and low resolution images respectively, based on the frequency of appearance in the dataset. The class will be gradually increased as the image data and annotation grow in the future.

II. Method

This work adopts YOLOv3, it has relatively good detection and classification results, and the training time is much less than other deep neural networks (such as ResNet-101). The network can also generate different weight files which can be used to compare the accuracy easily. The YOLOv3 network architecture uses successive 3×3 and 1×1 convolutional layers with some shortcut connections. It is better utilized for GPU, and thus faster and efficient to evaluate compared to other network structures. The comparison with ResNet-101 and ResNet-152 is shown Table I.

The support of this work in part by Create Electronic Optical Co., LTD, Taiwan, is gratefully acknowledged.

TABLE I
THE COMPARISON OF SEVERAL BACKBONES: ACCURACY, BILLION OF OPERATIONS, BILLION FLOATING POINT OPERATIONS PER SECOND, AND FPS OF VARIOUS NETWORKS.

Backbone	Top-1	Top-5	Bn Ops	BFLOP/s	fps
Darknet-19 [5]	74.1	91.8	7.29	1246	**171**
ResNet-101 [6]	77.1	93.7	19.7	1039	53
ResNet-152 [6]	**77.6**	**93.8**	29.4	1090	37
Darknet-53 [3]	77.2	**93.8**	18.7	**1457**	78

The system predicts bounding boxes using dimension clusters as anchor boxes [5]. It predicts 4 coordinates for each bounding box. An objectness score for each bounding box is predicted using logistic regression. If a bounding box prior overlaps a ground truth object by more than any other bounding box prior, then the score is 1. Each bounding box predicts the class using multi-label classification. YOLOv3 does not use a softmax since it is not necessary for good performance. Instead, independent logistic classifiers are adopted. During network training, YOLOv3 uses binary cross-entropy loss for class prediction. A multi-label approach is used to better model the data.

III. Experiments

The experiment is first carried out using the original image size of 2048×2048, but the mAP is low due to resizing the images as inputs to the network. We then crop the images in 512×512 for network training and testing, and use TT100K dataset [4]. To make the approach suitable for common Taiwan road scenes, we add 2 more traffic sign types to increase the dataset to 23 classes. In the experiments, still images and videos are used for testing. The still image test on TT100K is carried out with the resolution of 2048×2048 and 512×512, and the video test on our own dataset is performed using the image size of 1280×750 and 612×512.

Table II shows the mAP results of different training and testing image sizes. If the training images are large (e.g. 2048×2048) and the testing images are small (e.g.608×608 or 1024×1024), the mAP drops due to the features suppressed. While

TABLE II
THE MAPS OF DIFFERENT TRAINING AND TESTING IMAGE SIZES.

Number of classes	Training size	Testing size	mAP
23	2048×2048	608×608	48.44%
23	2048×2048	1024×1024	47.35%
21	512×512	512×512	**92.99%**

TABLE III
THE MAPS OF DIFFERENT TESTING IMAGE SIZES WITH OR WITHOUT RESIZING USING THE NETWORK TRAINED BY 512×512 IMAGES.

Number of classes	Testing size	Input resize	mAP
21	512×512	512×512	92.99%
21	2048×2048	512×512	77.72%
21	2048×2048	1024×1024	83.11%

the resolution of the training and testing images is equivalent, the mAP is much higher as 92.99%. Table III shows the mAP results of different testing image sizes with or without resizing using the network trained by 512×512 images. In the case of 2048×2048 images, resizing to a smaller resolution (say, 512×512) results in a lower mAP (77.72% compared to 83.11% of the 1024×1024 input).

Table IV shows the video test results. The resolution of input video sequences is 1280×750, and an additional ROI for testing is set as 612×512, as shown in Fig. 1. The mAP for the original input size is high, but it requires more computation and results in the processing speed of 8.9 fps. If the input is resized to 512×512, the fps can increase to 30 but the mAP dropped significantly. Finally, there is a good balance between the processing time and mAP for the input ROI of 612×512.

Further tests are carried out with more annotated Taiwan road scene images. The dataset is increased by 1,686 images to a total of 31,045 training samples. Fig. 2 shows the results with and without additional training data of Taiwan road scene images. The mAP is improved significantly from 25.98% to 97.52%. Fig. 3 shows the results on test videos. Although the mAP is generally lower than the image cases, it is improved with the training data from Taiwan road scenes.

TABLE IV
THE VIDEO TEST RESULTS WITH DIFFERENT ROIS.

Video Input ROI	Input resize	mAP	fps
1280×750	1024×1024	37.85%	8.9
1280×750	512×512	8.26%	30
612×512	512×512	30.9%	30

Fig. 1. Left: The original image resolution of the test videos is 1280×750. Right: The ROI of 612×512 is used to improve the detection rate and fps.

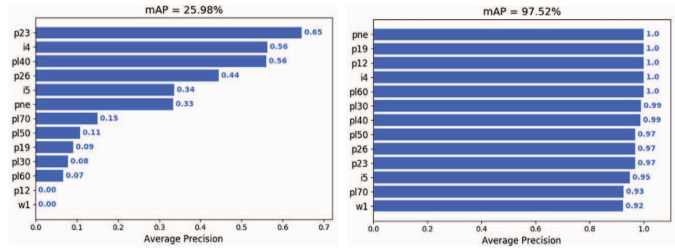
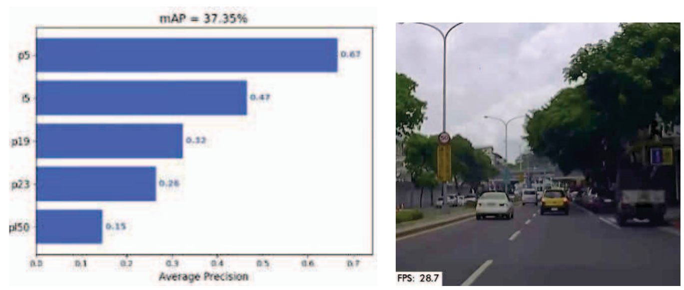
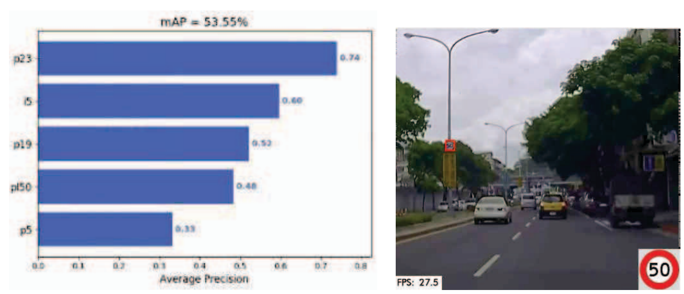


Fig. 2. Left: Only TT100K is used as the training data. Right: Additional Taiwan road scene images are used for training.



(a) (b)



(c) (d)

Fig. 3. The video test results. (a) and (b) are from the TT100K training dataset. (c) and (d) are the results from including the Taiwan road scene images for training.

REFERENCES

[1] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler, "A system for traffic sign detection, tracking, and recognition using color, shape, and motion information," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, June 2005, pp. 255–260.

[2] H. Luo, Y. Yang, B. Tong, F. Wu, and B. Fan, "Traffic sign recognition using a multi-task convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1–12, 2017.

[3] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *CoRR*, vol. abs/1804.02767, 2018.

[4] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-sign detection and classification in the wild," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 2110–2118.

[5] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, 2017, pp. 6517–6525.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 770–778.

Traffic Light Detection using Convolutional Neural Networks and Lidar Data

Tien-Wen Yeh, Ssu-Yun Lin and Huei-Yung Lin
Department of Electrical Engineering
National Chung Cheng University
Chiayi 621, Taiwan
{tw950107,sara850418}@gmail.com, lin@ee.ccu.edu.tw

Sheng-Wei Chan, Che-Tsung Lin and Yan-Yu Lin
Mechanical and Mechatronics Systems Research Laboratories
Industrial Technology Research Institute
Hsinchu 31040, Taiwan
{shengweichan,AlexLin,YanYuLin}@itri.org.tw

***Abstract*—This paper presents a traffic light detection system based on convolutional neural networks and lidar data. The proposed approach contains two stages. In the first detection stage, the map information is adopted to assist the object detection, and two cameras with different focal lengths are used to detect traffic lights at different distances. In the following recognition stage, we combine detector and classifier to deal with the problem of many light states. A dataset is created with urban road scenes in Taiwan. The experiments have demonstrated the advance in traffic light detection and recognition.**

***Index Terms*—traffic light detection, state recognition, ADAS**

I. Introduction

The techniques for traffic light detection are mainly divided into two approaches. The first method is to make traffic lights have communication capabilities with vehicles through V2I (vehicle-to-infrastructure) [1]. The second one is to detect the position and state of the traffic lights by vehicle on-board sensors. These two methods have both pros and cons. However, the first one requires the installation or replacement of basic equipment, which are normally expensive and time-consuming. The sensor based method has been developed for a long time [2]. However, the outdoor environment has many effects in detection, which results in a limited accuracy. With the success of deep learning in recently years, researchers have applied convolutional neural networks (CNNs) to traffic light detection [3] to cope with these problems.

In this work, we present a traffic light detection system for Taiwan road scenes using CNNs and lidar data. There are mainly two technical challenges in our application scenarios. First, the public datasets currently available contain the traffic lights arranged vertically, which are different from the horizontally arranged case we have to deal with. Second, the arrow lights for the direction instruction are very common in Taiwan's traffic scenes, but most existing works use classifiers to recognize the circle lights only. These issues lead to the solution of collecting the traffic data by ourselves. Moreover, data unbalance among multiple classes due to the lack of sufficient arrow light images further complicates the network training. Thus, a method by combining the object detector and classifier for light state recognition is proposed.

The support of this work in part by Industrial Technology Research Institute, Taiwan, is gratefully acknowledged.

II. Method

The proposed technique is to detect the traffic lights in Taiwan road scenes. After investigating the public datasets, there are no suitable resources specifically for our purpose. In this work, we cooperate with ITRI/MMSL to collect our own traffic dataset on two commuter routes. One is from ITRI to THSR Hsinchu station (16 km, 40 minutes), and the other is from CCU to THSR Chiayi station (30 km, 50 minutes). Two cameras, with 3.5mm and 12mm lenses respectively, are mounted below the rear view mirror for video recording. The images are captured at 36 fps with the resolution of 2048×1536. The lidar data are also acquired, which are used to provide the traffic light distances. We also use the HD map with annotated information about the traffic lights, such as ID, position, horizontal and vertical angles.

The first route is recorded three times, and the second route is recorded once. We pick 5 images per second, and label the positions of the traffic lights and the classes of the light states. There are totally 26,868 frames and 29,963 traffic lights labeled. The traffic lights are divided into 14 classes. When the vehicle is driving, the position between the vehicle and the traffic light can be determined by the lidar data and HD map. We use this information to crop the image, and find the accurate position of the traffic light and the light state.

Our method is a two-stage approach. The first stage detects the position of the traffic light, and the second stage recognizes the light state. We adopt YOLOv3 as the first stage traffic light detector [4] since the accuracy is fairly good with real-time processing capability. For the second stage, we adopt YOLOv3-tiny to detect the light states (red, green, yellow, and arrows), followed by LeNet to distinguish the arrows with different directions. This approach is expected to solve the data class unbalance problem. In addition, it is also flexible for the detection of less common light states.

Since the network detection is based on the cropped images, the training data are also cropped to simulate the lidar processing. Fig. 1 shows some sample images. Each traffic light image is cropped 3 times, and we let the traffic light randomly distributed at a different position of the cropped image. Because our dataset mainly contain red, yellow, green, straight, and straight-right classes, we employ data augmentation to

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Fig. 1. Examples of the cropped training sample.

Fig. 2. Examples of data augmentation images with arrow lights.

generate more training images of other classes. One important case is the rotation of arrow lights as shown in Fig. 2.

III. Experimental Results

Table I shows the comparison of two traffic detection techniques using our datasets. The first method uses YOLOv3 to detect the traffic lights and AlexNet [5] to classify the light states. The second method is presented in this paper. In addition to traffic light detection using YOLOv3, it adopts YOLOv3-tiny for light state recognition and LeNet for arrow light classification [6]. The table shows that our approach with YOLOv3+YOLOv3-tiny+LeNet outperforms YOLOv3+AlexNet, but requires more computation time.

Table II shows the precision, recall, mAP, and F1-score for each stage of the network. The values of 'State' and 'Type' are based on the results of the previous stage. There are larger errors in 'State' classification, especially for the Green class. This could be caused by its similar appearance to the arrow lights at a far away distance. Table III shows the precision, recall, mAP, and F1-score of the network final outputs. The classes with sufficient training samples such as red, green and yellow lights, the mAPs are higher as expected. For the classes with insufficient training data, the results are not improved even with data augmentation.

Finally, we compare the results with or without the map information. If the HD map and lidar data used to provide a rough ROI in the image, the mAP is 0.97. It is significantly higher than the result of 0.78 without the HD map and lidar information. Some of the traffic light detection results are shown in Fig. 3.

TABLE II
The mAP for each stage of the network.

	Class	Precision	Recall	mAP	F1-Score
Detection	Traffic Light	0.19	0.99	0.97	0.31
State	Red	0.44	0.96	0.93	0.60
	Yellow	0.15	0.98	0.90	0.26
	Green	0.07	0.97	0.64	0.14
	Arrow	0.15	0.97	0.91	0.26
Type	Left	0.92	0.89	0.87	0.90
	Straight	0.98	0.98	0.98	0.98
	Right	0.96	0.98	0.97	0.97

TABLE III
The mAP for each class.

Class	Precision	Recall	mAP	F1-Score
All	0.74	0.77	0.67	0.76
Close	0.59	0.67	0.43	0.59
Red	0.96	0.80	0.78	0.87
Yellow	0.84	0.89	0.79	0.87
Green	0.97	0.78	0.76	0.86
Left	N/A	N/A	N/A	N/A
Straight	0.73	0.63	0.55	0.67
Right	N/A	N/A	N/A	N/A
RedLeft	0.62	0.76	0.55	0.63
RedRight	0.62	0.59	0.45	0.57
StraightLeft	0.61	0.82	0.64	0.69
StraightRight	0.94	0.89	0.87	0.92
LeftRight	0.89	0.92	0.84	0.90
RedRightLeft	N/A	N/A	N/A	N/A
StraightRightLeft	0.91	0.74	0.69	0.81

Fig. 3. Some of the traffic light detection results. (See also https://www.youtube.com/watch?v=BFr4COeB_4w&feature=youtu.be)

References

[1] K. Abboud, H. A. Omar, and W. Zhuang, "Interworking of dsrc and cellular network technologies for v2x communications: A survey," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 9457–9470, 2016.

TABLE I
Test result comparison on our dataset.

Network	YOLOv3+AlexNet	YOLOv3+YOLOv3-tiny+LeNet
Precision	0.54	0.72
Recall	0.41	0.61
mAP	0.32	0.49
F1-Score	0.40	0.62
Speed (ms)	31	52

[2] M. B. Jensen, M. P. Philipsen, A. Møgelmose, T. B. Moeslund, and M. M. Trivedi, "Vision for looking at traffic lights: Issues, survey, and perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, pp. 1800–1815, 2016.

[3] M. Weber, P. Wolf, and J. M. Zöllner, "Deeptlr: A single deep convolutional network for detection and classification of traffic lights," *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 342–348, 2016.

[4] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *CoRR*, vol. abs/1804.02767, 2018.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, pp. 84–90, 2012.

[6] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

Real-Time Forward Collision Alert System using Raspberry Pi

Wai Chun Phoon and Phooi Yee Lau
Universiti Tunku Abdul Rahman, Centre for Computing and Intelligent System (CCIS), 31900 Kampar, Malaysia
phoon30@gmail.com, laupy@utar.edu.my

Abstract—*Most new luxury cars often comes with some form of driving assistance systems to reduce the risk of car accidents – such as the lane departure warning system, automatic parking, anti-lock braking system(ABS), blind spot monitor and forward collision warning system. These systems are developed to enhance the car safety and to provide a safer driving conditions. Due to its cost and design, these systems, often are only available in luxury cars. The main purpose of this work is to propose a real-time low-cost Forward Collision Alert System (FCAS). The FCAS are able to alert drivers, when their cars are getting too close to the front vehicle by estimating the speed of the car in front, using a loud beeping sound. The FCAS is then embedded onto a Raspberry Pi for real-time. Experimental results shows that the system can reliably alert the driver, in real-time, when a moving vehicle in front position is too close.*

I. INTRODUCTION

These days almost everyone owns a vehicle, resulting in the increase of vehicle on the road – attracting more accidents yearly [1]. Often, accidents happen due to: 1) complex and uncertain road environment, and 2) driving behavior. Many systems are being developed to provide assistance while driving, and these systems should enable driver to have a safer inter-car space, which comes with an alert system, but unfortunately, most of these system are often only available in top luxuries cars, due to its development cost and installation.

In 2016, a set of lane detection techniques to detect lanes, and distance estimation algorithm to estimate the distances, for a moving car was proposed by Ileri, K, et al [1]. In 2016, Alpar et al proposed a technique named Corona Segmentation [2]. This technique could discriminate vehicle's brake light from other light source to prevent forward collision. Prior, in 2015, Garethiya et al proposed a predictive vehicle collision avoidance system using Raspberry Pi and ultra-sonic sensors [3]. In the same year (2015), Ujjainiya et al conducted experiments to compare the performance of Laplacian edge detection, Sobel edge detection and Canny edge detection in in vehicle detection. The experimental results indicate that the Canny edge detection technique performs better [4]. In 2017, Huang et al proposed a simple method that is which could efficiently detects vehicle and estimates the inter-vehicle distance using single lens camera [5]. However, all these systems focused on detection and lacks of on-the-road real-time performance evaluation. In this paper, a real-time low-cost Forward Collision Alert System (FCAS) is proposed. The real-time driver alert system consists of (1) a FCAS module, and, (2) Raspberry Pi Deployment (RPD) module.

The rest of the paper is organized as follows: Section II which describes and discusses proposed approach, and Section III which show the implementation results with discussions, future works, and conclusions.

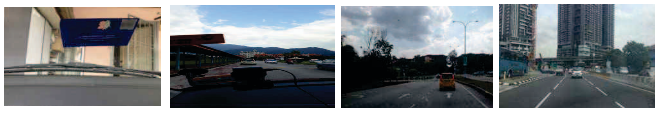
II. METHODOLOGY

The FCAS proposed to monitor vehicles in front under different environment, and to provide sound alert to warn drivers when vehicle in front is getting too close. Fig. 1 show how the Forward Collision Alert System (FCAS) works. The FCAS consists of 5 major steps, being: (1) *Step 1: FCAS-Video acquisition*, (2) *Step 2: FCAS-Speed acquisition*, (3) *Step 3: FCAS-ROI segmentation*, (4) *Step 4: FCAS-Vehicle detection*, and (5) *Step 5: RPD-Alert System*.

Fig. 1. Simple illustration of how FCAS work.

A. *Step1:FCAS-Video Acquisition*

As shown in Fig. 2, the acquired video stream in full HD is resized to 640x480 pixels, 30 frames per second and H264 encoding format. This is to ensure the system work on different scenarios, light scenario – during daylight, and lowlight scenario - during night.



(a) (b) (c) (d)

Fig. 2. Video acquisition.

The daylight scenario will be set between 07:30 (UTC +8) to 19:30(UTC +8) and the lowlight scenario will be set between 19:31(UTC +8) to 07:29(UTC +8). In this step, it is important to (a) avoid capturing any unwanted information such as road tax sticker and wiper – see Fig 2(b), and (2) ensure that when mounting the camera during installation, the camera should be facing its own lane, in a straight line, as shown in Figure 2(d) and not diagonally, as shown in Fig. 2(c).

B. *Step2: FCAS-Speed Acquisition*

Obtaining the vehicle speed is important, as the alert system should not be triggered during low speed or stopping condition, i.e. due to red traffic light signs or when a car is

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

stuck in a traffic jam, i.e. the system will only operate between 40-100km/h. In this paper, an Android smartphone built-in GPS chip is being used to detect the car speed, via UDP through a USB cable, every 3 seconds.

C. Step3: FCAS-ROI Segmentation

Here, the main Region of Interest (ROI) is being identified – as shown in Fig. 3 (a). The ROI setting, i.e. at the bottom is to ensure the vehicle, often located at the lower part, can be detected – as usually the upper part represents the sky. This is also because it is impossible to find vehicle at the upper part of the image.

D. Step4: FCAS-Vehicle Detection

The system considered two types of scenarios, i.e. daylight scenario and lowlight scenario. In daylight scenario, the shadow for the vehicle in front is considered, by detecting the horizontal lines (assuming being the shadow): input image → grayscale image → threshold → line detection using Canny Edge detection → removal of vertical lines → removal of white noise → enhance the vehicle shadow. Depending on the ratio, some of the horizontal lines will be deleted, as the vehicle's shadow are too short.

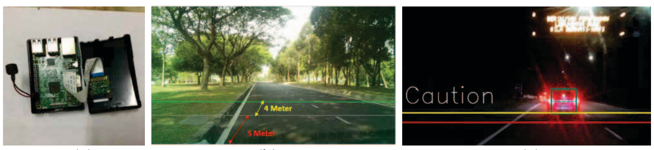
For the lowlight scenario, the tails light for the vehicle in front is considered, by detecting a horizontally same size spot: input image → grayscale image → threshold → multiple white spot detected (representing light source) → filter the minimum/maximum white spot size → calculate the white spot between the defined sizes (to filter unwanted noise) → stored as contours. As a vehicle's tail light is mostly the same size and only found horizontally, three conditions is proposed to detect a tail light; 1) contours are parallel to each other; 2) the contours have similar size; and 3) the distance between the two contours. Fig. 3(b) and 3(c) shows the system performance in detecting vehicle during daylight and lowlight.

(a) (b) (c)

Fig. 3 (a) ROI cropped-region. (b) – (c) Output of vehicle detection.

E. Step5: RPD-Alert System

This step focus on (1) embedding the FCAS system onto the Raspberry Pi unit, with a Raspberry Pi camera and a buzzer, to capture the driver's condition, and (2) mounting the system in the car, to alert the driver by providing a medium loud beeping sound – see Fig. 4(a). Here, we considered that an image could only contain one vanishing point, i.e. usually on the horizontal line.



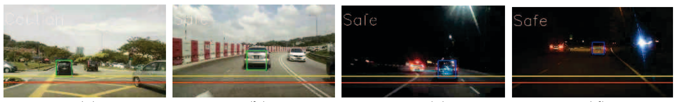
(a) (b) (c)

Fig. 4. (a) Recommended mounting style for Raspberry Pi. (b) – (c) Output: Alert.

Therefore, the alert systems can be created using two horizontal lines, being, (1) first alert (9 meters away), using a low frequency beeping to alert the driver, and (2) second alert (5 meters away), using a high frequency beeping sound to alert the driver – see in Fig. 4(b) and 4(c).

III. Discussions and Conclusions

In this section, a few sets of experiment with different conditions were conducted: (1) daylight scenario, and (2) lowlight scenario with few vehicle in front. As shown in Fig. 5(a) and 5(b), the system performs well in the following condition (a) road in housing area, (b) multiple vehicle in front, (c) Multi-lane road, and (e) drain cover. As shown in Fig. 5 (c) and 5(d), the system performs well in the lowlight condition.



(a) (b) (c) (d)

Fig. 5. (a) – (b) Output of the system in daylight. (c) – (d) Output of the system in lowlight.

Further, we evaluate the effectiveness of the proposed FCAS system captured using the Raspberry Pi 3 Model B with a resolution of 480x360 and 30fps. The experimental results shown in TABLEI indicate that the proposed FCAS framework could reliably alert the driver, in emergency condition, with an average accuracy rate of 86.8%.

TABLE I. ACCURACY OF SYSTEM DETECTION

Video	*True Positive*	*True Negative*	*False Positive*	*False Negative*	*Accuracy*
Video 1	795	21	66	12	91.3%
Video 2	733	33	122	5	84.3%
Video 3	736	20	130	4	84.9%

In contrast, some limitations is being highlighted during deployment, such as the lightning conditions and the environment noise would greatly influence the system accuracy. For future works, we proposed to integrate the lane detection system from Chee et al. to improve the system [6]. By employing the lane detection, the ROI would be fixed according to the lane instead of focusing on the vehicle in front.

References

[1] Ileri, K., Duru, A and Gorgunoglu, S.. Single Board Computer Based real Time Forward Collision Warning System For Driver Safety, The IIER International Conference,(73), pp.19-22.

[2] Alpar, O. Corona. Segmentation for Nighttime Brake Light Detection. IET Intelligent Transport Systems, 10(2), pp.97-105.

[3] Garethiva, S., Ujjainiya, L. and Dudhwadkar, V.. Predictive Vehicle Collision Avoidance System using Raspberry Pi. Asian Research Publishing Network, 10(8), pp.3655-3659.

[4] UJJainiya, L. and Chakravarthi, M.. Raspberry Pi Based Cost Effective Vehicle Collision Avoidance System Using Image Processing. Asian Research Publishing Network, 10(7), pp.3001-3005.

[5] Huang, D., Chen, C., Chen, T., Hu, W. and Feng, K.. Vehicle Detection and Inter-vehicle Distance Estimation Using Single-lens Video Camera on Urban/Suburban roads. Journal of Visual Communication and Image Representation, 46, pp.250-259.

[6] Chee, W., Lau, PY. and Park, S.. Real-time Lane Keeping Assistant System on Raspberry Pi. IEIE Transactions on Smart Processing and Computing, 6(6), pp.379-386.

Bounding Box based Annotation Generation for Semantic Segmentation by Boundary Detection

Xiaolong Xu, Fanman Meng, Hongliang Li, Qingbo Wu, Yuwei Yang and Shuai Chen
School of ICE, University of Electronic Science and Technology of China, Chengdu, China

***Abstract*—This paper proposes a new method to generate pseudo-annotations from manual bounding boxes for semantic segmentation. Different from traditional local data driven based methods such as Conditional Random Field (CRF) and GrabCut, we aim at using class-agnostic bounding box based segmentation models. To this end, we propose a new segmentation network, which formulates segmentation task as a sparse boundary point detection task rather than dense pixel label prediction task, and therefore can provide new type of pseudo-annotations. Furthermore, we detect object boundary based on direction, and use multiple directions to handle various shapes of objects. Moreover, we further enhance the pseudo generation by combining different types of segmentation masks. Classical Fully Convolutional Networks (FCN) network based on dense prediction is also modified to generate diverse foreground masks. A simple fusion method based on intersection operation is proposed to combine the two types of pseudo-annotations. We verify the effectiveness of our method on PASCAL VOC 2012 validation dataset. The mIoU value is 67.9%, which outperforms the state-of-the-art method by 1.1%.**

***Index Terms*—Weakly Supervised Semantic Segmentation, Pseudo-annotation, Bounding Box, Boundary Prediction**

I. Introduction

Semantic segmentation requires a large number of pixel-level annotations [1]. However, manually generating the annotations is an extremely time-consuming task. To this end, researchers replace pixel-level manual annotations with simple annotations such as windows, scribbles, and image-level labels [2], [3]. An efficient method is to generate pseudo-annotations from manual bounding boxes, and then uses the pseudo-annotations to train the segmentation model [4].

Good pseudo-annotation generation for bounding box based method relies on accurate transformation from window region to object region. By observing the fact that single segmentation method is hard to ensure the robust transformation, researchers propose fusion method that combines different types of pseudo-annotations generated by diverse methods to form more accurate pseudo-annotations. Experiments show that fusion method improves the quality of pseudo-annotation well [4]. However, the performance of fusion method depends on the diversity of initial pseudo-annotations.

Based on the fusion strategy, we try to build class-agnostic segmentation models to generate different types of pseudo-annotations. Two pseudo-annotation generation networks are proposed. One is boundary point detection network which formulates segmentation task as sparse boundary point detection task. Such method can not only provide new type of pseudo-annotation, but also reduce the prediction cost by three times. Meanwhile, the boundary point prediction is implemented in multiple directions, and their combination can ensure the robust segmentation of objects with diverse shapes. The other is classical FCN model which is dense prediction of pixel labels, and achieves segmentation by learning the mapping between windows and annotations. Since the two kinds of pseudo-annotation models formulate segmentation problem as sparse boundary point detection problem and dense pixel label prediction problem respectively, with different segmentation processes and segmentation results, their results are totally different, and their fusion can ensure high-quality generation of pseudo-annotations. Experiments demonstrate the effectiveness of the proposed method.

II. The Proposed Method

A. Overview

The proposed method consists of two steps: foreground mask generation and foreground mask combination. Two class-agnostic segmentation networks are firstly introduced. Then, the fusion methods for both multiple bounding boxes and multiple semantic segmentation annotations are proposed to form the final pseudo-annotations.

B. Foreground Mask Generation

1) Boundary Point Detection Network: The pipeline and details of the proposed segmentation network are shown in Fig. 1, where the idea is to formulate the segmentation problem as the boundary point detection problem. Our method predicts the boundary points based on a certain direction, such as the horizontal direction. Moreover, in order to achieve sparse prediction for low computation burden, we uniformly sample n lines in the direction and predict two edge points (such as left and right point for horizontal direction) on each line. Therefore, the segmentation problem is transformed into a prediction problem of $2 \times n$ points. By setting n to be a small value, it is possible to achieve sparse and fast prediction. By connecting the boundary points, the object region is obtained.

Meanwhile, single direction cannot predict inner boundary points of concave shape. To overcome such shortcoming, we employ multiple directions, such as horizontal and vertical directions. The advantage of using multiple predictions is that the edge of failure in one direction will be compensated by other directions. Better boundary can be generated by combining multiple directions. Specifically, we use the prediction method to generate boundary points on each direction. Then, we combine the results of all directions to produce the final

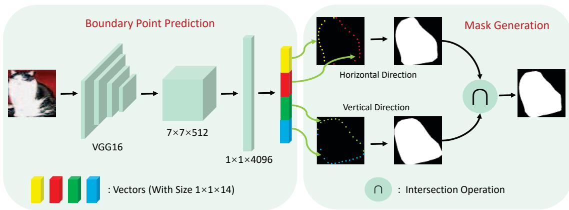


Fig. 1. The pipeline of our proposed boundary point detection network.

TABLE I
THE MIOU VALUES ON PASCAL VOC 2012 VALIDATION DATASET. OUR BACKBONE NETWORK IS SET TO VGG-16 FOR FAIR COMPARISON.

Methods	mIoU
BoxSup$_{Box}$(ICCV'2015) [6]	52.3
WSSL$_{CRF}$(ICCV'2015) [7]	60.6
GAIN (CVPR'2018) [8]	55.3
MCOF (CVPR'2018) [9]	56.2
AffinityNet(CVPR'2018) [10]	58.4
DSRG(CVPR'2018) [2]	59.0
MDC(CVPR'2018) [11]	60.4
FickleNet(CVPR'2019) [3]	61.2
SDI$_{M+G}$(CVPR'2017) [4]	65.7
BCM-FR$_{CRF}$(CVPR'2019) [12]	66.8
Ours	**67.9**

object boundary. We simply use the intersection of regions to combine the regions of multiple directions. We use mean square error loss function to train the model.

2) FCN Network: FCN Network adopts the classical FCN in [5], with simple modification that the input is the cropped regions, and the output is binary mask. The binary cross entropy loss function is used to train the model.

C. Forming semantic segmentation Annotation

For each network, we next form the semantic segmentation annotations based on the results of the bounding boxes. Our idea is to put the segmentation result into the image mask according to the position of the window. Note that some pixels may be segmented into different foregrounds (For example, when the "Cat" is on the "Sofa", the Cat's area will be assigned to the labels of "Sofa" and "Cat" simultaneously). For such case, we merge the results according to the size of the region and assign pixel with the label of small object. Experiments show that such method can ensure high-quality training.

D. Combining the Annotations

Given two pseudo-annotations, we merge them by keeping their same labels, and filtering out the different labels (ignored when calculating losses). We use the pseudo-annotations generated above to train and establish the semantic segmentation model.

III. Experimental Results

We verify our method on PASCAL VOC 2012 validation dataset. For fair verification, we use all the images in MS COCO 2017 except the 20 classes in PASCAL VOC 2012 to train the two class-agnostic segmentation networks. Each window region is resized to 224×224. We set $n = 14$. Horizontal and vertical directions are used empirically. The performance is measured by the mean intersection-over-union (mIoU) value.

We compare our proposed method with several semantic segmentation methods which are based on bounding boxes or other weak annotations. The mIoU values are shown in Table I. It is seen that the mIoU value of the proposed method is 67.9%, which is 1.1% larger than the mIoU value 66.8% of the state-of-the-art method such as BCM-FR$_{CRF}$.

IV. Conclusion

This paper proposes a method to generate pseudo-annotations from bounding boxes for weakly supervised semantic segmentation. Boundary Point Detection network and FCN network are firstly proposed to generate regions. Then simple fusion strategies are proposed to form pseudo-annotations. The experimental results on PASCAL VOC 2012 dataset demonstrate the effectiveness of the proposed method.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 61871087, Grant 61502084, Grant 61831005, and Grant 61601102, and supported in part by Sichuan Science and Technology Program under Grant 2018JY0141.

References

[1] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *PAMI*, vol. 40, no. 4, pp. 834–848, 2017.
[2] Z. Huang, X. Wang, J. Wang, W. Liu, and J. Wang, "Weakly-supervised semantic segmentation network with deep seeded region growing," in *CVPR*, 2018, pp. 7014–7023.
[3] J. Lee, E. Kim, S. Lee, J. Lee, and S. Yoon, "Ficklenet: Weakly and semi-supervised semantic image segmentation using stochastic inference," in *CVPR*, 2019, pp. 5267–5276.
[4] A. Khoreva, R. Benenson, J. Hosang, M. Hein, and B. Schiele, "Simple does it: Weakly supervised instance and semantic segmentation," in *CVPR*, 2017, pp. 876–885.
[5] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *PAMI*, vol. 39, no. 4, pp. 640–651, 2014.
[6] J. Dai, K. He, and J. Sun, "Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation," in *ICCV*, 2015, pp. 1635–1643.
[7] G. Papandreou, L.-C. Chen, K. Murphy, and A. Yuille, "Weakly-and semi-supervised learning of a dcnn for semantic image segmentation (2015)," *arXiv preprint arXiv:1502.02734*.
[8] K. Li, Z. Wu, K.-C. Peng, J. Ernst, and Y. Fu, "Tell me where to look: Guided attention inference network," in *CVPR*, 2018, pp. 9215–9223.
[9] X. Wang, S. You, X. Li, and H. Ma, "Weakly-supervised semantic segmentation by iteratively mining common object features," in *CVPR*, 2018, pp. 1354–1362.
[10] J. Ahn and S. Kwak, "Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation," in *CVPR*, 2018, pp. 4981–4990.
[11] Y. Wei, H. Xiao, H. Shi, Z. Jie, J. Feng, and T. S. Huang, "Revisiting dilated convolution: A simple approach for weakly-and semi-supervised semantic segmentation," in *CVPR*, 2018, pp. 7268–7277.
[12] C. Song, Y. Huang, W. Ouyang, and L. Wang, "Box-driven class-wise region masking and filling rate guided loss for weakly supervised semantic segmentation," in *CVPR*, 2019, pp. 3136–3145.

Hardware Architecture of Emotion Recognition from Speech Features using Recurrent Neural Network and Backpropagation Through Time

Joshua Gunawan, Teresia R. S. Putri, Yashael F. Arthanto
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
joshuagnwn@gmail.com

Trio Adiono
University Center of Excellence on Microelectronics
Institut Teknologi Bandung
Bandung, Indonesia
tadiono@gmail.com

***Abstract*— Emotion recognition from speech feature is one of the application where the system needs temporal information in order to produce a correct prediction. On the other hand, recurrent neural network has the advantage of retaining temporal information. This paper proposed a hardware architecture design for emotion recognition system using LSTM (Long Short Term Memory) and BPTT (Backpropagation Through Time). For this application, the proposed architecture consists of a two-layer stacked LSTM with 53 cells on the first layer and 8 cells on the second layer. The architecture is implemented and verified using Verilog language and FPGA.**

Keywords*— *emotion, recognition, recurrent neural network, speech, Verilog.

I. INTRODUCTION

There are a few architectures that could be implemented to achieve high-performance deep learning. One of this architecture is recurrent neural networks (RNN). This architecture is best when temporal context is needed, where a standard neural network model is lacking. Emotion recognition from the speech is a good application of a system that needed temporal information. One can only define the emotion from a person when we pay close attention through time, not at any instant sampling.

RNN works great with temporal information because every output is used as an additional input for the next sampling data. This concept requires a different technique of backpropagation model called backpropagation through time (BPTT) [6]. RNN is not without problem, it often faces the vanishing gradient problem where the model stops to learn before the minimum point [5]. This problem is solved with a few methods like LSTM and GRE. For this implementation, we use the LSTM model to classify speech features into eight different emotions: neutral, calm, happy, sad, angry, fearful, disgust, and surprise. A total of 53 features includes MPCC, LPCC, Chroma Vector, and other spectral analysis. In order to achieve real-time processing of LSTM, we implement this model in an FPGA using Zybo Zynq-7000.

II. EMOTION RECOGNITION SYSTEM

In order to achieve high accuracy in proposed speech emotion detection, we construct two layers LSTM cell with the first layer consisted of 53 LSTM cells works as an input layer and 8 LSTM cells in the second layer as an output layer.

Each cell has a weight parameter with the same length as the input data size. In the first layer, the input has a dimension of 53×1. In general, LSTM forward propagation has 4 gates defined as input squashing, input gate, forget gate and output gates [1], each consists of the activation function. Based on this basic LSTM structure, the total parameter needed to build the proposed network is 24,668. Thus, simple and effective architecture is needed in order to fit the model into the FPGA.

III. ARCHITECTURE

A. LSTM Network Structure

The original structure of 2 layers LSTM consisted of 53 cells in the input and 8 cells in the output as shown in Fig. 1. Because each LSTM cell has exactly the same structure, the network can be simplified into only two LSTM cell as shown in Fig. 2.

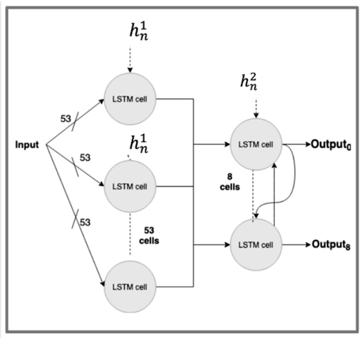


Fig. 1. Original LSTM network structure

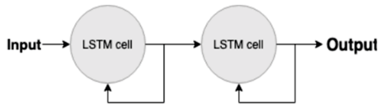


Fig. 2. Proposed LSTM network structure

B. LSTM Core

In order to reduce the number of utilized components, we simplify the complex computation with 1) Piecewise linear approximation to replace complex exponential function in activation function into addition and bit shifting function with maximum error resulted is 8%, 2) Multiply accumulator (MAC) to do dot products in LSTM cell, 3) Hardware sharing, by utilizing MAC inside the LSTM cells to do both forward propagation and backpropagation calculation. This hardware sharing architecture is later called an LSTM core as shown in Fig. 3.

C. Backpropagation Stage

The backpropagation calculation is divided into two parts: 1) The derivative of error in respect to gates is calculated using the delta gates calculation module as shown in Fig. 4., that works by simplifying complex equations into simpler ones and calculating them concurrently. This architecture only needs 12 cycles to produce the derivatives. 2) The derivative

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

of error in respect to cells' inputs, which requires matrix multiplications, are calculated using the LSTM core. To support this architecture, "DELTA X MEM" and "DELTA GATES MEM", and "BP MEM" are used to do complete backpropagation calculation as shown in Fig. 3.

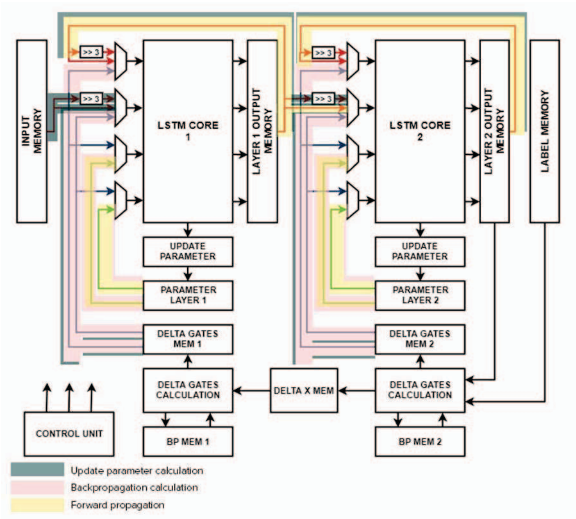


Fig. 3. The data path on Forward Propagation, Backpropagation, and Update Parameter Stage.

D. Update Parameter Stage

Utilizing the derivative of error in respect to gates from backpropagation stage, we can update the parameter of the system. The difference of parameters is calculated using the formula in [1]. After the calculation is done, the parameters are updated simultaneously with a learning rate of 0.125. The calculation itself utilizes the LSTM Core as shown in Fig. 3.

IV. Results

A. Software Modeling

Software modeling is done to find the best structure and hyper parameter that resulted in the highest prediction accuracy. Software modeling is done using Keras, and the hyper parameter and prediction accuracy are obtained as shown in TABLE I.

TABLE I. Summary of Software Implementation

Optimizer	Adam
Learning rate	0.001
Loss function	Mean Square Error
Epoch	150
Accuracy	92%

B. Hardware Implementation

There is a slight difference between hardware implementation and software model to reduce complexity. The proposed design has no optimizer and a learning rate of 0.125. From the timing analysis, the worst case path delay is 39.191 ns and therefore the maximum frequency is 25.5 MHz This frequency is slower than standard CPU but has a faster computation speed due to dedicated architecture and control unit. One cycle of the training process will require approximately 100,000 clock cycle.

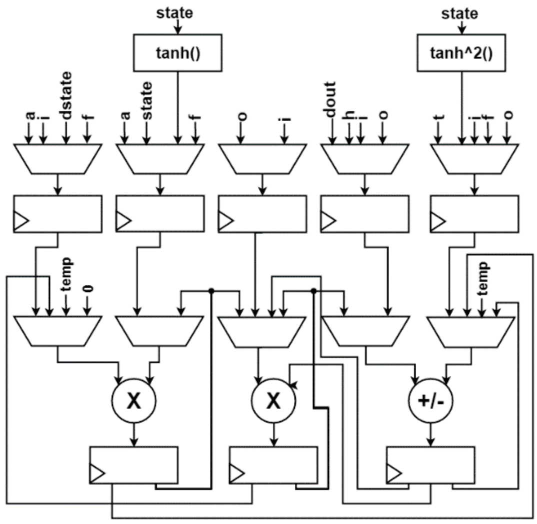


Fig. 4. Hardware sharing on delta gates calculation.

If the design is implemented with 25MHz then it will require 4 ms rather than 1 s performed by python using Intel Xeon. Our proposed design was implemented in the Zybo Zynq7000 as shown in Fig. 5. The hardware sharing concept makes it possible to implement complex computation in a hardware-based platform.

Resource	Estimation	Available	Utilization %
LUT	6709	17600	38.12
LUTRAM	192	6000	3.20
FF	3014	35200	8.56
BRAM	51	60	85.00
DSP	52	80	65.00
IO	28	100	28.00
BUFG	1	32	3.13

Fig. 5. FPGA utilization Report

V. Conclusion

This proposed design shows the architecture of LSTM for emotion recognition application. The hardware implementation works faster the standard processor (1.8 - 5.0 GHz) for this application using a very low system frequency (25 MHz). Faster processing is achieved due to the dedicated hardware and control unit design. Meanwhile, the design size is minimized by implementing hardware sharing architecture.

References

[1] Gang Chen, "A gentle tutorial of recurrent neural network with error backpropagation," *arXiv preprint arXiv:1610.02583*, 2016.

[2] S. R. Livingstone and F. A. Russo, "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English," *Plos One*, vol. 13, no. 5, 2018.

[3] M. Bhatti, Y. Wang, and L. Guan, "A neural network approach for human emotion recognition in speech," *2004 IEEE International Symposium on Circuits and Systems (IEEE Cat. No.04CH37512)*.

[4] P. Shen, Z. Changjun, and X. Chen, "Automatic Speech Emotion Recognition using Support Vector Machine," *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, 2011.

[5] R. Pascanu, T. Mikolov, and Y. Bengjo, "On the difficulty of training Recurrent Neural Networks," *Proceedings of the 30th International Conference on Machine Learning 2013 (ICML'13)*, pp. 1310– 1318, 2013.

[6] P. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

A Metropolis within Gibbs Algorithm for Estimating the Bayesian Reduced Reparameterized Unified Model

***Abstract* - The Reduced Reparameterized Unified Model (RRUM) is a model frequently seen in studies on educational data mining. The objective of this research is to advance an MCMC algorithm for the Bayesian RRUM. The algorithm starts with estimating correlated attributes. Using a saturated model and a binary decimal conversion, the algorithm transforms possible attribute patterns to a Multinomial distribution. Along with the likelihood of an attribute pattern, a Dirichlet distribution is used as the prior to sample from the posterior. Correlated attributes of examinees are generated using the inverse transform sampling. Model parameters are estimated using the Metropolis within Gibbs sampler sequentially. A simulation is conducted to evaluate the performance of the algorithm.**

Index Terms - RRUM, CDM, Bayesian, MCMC.

I. Introduction

Educational Data Mining is an emerging discipline that concerns with developing methods for better understanding students and the settings in which they learn. Cognitive diagnostic assessment (CDA) is a framework that aims to evaluate whether an examinee has mastered a particular cognitive process called attribute [1]. In CDA, exam items are each associated with attributes that are required for mastery. Using examinees' attribute states, CDA provides effective information for examinees to improve their learning and for educators to adjust their teaching.

A few cognitive diagnosis models (CDMs) have been developed, including the deterministic input, noisy-and gate (DINA) model [2], the noisy input, deterministic-and gate (NIDA) model [3], and the reparameterized unified model (RUM) [4]. All these models use the Q-matrix [5] to measure attribute states of examinees. The Q-matrix is represented as a J by K binary matrix, $Q_{J\times K} = (q_{jk})_{J\times K}$. If attribute k is required by item j, $q_{jk} = 1$. If attribute k is not required by item j, $q_{jk} = 0$. Suppose that there are a total of I examinees taking the exam that measures K attributes. To reveal the attribute state of an examinee, we use the the following binary matrix $A_{I\times K} = (\alpha_{ik})_{I\times K}$ to represent. If examinee i does not master the attribute k, then $\alpha_{ik} = 0$. If examinee i masters the attribute k, then $\alpha_{ik} = 1$.

Extending the NIDA model, [3] proposed a model that attempts to estimate the slip and guess parameters for different items. That is, the the slip and guess parameters have subscripts for both items and attributes. To improve this model, [6] advances the unified model that incorporates a unidimensional ability parameter. However, these two models are not statistically identifiable. [4] reparameterizes the unified model so that the parameters of the model can be identified while retaining their interpretability. As is expected, this reparameterized unified model is a more complicated conjunctive CDM. The RRUM has a more complex item response function (IRF) than other CDMs. The IRF of the RRUM is

$$P\left(X_{ij} = 1 \middle| \boldsymbol{\alpha}, \boldsymbol{\pi}^*, \mathbf{r}^*\right) = \pi_j^* \prod_{k=1}^{K} \left(r_{jk}^{*(1-\alpha_{ik})}\right)^{q_{jk}}$$

This research proposes a Metropolis within Gibbs algorithm for estimating the Bayesian RRUM. Specifically, a saturated model using the inverse transform sampling is used to estimate correlated attributes, and the Metropolis with Gibbs sampling is adopted to estimate the π^* and r^* parameters. The proposed algorithm is implemented in R [7].

II. Proposed Metropolis within Gibbs Sampling Algorithm

Step 1: Binary Decimal Conversion

The setting for the estimation is comprised of responses from I examinees to J items that measure K attributes. Given a J by K Q-matrix, the following steps perform sequentially at iteration t, t = 1, ..., T.

With K attributes, there are a total of 2^K possible attribute patterns for examinee i. Let $2^K = M$, and let the matrix **x** be the matrix of possible attribute patterns. Each of the M rows in **x** represents a possible attribute pattern, which is converted to a decimal number. After the conversion, these M possible attribute patterns become a Multinomial distribution. To estimate correlated attributes, a saturated Multinomial model is used that assumes no restrictions on the probabilities of the attribute patterns.

Step 2: Updating Probability of Attribute Pattern

Let **y** and **q** be the data and the Q-matrix. As the conjugate prior for a Multinomial distribution is also a Dirichlet distribution, Dirichlet(1, 1, ...,1) is used as the prior and therefore the conditional posterior is distributed as Dirichlet($1 + y_1$, $1 + y_2$, ...,$1 + y_M$), where y_l (l = 1, ... , M) is the number examinees possessing the l^{th} attribute pattern obtained from iteration t - 1.

Step 3: Updating Attribute

The full conditional posterior distribution is sampled using the discrete version of inverse transform sampling. Let the posterior (p_1, p_2, ..., p_M) be the PMF of the M possible attribute patterns. The CDF is computed by adding up the probabilities for the M points of the distribution.

Updating the attribute state of examinee i is achieved by checking which subinterval the value u falls into. This subinterval number is then converted to its corresponding binary number (see step 1) that represents the attribute state of examinee i. After step 3 is applied to each examinee, attribute states for all examinees.

*Step 4: Updating **r*** and **π*** Parameters*

A Metropolis within Gibbs algorithm is used to sample π* and r*. The non-informative Beta (1,1) prior is applied. Candidate values for **r*** is sampled from Uniform ($\mathbf{r}^{*(t-1)}$ - δ, $\mathbf{r}^{*(t-1)}$ + δ). It should be noted that candidate values for r* are restricted to the interval (0, 1), and that δ is adjusted so that the acceptance rate is between 25% and 40% (see Gilks et al., 1996). The updated **α** from step 3 is carried to step 4. As **π*** and **r*** are assumed to be independent of each other, p(**π***, **r***) = p(**π***)p(**r***).

TABLE I
Q-MATRIX FOR SIMULATION

Item	Attribute				
	1	2	3	4	5
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	0	0	1	0
5	0	0	0	0	1
6	1	0	0	0	0
7	0	1	0	0	0
8	0	0	1	0	0
9	0	0	0	1	0
10	0	0	0	0	1
11	1	1	0	0	0
12	1	0	1	0	0
13	1	0	0	1	0
14	1	0	0	0	1
15	0	1	1	0	0

E. *Simulation Design*

The Q-matrix for simulation is shown in Table 1. 30 items that measure 5 attributes comprise this artificial Q-matrix, which is constructed in a way that each attribute appears alone, in a pair, or in triple the same number of times as other attributes. An attribute is on average measured by 12 items. This Q-matrix is complete, containing at least one item devoted solely to each attribute.

III. RESULTS

The δ is set to 0.052 in step 4, so that the acceptance rate is around 35%. Table 2 presents the results from the simulation. For this complete and balanced Q-matrix, the recovery rates range from 0.919 to 0.941. It should be noted that using the independent model for simulation with sample size 2000 and correlation 0.5, we notice that the average of recovery rates from 20 data sets drops to 0.835, indicating that using the saturate model for correlated attributes is indeed improving the accuracy of attribute estimates.

TABLE 2
TRESULTS FOR THE SIMULATION

Size	Correlation		
	0.1	0.3	0.5
500	0.919	0.925	0.928
1000	0.922	0.929	0.936
2000	0.926	0.931	0.941

IV. DISCUSSION

The current research proposes an MCMC algorithm for estimating parameters of the RRUM in a Bayesian framework. The algorithm is summarized as follows. Using the binary decimal conversion, possible attribute patterns are transformed to a Multinomial distribution (step 1). Along with the likelihood of an attribute pattern, a Dirichlet distribution is used as the prior to sample from the posterior. The Dirichlet distribution is constructed using Gamma distributions (step 2), and attributes of examinees are updated using the inverse transform sampling (step 3). Sequentially, r* and π* are generated using the Metropolis within Gibbs sampler (step 4).

Like most of the studies, the simulation uses a complete and balanced Q-matrix. The measure of accuracy is on average 0.929. It is suggested that future research compare the estimated examinees' attribute patterns with the estimate from other CDMs such as the popular DINA model.

REFERENCES

[1] J.P., Leighton, and M.J. Gierl, (Eds.)., Cognitive diagnostic assessment for education.Theory and applications. Cambridge, MA: Cambridge University Press, 2007.

[2] B.W., Junker and K. Sijtsma, "Cognitive assessment models with few assumptions, and connections with nonparametric item response theory," Applied Psychological Measurement, 25, 258-272., 2001

[3] E. Maris, "Estimating multiple classification latent class models," Psychometrika, 64, 187–212, 1999

[4] S. Hartz, "A Bayesian framework for the Unified Model for assessing cognitive abilities: Blending theory with practicality (Doctoral dissertation)," University of Illinois, Urbana-Champaign., 2002

[5] K. K. Tatsuoka, "Rule space: An approach for dealing with misconceptions based on item response theory," Journal of Educational Measurement, 20, 345–354., 1983

[6] L. DiBello, L. A. Roussos, and W. Stout, "Review of cognitively diagnostic assessment and a summary of psychometric models," In C. V. Rao & S. Sinharay (Eds.), Handbook of statistics (Vol. 26, Psychometrics, pp. 979-1027). Amsterdam, the Netherlands: Elsevier., 2007

[7] R Development Core Team, R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.r-project.org., 2017

A Design Framework for Hardware Approximation of Deep Neural Networks

Wei-Hung Lin, Hsu-Yu Kao, Shih-Hsu Huang
Department of Electronic Engineering
Chung-Yuan Christian University
Taoyuan, Taiwan
{weihung,g10502605,shhuang}@cycu.edu.tw

***Abstract*— For real-time edge AI applications, there is a need to implement deep neural networks (DNNs) in hardware for high speed. The trade-off between computing accuracy and hardware cost must be made during the implementation of hardware approximation. In this paper, we propose a design framework for DNN hardware approximation. The proposed framework provides a behavior model library of approximate logic circuits (e.g., approximate multipliers) for the designers to utilize them. Moreover, the proposed framework also supports the dynamic fixed-point arithmetic for hardware simulation. To save the required total bit width, we develop an integer length tuning method in the framework to maximize the computing accuracy under a constraint on the total bit width. Experimental results on ICNet show that the proposed framework can achieve high computing accuracy with small hardware cost.**

Keywords—Neural Networks, Hardware Approximation, Hardware Simulation, Computing Accuracy, Bit Width, Design Methodology

I. Introduction

In recent years, a lot of deep neural network (DNN) models have been developed for various application domains, including semantic segmentation [1], image classification [2], speech recognition [3], and so on. For real-time AI edge applications, it is necessary to implement DNN models in hardware for high speed. To implement DNN models by hardware, there is a demand to make the trade-off between computing accuracy and hardware cost.

The proposed design framework is based on TensorFlow toolkit (an open source machine learning library) [4]. We have integrated the following three extra features into TensorFlow toolkit [4] for the simulation of hardware implementation.

(1) We provide a behavior model library that includes many approximate logic circuits (e.g., approximate multipliers) [5,6] for designers. Thus, designers can use behavior level descriptions to perform hardware simulation.

(2) The fixed-point representation is popularly adopted because of low cost. Thus, in the proposed design framework, we support dynamic fixed point arithmetic [7] for hardware simulation.

(3) To save hardware cost, we also develop an integer length (IL) tuning method to maximize the computing accuracy under a constraint on the total bit width.

We have used ICNet [1] to test the effectiveness of the proposed design framework. Experimental results show that the proposed design framework can achieve high computing accuracy with low hardware cost.

II. Proposed Design Framework

The proposed design framework is developed based on TensorFlow toolkit. Note that there is no hardware simulation in TensorFlow toolkit. However, TensorFlow toolkit supports C/C++ libraries that we can utilize them to create customized operations (i.e., the behavior model of hardware) for DNN hardware simulation (or hardware/software co-simulation). We can utilize these customized operations to simulate hardware behavior.

The proposed design framework provides three features for hardware simulation in TensorFlow. The details of these three features are addressed below.

A. Behavior-Model Library of Approximate Circuits

In order to save hardware cost and power consumption, approximate logic circuits are often used in DNN hardware implementation. Thus, to correctly analyze the behavior of a DNN hardware implementation, it is necessary to take all the approximations into account during simulation.

The proposed design framework provides a library that includes many approximate logic circuits for designers. For each approximate logic circuit, there are two models: behavior model and RTL model. The behavior level description (behavior model written in C++ language) is used for the simulation in TensorFlow. The RTL description (RTL model written in Verilog Language) is used for FPGA synthesis.

Fig. 1 illustrates a library of approximate circuits [5,6]. This library includes two multipliers and three activation functions (AFs). Designers can use behavior models for simulation and then use RTL models for FPGA synthesis.

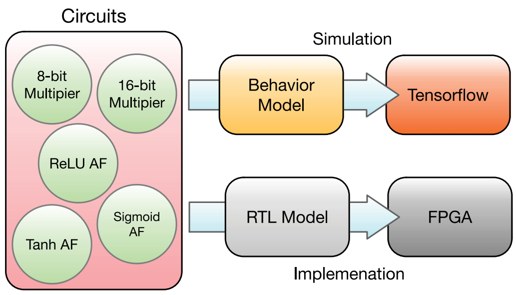


Fig. 1. A library of approximate logic circuits.

B. The Support of Dynamic Fixed Point Arithmetic

Previous work has shown that dynamic fixed point representation is useful for the quantization of DNN model. Especially, in hardware design, the dynamic mechanism of dynamin fixed point representation [7] can save total bit width. In the proposed design framework, we support the mechanism of dynamic fixed point arithmetic for hardware simulation.

To perform dynamic fixed point quantization with high accuracy, we need to correctly determine parameters of each convolution layer in advance, such as the integer length (IL) of input, output, and kernel. Here we use TensorFlow simulation to perform the analysis process. Then, from simulation results, we can determine parameters for quantization. Fig. 2 demonstrates the process flow for each convolution layer. For example, if all the input values during simulation are between -120 to 120, the IL of input can be determined to be 8 bits (including 1 sign bit). Suppose the total bit width of a real number is 16. If the IL of input is 8, then the number of bits for fraction will be 8 (i.e., 16-8 = 8).

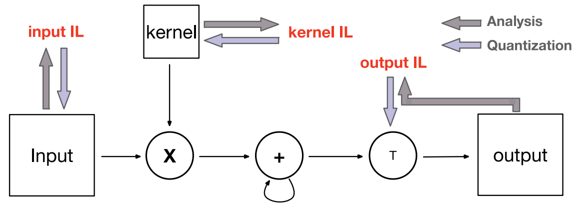


Fig. 2. The process of determining and using parameters.

C. Integer Length Tuningt

Table I gives an analysis to the minimum required IL during simulation. The row *Min IL* denotes the minimum required IL. The row *Amount* denotes the number of times that such a minimum required IL is required during simulation. In Table I, IL = 8 is required only 10 times during a long simulation. Thus, we may suppress these 10 extreme values (e.g., use the maximum or minimum value of IL = 7 to represent these 10 extreme values) with a small accuracy loss.

Table I. An analysis to the minimum IL during simulation.

Min IL	8	7	6	5
Amount	10	759907	8705622	13296450

Designers is asked to specify a ratio α as the lower bound (i.e., for the number of times appear during simulation). Let u and v denote the largest required IL and the smallest required IL, respectively, during simulation. Let M_k denote the number of times that the minimum required IL is k. Let T_k denote the summation of M_k, M_{k+1}, ……, and M_u. Note that T_k implies the number of times that the minimum required IL is larger than or equal to k. Since v is the smallest required IL, T_v also implies the number of total simulation. Then, the IL will be determined to be value n, where n satisfies both $T_n / T_v \geq \alpha$ and $T_{n+1} / T_v < \alpha$.

Suppose that α = 0.0001, the largest required IL is 8 and the smallest required IL is 5. From Table I, we have M_8 = 10, M_7 = 759907, and so on. Thus, T_8 = 10, T_7 = 759917, and so on. The number of total simulation (i.e., T_5) is 22761989. Since α = 0.0001, $T_7/T_5 \geq \alpha$, and $T_8 / T_5 < \alpha$, the IL is determined to be 7.

Note that the IL tuning method can be used to reduce the total bit width. Even under a constraint on the total bit width, the method can also be used to maintain the bit width for fraction.

III. Experiment Results

In this section, we demonstrate our results on ICNet [1]. The trainval model in ICNet is used for experiments.

First, we perform the simulation of dynamic fixed point arithmetic. Table II gives the mIoU (mean intersection-over-union) accuracy with respect to different total bit widths. The column Org denotes the original mIoU (e.g., 32-bit floating point representation). We find that, if the total bit width is reduced to 9 bits, the mIoU accuracy becomes 44.83%. Compared with the original mIoU accuracy (i.e., 80.90%), the accuracy loss is large.

Table II. The mIoU Results of different total bit widths.

	Org	Total Bit Width				
		16	12	11	10	9
mIoU (%)	80.90	80.66	80.56	79.80	75.29	44.83

Next, we try to use the IL tuning method to improve the mIoU accuracy. Fig. 3 gives an analysis. The blue line denotes the dynamic fixed point representation without IL tuning. The green line denotes the dynamic fixed point representation with IL tuning. As shown in Fig. 3, when the total bit width is 9, by using the IL tuning method, the mIoU accuracy can be improved from 44.83% to 72.98%. Thus, the IL tuning method can improve the accuracy.

Fig. 3. An analysis to the IL tuning method.

Fig. 4(a) gives the original result (i.e., 32-bit floating point representation). Fig. 4(b) gives the result of 16-bit fixed point representation without IL tuning. Fig. 4(c) gives the result of 9-bit fixed point representation without IL tuning. Fig. 4(d) gives the result of 9-bit fixed point representation with IL tuning. As shown in Fig. 4(d), the IL tuning method can achieve high accuracy with only 9-bit width fixed point representation.

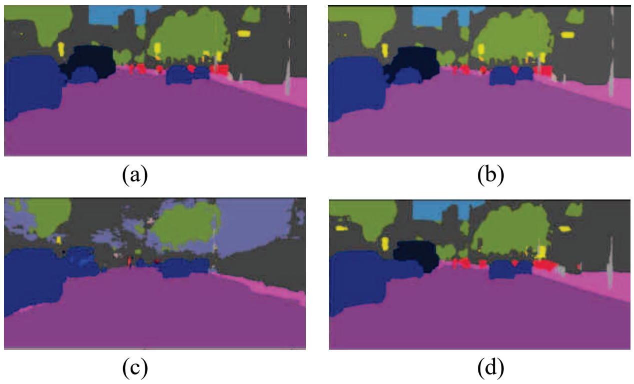


Fig. 4. Segmentation results of different quantization methods.

IV. Conclusions

This paper proposes a new design framework, which is based on TensorFlow, for the development of DNN hardware approximation. We have integrated three extra features into TensorFlow toolkit for the simulation of hardware implementation. Experiments on ICNet show that the proposed design framework is useful for the trade-off between computing accuracy and hardware cost. Especially, the IL tuning method can greatly improve the computing accuracy.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology of Taiwan, R.O.C., under grant number MOST 108-2218-E-033 -002.

References

[1] H. Zhao, et. al., “ICNet for Real-Time Semantic Segmentation on High-Resolution Image”, Proc. of IEEE ECCV, 2018.
[2] T. Guo, et al., “Simple Convolutional Neural Network on Image Classification”, Proc. of IEEE International Conference on Big Data Analysis, 2017.
[3] A.B. Nassif, et al., “Speech Recognition Using Deep Neural Networks: A Systematic Review”, IEEE Access, vol. 7, 2019.
[4] https://www.tensorflow.org/
[5] C.W. Tung and S.H. Huang., “Low-Power High-Accuracy Approximate Multiplier Using Approximate High-Order Compressors”, Proc. of IEEE ICCET, 2019.
[6] C.H. Chang, et al., "Hardware Implementation for Multiple Activation Functions", Proc. of IEEE ICCE-TW, 2019.
[7] P. Peng, et al., “Running 8-bit Dynamic Fixed-Point Convolutional Neural Network on Low-Cost ARM Platforms”, Proc. of the IEEE Chinese Automation Congress, 2017.

A Cooperation-based Scheme for Secondary Users to Enhance the Utilization of TV Spectra

David Shiung
Graduate Institute of Communications Engineering, National Changhua University of Education (NCUE), Changhua, 500, Taiwan
davids@cc.ncue.edu.tw

Ya-Yin Yang
Department of Electrical Engineering, National Cheng Kung University (NCKU), Tainan, 701, Taiwan
ivyyang64@gmail.com

Chi Ou-Yang
Graduate Institute of Communications Engineering, National Changhua University of Education (NCUE), Changhua, 500, Taiwan
m0557007@mail.ncue.edu.tw

***Abstract*—Space shadowed by obstacles, e.g., buildings and indoor structures, is traditionally considered as occupied and is unable to provide any useful incumbent services due to poor incumbent signaling. The time-frequency slots in this case are ignored for a long time and can be further used if an agile coexisting scheme is applied by secondary users (SUs). In this paper, we propose a new scheme for SUs to reuse these spatial opportunities in an incumbent digital TV (DTV) environment. These spatial opportunities can be used either by the SUs for private applications or by TV viewers (TVVs) provided that TV programs are suitably relayed. We analyze the applicability of the proposed scheme, which paves a way for enhancing the spectral efficiency in the spatial domain.**

***Index Terms*—TV white spectrum, spectral efficiency, coexisting scheme.**

I. Introduction

Tremendous development of wireless communication technology already turned the available spectrum into a scare resource [1]. According to the study conducted by Federal Communications Commission (FCC), the spectrum utilization is quite inefficient and at some time and at some places is merely 10% [2]. This paradox empowers the development of cognitive technology, which allows cognitive radios (CRs), also known as secondary users (SUs), to access the licensed spectra when the primary users (PUs) do not use them [3]. This idea regards spectrum holes in the time and in the frequency domains and does improve spectrum utilization [1], [4]. Nevertheless, there exist spectrum holes that are related to “distance”. This phenomenon can be seen when the distance to the primary transmitter is far enough or there are regions where the radio signals of the primary transmitter are poorly covered due to obstacles. The objective of this paper is to purpose an agile scheme to further enhance the spectral efficiency in the spatial domain.

The incumbent service provided in the TV bands is digital TV (DTV) broadcasting and the PUs are TV viewers (TVVs). Although the idea of cognition is attractive and the standardization activity is intense, lack of reward for the TVVs does hamper the deployment of SUs in TV white space (TVWS) [4].

In this paper, we redefine TVWS as a triplet of three parameters, *i.e.*, time, frequency, and space. Thus, TVWS is available even if the TV tower (TVT) is broadcasting. SUs in our new scheme have three roles to play: (i) relay (amplify and forward) TV programs for TVVs; (ii) use TVWS for private applications; and (iii) idle. The permissibility to use TVWS for each SU joining the game is constrained by its credit and it can be earned by relaying TV programs. On the other hand, the credit is lost due to consuming TVWS.

We model the variation of credit of a SU as random walks. The statistics is general such that simple random walks [5] (*i.e.*, credit variation with steps +1, -1, or 0) becomes a special case of our results.

II. Proposed Scheme

In order to fully utilize the spare spectra in all the time, frequency, and space domains, we give a new definition of TVWS. Then the work flow for the SUs to enable the cooperation-based coexisting scheme is presented.

A. Define TVWS

Consider the meaning of TVWS at locations $\vec{s} = (s_x, s_y, s_z)$ of a three-dimensional (3-D) Cartesian space. Since TV broadcasts use TV channels of identical bandwidth, it is convenient to slice both the time and frequency domains into grids of fixed size for analysis. Let a resource block (RB) be located at the uth time slot and the vth TV channel, where u and v are non-negative integers. We build a triplet vector $\vec{z} = (\vec{s}, u, v)$ to identify the TVWS regarding geographic locations, time slots, and TV channels of interest. Let $\vec{z}$ be any event. The indicator random variable (RV) for TVWS with a constrained outage probability ϵ is defined as

$$\boldsymbol{I}_{ws}(\vec{z}) \triangleq \begin{cases} 1, & \Pr(P_{TV}(\vec{z}) \geq \eta) < \epsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The parameter η denotes the threshold for TV-signal strength. $P_{TV}(\vec{z})$ denotes the TV signal power at $\vec{z}$. The neighbor indicator RV of $\vec{z}$ with respect to $\vec{z}_0 = (\vec{s}_0, u_0, v_0)$ is defined as

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

$$\boldsymbol{I}_{nbr,\delta,\vec{z}_0}(\vec{z}) \triangleq \begin{cases} 1, & d(\vec{s},\vec{s}_0) \le \delta, \boldsymbol{I}_{ws}(\vec{z}) = 1, \text{and } (u,v) = (u_0,v_0), \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

$d(\vec{s},\vec{s}_0)$ is the Euclidean distance between $\vec{s}$ and $\vec{s}_0$, which is equal to the length of the line segment connecting them. The TVWS-covered cluster spanned by an initial vector $\vec{z}_0$ is defined as

$$\mathbf{C}_{ws,\delta,\vec{z}_0} \triangleq \cup_{i=0}^{\infty}\{\vec{z} | \boldsymbol{I}_{nbr,\delta,\vec{z}_i}(\vec{z}) = 1\}, \quad (3)$$

where the subsequent vector $\vec{z}_i$, $i = 1,2,\cdots$ are spanned by $\vec{z}_0$ and are numbered sequentially. Notice that the cluster $\mathbf{C}_{ws,\delta,\vec{z}_0}$ is for the u_0th time slot and the v_0th TV channel.

B. Work Flow for SUs

The work flow for the cooperation-based scheme is graphically shown in Fig. 1. The TVT is the licensed transmitter, which broadcasts TV programs according to a regular schedule. The TVWS database also contains the information regarding the spatially available spare spectra, as defined in (3), which is periodically estimated and updated by the arbiter.

Consider a SU playing the work flow for a continuum of n time slots and starting the game from an initial credit x. The SU gained from the arbiter at the nth time slot is X_n. Actually if X_n<0 the SU pays $-X_n$ to the arbiter for the expense of using TVWS. If X_n=0 the SU just enters idle mode. Let S_n, $n \ge 0$, denote the credit number after n time slots. Then we have S_0=x and

$$S_n = x + X_1 + \cdots + X_n, n \ge 1. \quad (4)$$

In order to construct the cooperation between the SU and the broadcasting network, the arbiter makes a bargain with the SU to relay TV broadcasts when his credit is not greater than a, while to use TVWS when his credit is not less than b. Then the number of time slots before he is forced to either gain or consume credit is a RV defined by

$$T = \min(n \ge 0 | S_n \le a \text{ or } S_n \ge b). \quad (5)$$

To guarantee $S_n \le a$ or $S_n \ge b$ for some n, we further assume that $a \le x \le b$, $a < b$ and $\Pr(X_k = 0) < 1$. Set the credit gained at the nth time slot X_n as

$$X_n = \begin{cases} d' = d + 1, & \text{with probability } p, \\ -c' = -(c+1), & \text{with probability } q, \\ 0, & \text{with probability } r, \end{cases} \quad (6)$$

where $p \ge 0$, $q \ge 0$, $r \ge 0$, and $p + q + r = 1$. Note both d' and c' are positive and are determined by an economic demand-and-supply rule.

The statistics of S_T and T are interesting from an engineering perspective. We can obtain the probability of hitting the upper bound and lower bound as

$$\frac{x-a}{b-a+d} \le \Pr(S_T \ge b) \le \frac{x-a+c}{b-a+c} \quad (7)$$

and

$$\frac{b-x}{b-a+c} \le \Pr(S_T \le a) \le \frac{b+d-x}{b-a+d}, \quad (8)$$

respectively.

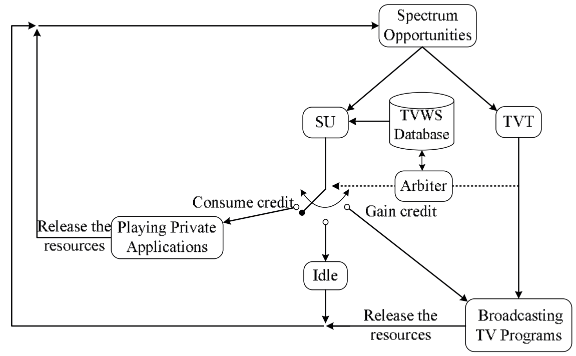


Fig. 1. Work flow for a SU gaining/consuming credit.

III. Numerical Examples and Discussions

Consider the examples with x=3, a=0, and b=10 for various settings of c and d shown in Table 1. Comparing case (i) with case (ii), we find that the probabilities of Pr(S_T≧b) and Pr(S_T≦a) increase for a large of c and d. This provides a preliminary evaluation of the proposed scheme.

Table 1. Numerical examples.

	(i) c=1 d=1	(ii) c=2 d=2	(iii) c=1 d=2	(iv) c=2 d=1
Pr(S_T≧b)	0.4545 ~0.5455	0.4167 ~0.5833	0.4167 ~0.5455	0.4545 ~0.5833
Pr(S_T≦a)	0.4545 ~0.5455	0.4167 ~0.5833	0.4545 ~0.5833	0.4167 ~0.5455

References

[1] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.

[2] FCC Spectrum Policy Task Force, *FCC Report of the Spectrum Efficiency Working Group*, Nov. 2002. [Online]. Available: http://www.fcc.gov/sptf/files/SEWGFinalReport1.pdf

[3] H. Maloku, E. Hamiti, Z. L. Fazliu, V. P. Lesta, A. Pitsillides, and M. Rajarajan, "A decentralized approach for self-coexistence among heterogeneous networks in TVWS," *IEEE Tr. Veh. Technol.*, vol. 67, no. 2, pp. 1302-1312, Feb. 2018.

[4] D. Shiung and Y.-Y. Yang, "On using TV white spectrum: A biological mutualism approach," *The 11th IEEE Vehicular Technology Society Asia Pacific Wireless Communications Symposium* (*VTS APWCS*), Taiwan, Aug. 2014.

[5] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to Probability Theory*, Houghton Mifflin Company, 1971.

Lattice-Superposition NOMA for Near-Far Users

Chi Wan Sung
Department of Electrical Engineering
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong
Email: albert.sung@cityu.edu.hk

Kenneth W. Shum
School of Science and Engineering
The Chinese University of Hong Kong (Shenzhen)
Shenzhen, China
Email: wkshum@cuhk.edu.cn

***Abstract*—A novel non-orthogonal multiple access scheme based on lattice superposition is proposed. Sixteen points from the D_4 lattice are picked as the signal constellation in the four-dimensional real space. Superposition coding for two users is applied at the downlink of the communication system. While the far user detects its intended signal directly, the near user performs interference cancellation before detects its own intended signal. The performance of our proposed scheme is compared with a benchmark orthogonal scheme using 16-QAM. When the link gain difference between the two users is large, our proposed lattice superposition scheme outperforms the benchmark significantly. In particular, when the link gain of the far user is 20 dB smaller than that of the near user, with a word error probability of 10^{-3}, an energy gain of 6 dB can be obtained.**

***Index Terms*—Non-orthogonal multiple access (NOMA), lattice, superposition coding, interference cancellation.**

I. Introduction

Non-Orthogonal Multiple Access (NOMA) is a promising candidate for 5G wireless communications because of its high spectral efficiency [1]. According to [2], NOMA can be broadly classified into power-domain NOMA and code-domain NOMA, depending on whether users are separated by using different power levels or using different codes. In this work, we consider power-domain NOMA. The principle of power-domain NOMA is based on superposition coding and successive interference cancellation, which has been known for many decades. From the viewpoint of information theory, it achieves the capacity of the Gaussian broadcast channel if Gaussian codebook with long block length is used [3]. In practice, it is inconvenient to use Gaussian codebook because of the lack of fast decoding algorithm. Besides, the use of long block length incurs long delay, which is undesirable considering the recent emphasis of short-delay applications. For this reason, we propose a simple practical scheme based on the use of lattices [4]. Different from the works [5], [6], which use lattice partition, our scheme is based on lattice superposition, which will be described in the next section. Afterwards, we will present our simulation results on its error performance. Our main contribution is the design and evaluation of a new NOMA scheme which can effectively reap the near-far gain induced by the difference of link quality between two users.

This work was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11216416).

II. Lattice Superposition

Consider the downlink of a two-user NOMA system. The base station transmits a signal vector $\boldsymbol{x} \in \mathbb{R}^L$, whose Euclidean norm is given by $\sqrt{E_T}$, where E_T is the total energy. The received vector of user i, where $i = 1, 2$, is given by

$$\boldsymbol{y}_i = h_i \boldsymbol{x} + \boldsymbol{n}_i,$$

where h_i is the amplitude gain and $\boldsymbol{n}_i$ is a random noise vector. We assume that the components of $\boldsymbol{n}_i$ are independent and identically distributed, each of which is a Gaussian random variable with mean zero and variance $N_0/2$, where N_0 is the noise spectral density. Without loss of generality, user 1 is assumed the near user while user 2 is the far user. The relative link gain in energy between the two users is $\beta \triangleq h_2^2/h_1^2 < 1$.

Superposition coding is used for transmission, i.e.,

$$\boldsymbol{x} = \sqrt{\alpha E_T}\boldsymbol{x}_1 + \sqrt{(1-\alpha)E_T}\boldsymbol{x}_2,$$

where $\boldsymbol{x}_i$ is the signal vector for user i, which is of unit norm, and $\alpha \leq 1$ is the energy splitting factor between the two users.

We consider the case where $L = 4$. The signal constellation for each user consists of 16 points, which are chosen from the D_4 lattice. Recall that the D_4 lattice consists of all integer vectors in $\mathbb{R}^4$ with an even element sum [4]. We pick 16 points that are closest to the origin and normalize them by the factor of $1/\sqrt{2}$. The normalized 16 vectors are listed as the rows of the following matrix:

$$\boldsymbol{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \end{bmatrix}, \text{ where } \boldsymbol{C}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

and $\boldsymbol{C}_2$ is a 12×4 integer matrix, whose rows are the 12 weight-two vectors with one component equal to 1 and the other equal to -1. We call $\boldsymbol{C}$ the signal constellation matrix. Note that each of its rows is of unit norm, and the distance between any two rows is at least one. To transmit four data bits to user i, the transmitter let x_i be the row of $\boldsymbol{C}$ whose index has binary representation equal to the four data bits. In other words, eight data bits are transmitted in four-symbol duration, so the energy per bit, E_b, is given by $E_T/8$.

We assume that the receiver of user i knows the value of h_i. User 2, the far user, treats the intended signal for user 1 as noise, and performs nearest-neighbor detection by finding

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

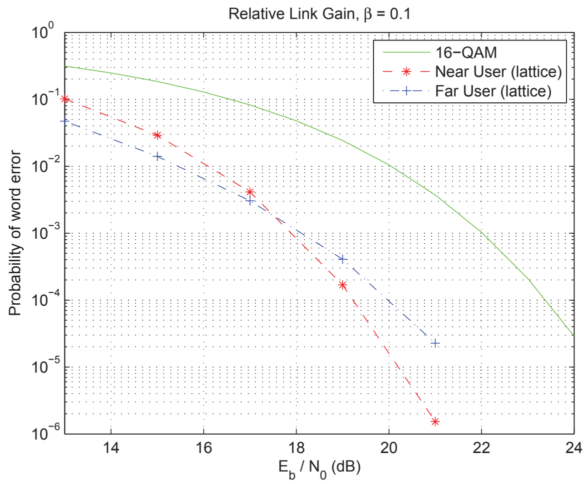


Fig. 1. Error Probabilities of lattice superposition and 16-QAM ($\beta = 0.1$).

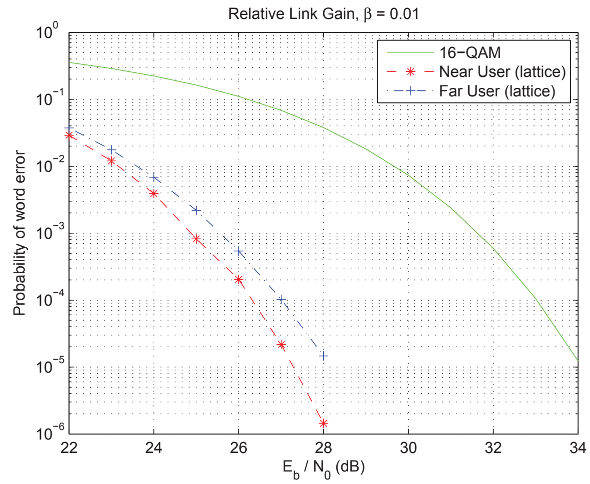


Fig. 2. Error Probabilities of lattice superposition and 16-QAM ($\beta = 0.01$).

the row vector in $h_2\sqrt{(1-\alpha)E_T}\boldsymbol{C}$ that is closest to $\boldsymbol{y}_2$. If the detected signal vector differs from the transmitted one, we say that a *word error* occurs at the far user, which means that one or more of the four data bits are in error. The probability of its occurrence is denoted by P_{w2}.

User 1, the near user, performs interference cancellation. It first detects the signal vector intended for user 2 by finding the row vector in $h_1\sqrt{(1-\alpha)E_T}\boldsymbol{C}$ that is closest to $\boldsymbol{y}_1$. Let it be $\hat{\boldsymbol{r}}$. It then subtracts $\hat{\boldsymbol{r}}$ from $\boldsymbol{y}_1$ to obtain

$$\boldsymbol{y}_1' = h_1\sqrt{\alpha E_T}\boldsymbol{x}_1 + \left(h_1\sqrt{(1-\alpha)E_T}\boldsymbol{x}_2 - \hat{\boldsymbol{r}}\right) + \boldsymbol{n}_1,$$

which is the received vector after interference cancellation. It then performs another nearest-neighbor detection by finding the row vector in $h_1\sqrt{\alpha E_T}\boldsymbol{C}$ that is closest to $\boldsymbol{y}_1'$. If the detected vector differs from the transmitted one, a *word error* occurs at the near user, which occurs with probability P_{w1}.

III. Simulation Results

In this section, we evaluate the performance of our proposed NOMA scheme with a benchmark orthogonal multiple access (OMA) scheme. Without loss of generality, the link gain for the near user is normalized to 1. We investigate how the probability of word error varies with E_b/N_0. For the OMA scheme, we assign two symbols to each of the two users. Each symbol is modulated by 4-PAM, so the two symbols together becomes 16-QAM, which represent four bits. Again a word error is defined as the event that one or more of the four bits are in error. The energy splitting factor, α, is chosen to be $\beta/(\beta+1)$, so that the two users under the OMA scheme have identical word error probability [7]:

$$P_w^{(OMA)} = 1 - \left[1 - \frac{3}{2}\,Q\left(\sqrt{\frac{\alpha E_T}{10N_0}}\right)\right]^2.$$

First, we consider the case where $\beta = 0.1$, which means that the link gain of the far user is 10 dB smaller than that of the near user. For the NOMA scheme, we let $\alpha = 0.05$. The error probabilities under the two schemes are plotted in Fig. 1, which shows lattice-superposition NOMA outperforms OMA. For example, to achieve a word error probability of 10^{-3}, lattice-superposition NOMA has 4 dB energy gain.

Next, we consider the case where $\beta = 0.01$, which means the link gain of the far user is 20 dB smaller. For the NOMA scheme, we let $\alpha = 0.01$. The error probabilities are plotted in Fig. 2. Since the difference between the link gains of the two users becomes larger, the advantage of NOMA increases. To achieve a word error probability of 10^{-3}, lattice-superposition NOMA has 6 dB energy gain for the far user and even more for the near user. This result agrees with the well-known fact that NOMA has larger benefit when the link gains between the two users are significantly different.

IV. Conclusion

A novel lattice-superposition NOMA scheme is constructed for the two-user Gaussian broadcast channel. It outperforms OMA especially when the link gain difference between the two users is large. We are extending the work to a larger system.

References

[1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for future radio access," in *Proc. IEEE VTC 2013-Spring*, Jun 2013, pp. 1–5.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Magazine*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[4] R. Zamir, *Lattice Coding for Signals and Networks*. Cambridge University Press, 2014.

[5] D. Fang, Y.-C. Huang, Z. Ding, G. Geraci, S.-L. Shieh, and H. Claussen, "Lattice partition multiple access: a new method of downlink non-orthogonal multiuser transmissions," in *Proc. IEEE GLOBECOM*, Dec 2016.

[6] M. Qiu, Y.-C. Huang, S.-L. Shieh, and J. Yuan, "A lattice-partition framework of downlink non-orthogonal multiple access without SIC," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2532–2546, Jun. 2018.

[7] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill Education, 2007.

Compare of Vehicle Management over the Air and On-Board Diagnostics

Taehyoung Kim
Department of Electronics and Computer Engineering
Hanyang University
Seoul, Republic of Korea
kth9922@nate.com

Sungkwon Park
Department of Electronics and Computer Engineering)
Hanyang University
Seoul, Republic of Korea
sp2996@hanyang.ac.kr

***Abstract*— In the future, the more advanced autonomous navigation and connected car technology, the more ECU and related software in the vehicle will increase. In the future, as the autonomous vehicle industry grows, information on software updates will increase. Autonomous vehicles are connected to the Internet and traffic-dedicated cloud servers through a connected vehicle system to achieve more efficient autonomous driving. Accordingly, update technology that continuously manages and improves the software will become important This paper focuses on vehicle management through Wireless Vehicle Software Update (OTA) and Wired Vehicle Software Update (OBD).**

Keywords—Vehicle Management, OTA, OBD-II, DTC, Software Update

I. Introduction

Recently, the value of automotives has been shifting from hardware to software in accordance with the progress of autonomous car development. As a result, vehicles are evolving around vehicle intelligence technologies based on ECU (Electronic Control Unit), and vehicles such as V2X (Vehicle-to-Everything) communication technology such as vehicle-infrastructure, connect connected technologies. Autonomous vehicles connect to the Internet and specific cloud servers in this connected-car mode to achieve more efficient autonomous driving. In the future, as the autonomous vehicle industry expands, more information will be needed to diagnose not only software update information. Previously, this information was updated through the workshop. However, it is very troublesome for the driver to visit the workshop to update periodic information. Therefore, even if driver do not visit the workshop, the update method is applied with a single click. As a result, a way to update the software without visiting the workshop has been developed. This paper compares the software management of vehicles through the existing wired network (using OBD-II) and the over-the-air (OTA) wireless network, which is currently applied to vehicles. In the second section of this paper, we introduces software updates using OTA technology in the vehicle and introduces the basics of OBD-II, the vehicle diagnostic standard. In third section, we compares the technologies of OBD-II and OTA and introduces each feature and future work.

II. Background

A. OTA(Over-the Air)

The OTA is mechanism for updating software over the wireless network without physically accessing the device. The software must support it to update the target device wirelessly. OTA can be divided into Software Over-The-Air (SOTA), which updates the software through a wireless network, and Firmware Over-The-Air (FOTA), which updates the microprogram that controls the hardware. The SOTA can be further divided according to the type of software. Most Infotainment and Telematics updates such as automotive infotainment apps and navigation maps are included. FOTA is a technology that can wirelessly update vehicle firmware and ECU software. Core ECU update using FOTA is currently being used by some vehicle Original Equipment Manufacturing(OEM) like Tesla and Hyundai. Figure 1 has divided OTA into SOTA and FOTA. In the era of fully autonomous driving in the future, most OEM will update vehicle performance or correct deficiencies via OTA.

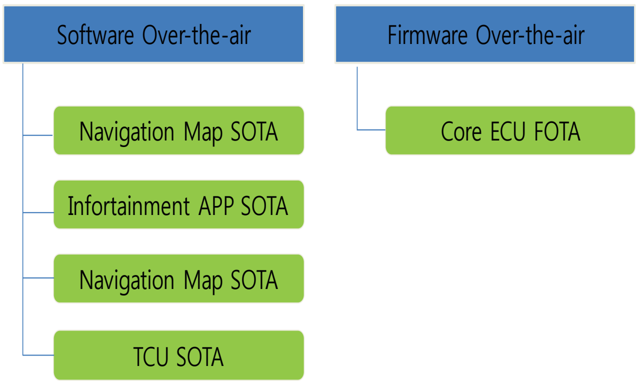


Figure 1. Example of SOTA and FOTA

The update scenario of vehicle software through OTA is as follows. [1][2]. In this scenario, new update information is generated from the OEM server and transmitted to the OTA Master of the user vehicle. OTA Master refers to the gateway embedded into the vehicle. In this case, the update information transmitted from the server is stored in the OTA Master. The stored update information will ask for approval of the vehicle ECU update via the Human Machine Interface (HMI) before the power is shut down (after the startup has finished). If the driver agrees, the OTA Master will send the update file to the corresponding ECU and the ECU will proceed with the software update. The point of this scenario is that the driver can select an ECU update and after the selection the vehicle is automatically shut down after the update is complete. Figure 2 shows OTA system architecture for ECU update.

We will send to the communication base station to transfer the update file from the OEM server. The base station that received the file transmitted from the server uses LTE to transmit to the vehicle that needs updating.

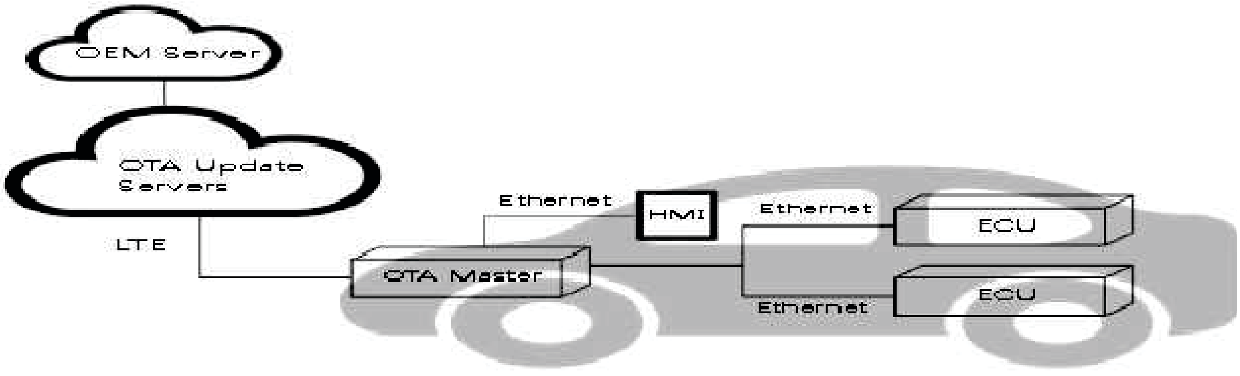


Figure 2. Architecture of OTA System

B. *OBD(On-Board Diagnostic)*

Currently, automotive diagnostic technology is based on the OBD standard[2]. The standard is a function that allows information on the main system of the car transmitted from the sensors attached to the vehicle to the ECU to be displayed on the vehicle's MLP (Malfunction Indicator Lamp) or an external device using the serial communication function. The OBD standard has evolved into a standardized diagnostic system called OBD-II[3]. Diagnosis by OBD-II can be informed of various alarms and diagnosis information of automobile by Diagnostic Trouble Code (DTC) generated in the car sensor and ECU. Automotive workshops use vehicle-specific diagnostic equipment called Vehicle Scanner to determine cause analysis and repair methods for alarms in various vehicle sensors and ECUs.

III. COMPARE OF VEHICLE MANAGEMENT METHOD

In order to diagnose and update the software information of the ECU and sensor of the vehicle, the driver visited the vehicle workshop to check for faults and update the new software via wired or USB. However, with the development of vehicle intelligence technology, the number of ECUs and sensors that require software in the vehicle is gradually increasing and becoming important. Therefore, if the OBD-II standard is applied at the workshop to manage all information, the driver will feel inconvenient. To address this inconvenience, OTA technology has been applied to vehicles. By applying OTA technology to the vehicle, the driver can update the software without restriction of position. In addition, automotive software updates as well as defects can be automatically reported to OEMs to prevent serious failures in advance. However, if the vehicle is updated with software using OTA technology, security risks may be involved[4]. The typical risks are as follows. In the event of an error in updating an important function such as an engine or brake airbag, the result is a serious consequence of human safety. To solve this problem, Tesla has developed a platform that is designed to be more open and secure for OTA update. In this respect, updating via OTA is a convenient way from the driver's point of view, but from a security perspective it is still more secure to update via the OBD standard.

IV. CONCLUSION

In this paper, we compared OBD-II technology for software update and vehicle diagnosis and OTA technology applied in recent vehicles. Previously, the driver periodically visited the workshop to update or diagnose the vehicle's software information. However, as the number of ECUs and sensors in the vehicle increases, so does the software for organically executing it. OTA technology has been applied to vehicles to periodically update and manage these software. OTA technology has unlimited potential to update various information such as the vehicle's ECU, as well as the Entertainment and Navigation. This paper introduces the features and advantages and disadvantages of software update through OTA technology newly applied to vehicles.

ACKNOWLEDGMENT

This work was supported by the Standard Technology Development and Spread Program of KATS/KEIT. [20001643 , Development of National Standards on Autonomous Vehicles and Use Case Guideline Provisioning]

REFERENCES

[1] Byungjoo Kim, Sungkwon Park, “ECU Software Updating Scenario Using OTA Technology through Mobile Communication Network”, 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS), pp. 67–72, 2018.

[2] Y. Onuma, Y. Terashima, S. Nakamura, R. Kiyohara, "A Method of ECU Software Updating", 2018 International Conference on Information Networking (ICOIN), pp. 298-303, 2018.

[3] Pooja Rajendra Sawant, Yashwant B Mane, “Design and Development of On-Board Diagnostic (OBD) Device for Cars”, 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 25 April 2019

[4] Alex Xandra Albert Sim, Benhard Sitohang, “OBD-II standard car engine diagnostic software development”, 2014 International Conference on Data and Software Engineering (ICODSE), 19 March 2015

[5] Hristos Giannopoulos, Alexander M. Wyglinski and Joseph Chapman, “Securing Vehicular Controller Area Networks: An Approach to Active Bus-Level Countermeasures”, IEEE Vehicular Technology Magazine, 2017, Vol 12, pp.60-68

Investigation on Distributed Joint Optimization of User Association and ICIC Based on PF Criteria in Small Cell Deployments

Takuya Matsumoto, and Nobuhiko Miki
Faculty of Engineering, Kagawa University

***Abstract*—To support the continuous growth of mobile traffic, small cell deployments are promising. In small cell deployments, combined usage of user association (UA) and inter-cell interference coordination (ICIC) is inevitable, which requires joint optimization of both techniques. Therefore, we propose joint optimization of UA and ICIC via distributed control. In our distributed control method, "distributed" control units (DCUs) were located at the macro base stations (BSs). Each DCU collected achievable data rates of users that were connecting to the macro BS and low-power (LP) BSs within the macro coverage. To obtain the high performance similar to centralized control, achievable data rates of users in the neighboring macro BSs were collected via the backhaul. Furthermore, we showed the effectiveness of our proposed scheme.**

I. Introduction

To support the continuous growth of mobile traffic, small cell deployments, that include a large number of low-power (LP) base stations (BSs) to heavy traffic area in addition to macro BSs, are promising. In small cell deployments, LP BSs are limited to a small coverage owing to the low transmission power. As a result, the radio resources of LP BSs are not effectively utilized, because the number of users connected to LP BSs is limited. Therefore, it is beneficial to offload users from macro BSs to LP BSs. However, offloaded users suffer from severe interference from macro BSs. Thus, to obtain an offloading gain, inter-cell interference coordination (ICIC) is essential. These two techniques, i.e., user association (UA) and ICIC, are strongly coupled and directly influence each other. Therefore, joint optimization of UA and ICIC is inevitable[1],[2][1]. Using various optimization criteria, in this study, we used a utility-based approach to employ "proportional fair (PF)" criteria, because it can achieve a good tradeoff between the cell throughput and cell-edge user throughput.

Joint optimizations described in many related papers have been performed via centralized control. In this type of configuration, the central control unit, which collects achievable data rates of all users in the network, is assumed. However, in an actual system, it is difficult to perform centralized control. Therefore, in this study, we employed a more realistic assumption, i.e., distributed control. Although there are many studies that have employed distributed controls, the joint optimization of UA and ICIC based on PF criteria has not yet been presented. For example, in [3], they performed joint optimization of UA and ICIC based on PF criteria via centralized control, and in [4], they performed UA optimization based on PF criteria via distributed control. Therefore, we propose joint optimization of UA and ICIC based on PF criteria via distributed control. In our distributed control method, "distributed" control units (DCUs) were located at the macro BSs. Each DCU collected achievable data rates of users that were connecting to the macro BS and LP BSs within the macro coverage. To obtain the high performance similar close to centralized control, achievable data rates of users in the neighboring macro BSs were collected via the backhaul (See details in the subsequent section). Furthermore, we performed a computer simulation to show the effectiveness of our proposed scheme.

[1]Reference [1] discusses the necessity of joint optimization, and includes related reference papers. Section III in [2] also includes a discussion of the necessity of joint optimization, and Table II provides a list of the reference papers.

II. Proposed Distributed Control

In this study, we assumed the cell deployments which the LP BSs were overlaid with macro BSs. The macro and LP BS groups were defined as $\mathcal{M}_m$ and $\mathcal{M}_p$, respectively. A 4-tier hexagonal cell layout that included 37 $(= 1 + 7 + 11 + 18)$ BSs and that employed an omni-sector with wrap-around was used for the macro BSs. $n_p = 6$ LP BSs were randomly installed within every one macro BS coverage. The n_p LP BSs were assumed to be connected to the macro BS by an optical fiber, which enabled joint optimization of UA and ICIC among each macro BS and n_p LP BS. This BS group was defined as $\mathcal{S}_m$, where m denotes the index of the macro BS. In the figure1, $\mathcal{S}_1 = \{$Macro BS1, LP BS1, LP BS2$\}$. Furthermore, macro BSs were assumed to be connected to the neighboring six macro BSs via the X2 interface. Although it is difficult to apply joint scheduling and UA between connecting macro BSs via the X2 interface owing to the large delay, it is possible to exchange information between macro BSs to utilize scheduling. We defined the macro BS group neighboring the m-th macro BS as $\mathcal{R}_m$. In the figure1, $\mathcal{R}_1 = \{$Macro BS2, LP BS3$\}$.

The system bandwidth was divided into $L = 2$ subbands. The macro BSs transmitted signals at the first subband, $\mathbf{B}_1$, with a transmission power spectrum density of P, whereas the macro BSs stopped transmission to reduce the inter-cell interference to the LP BSs in the second subband, $\mathbf{B}_2$. Furthermore, the LP BSs transmit signals at both subbands with a transmission power spectrum density of p.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

A. UA with Multiple BSs

Here, we will first describe the UA with multiple BSs. Each user first searches the BS with the maximum reference signal received signal power (RSRP) plus the common offset value for all LP BSs. Here the maximum RSRP plus the common offset value of the n-th user, $\gamma_n^{\max}$, is defined as $\gamma_n^{\max} = \max_{m\in\mathcal{M}}(\gamma'_{mn})$, where γ'_{mn} is defined as $\gamma'_{mn} = \gamma_{mn} + \delta \ (m \in \mathcal{M}_p)$, $\gamma'_{mn} = \gamma_{mn} \ (m \in \mathcal{M}_m)$, where γ_{mn} denotes the RSRP of the n-th user from the m-th BS. The δ value denotes the common offset value for all LP BSs.

In the next step, the n-th user determines the set of BSs for the n-th user, $\mathcal{T}_n$, which is described as $\mathcal{T}_n = \{m \in \mathcal{M} \mid \gamma'_{mn} \geq \gamma_n^{\max} - \phi\}$. Finally, the n-th user is associated with multiple BSs within $\mathcal{T}_n$.

B. Algorithm for the Proposed Distributed Control

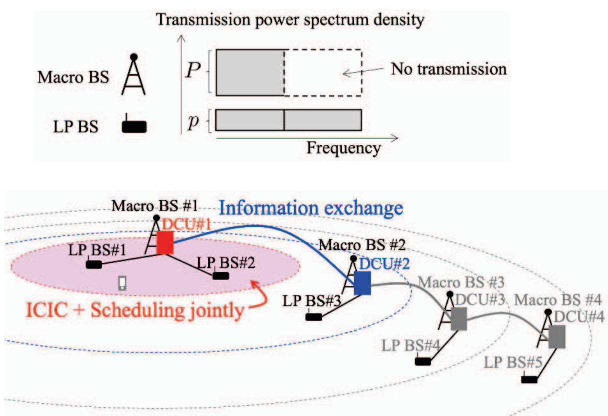


Fig. 1: Proposed distributed control.

The DCU located at the m-th macro BS (m-th DCU) performs joint user scheduling that belongs to $\mathcal{S}_m$. However, different macro BSs perform the scheduling independently. If the users are associated with different macro BSs, the independent scheduling causes a performance degradation. To avoid this degradation, the scheduling results are exchanged between neighboring macro BSs. More concretely, the m-th DCU schedules users associated with BSs that belong to $\mathcal{S}_m$ by utilizing the scheduling results of the m'-th DCUs $(m' \in \mathcal{R}_{m'})$.

III. Numerical Evaluation

In this section, we describe the evaluation of the performance of the proposed scheme based on a multi-cell simulation that employs macro and pico BSs (pico BSs were used as LP BSs in this study). Table I shows the simulation parameters. We assumed that 6 pico BSs and 24 users per one cell were randomly located within each cell with a uniform distribution. The δ value was set to be 21 dB.

Figure2 shows the cumulative distribution of the user throughput that employs the proposed distributed control scheme for the respective ϕ value. The cumulative distribution of the user throughput that employs the centralized control scheme is also plotted. This centralized control scheme is

TABLE I: Simulation Parameters.

Parameters	Value: Macro	Value: Pico
Cell layout	37 cell-sites	6 pico BSs per cell
Cell radius	289 m	
Distant-dependent path loss	$37.6 \log R$ +128.1 dB	$36.7 \log R$ +140.7 dB
Max. Tx power	46 dBm	30 dBm
Antenna gain	11 dBi	5 dBi
Carrier frequency	2 GHz	
Transmission bandwidth	10 MHz	

R: distance from the BS in km.

described in [5]. A proposed distributed control scheme with more than ϕ = 15dB can achieve comparable performance when the centralized control scheme is employed.

IV. Conclusion

This study proposed joint optimization of UA and ICIC via distributed control. We performed a computer simulation to show that the proposed distributed control scheme can achieve comparable performance when the centralized control scheme is employed.

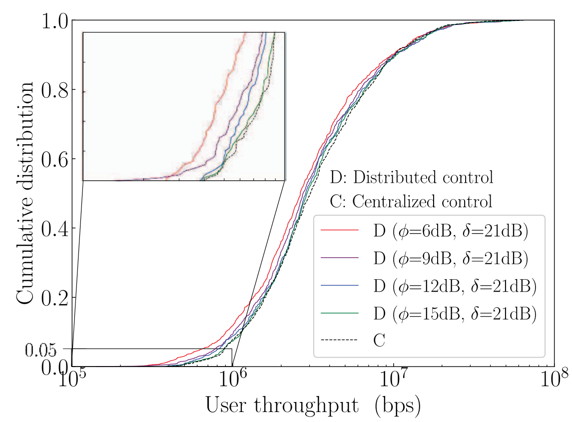


Fig. 2: Performance degradation of distributed control compared to centralized control.

References

[1] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," IEEE Transactions on Wireless Communications, vol. 13, no. 2, pp. 888–901, February 2014.

[2] D. Liu , L. Wang, Y. Chen, M. Elkashlan, K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," IEEE Communications Surveys & Tutorials, 2016.

[3] A. Bedekar and R. Agrawal, "Optimal muting and load balancing for eICIC," in 2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), May 2013, pp. 280–287.

[4] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," IEEE Transactions on Wireless Communications, vol. 12, no. 6, pp. 2706–2716, June 2013.

[5] N. Miki, Y. Kanehira, and H. Tokoshima, "Investigation on joint optimization for user association and Inter-cell interference coordination based on proportional fair criteria," in 2017 11th International Conference on Signal Processing and Communication Systems (ICSPCS'2017), Surfers Paradise, Australia, Dec. 2017.

Bended Differential Stripline Using Timing-Offset Differential Signal

Chih-Chen Yeh, Wen-Ju Chen, and Chun-Long Wang
Electronics Engineering
National Taiwan University of Science and Technology
Taipei, Taiwan
M9902263@mail.ntust.edu.tw; sabrina19960214@gmail.com; clw@mail.ntust.edu.tw

***Abstract*—In this paper, a bended differential microstrip line using the timing-offset differential signal is proposed. With the help of the timing-offset differential signal, the time-domain through common-mode noise could be greatly reduced from 0.068 V to 0.026 V. By replacing the microstrip line with the stripline, a bended differential stripline using the timing-offset differential signal is formed. With the help of the timing-offset differential signal, the time-domain through common-mode noise could be much better reduced from 0.051 V to 0.013 V, which is much smaller than 0.026 V.**

Keywords*—*bended differential stripline, timing-offset, differential signal

I. Introduction

Due to the increasing data rate of the digital signal and the decreasing size of the digital circuit, the electromagnetic interference between the traces on the printed circuit boards has become more and more serious than before. In order to alleviate the electromagnetic interference between traces, a differential trace has been widely applied in the circuit layout. [1]–[6].

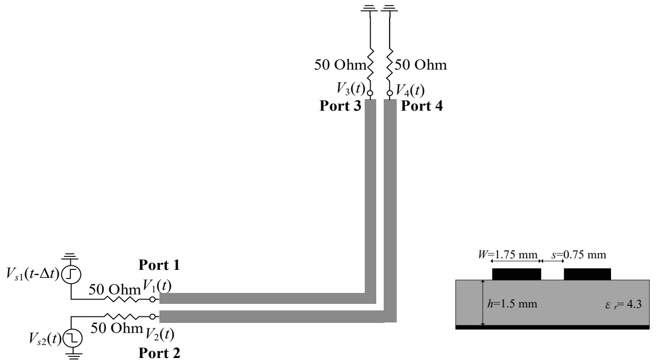
However, the differential trace may give rise to common-mode noise when the differential trace makes a turn in the printed circuit board. In order to eliminate the common-mode noise caused by the turn, various techniques have been proposed [7]–[15]. First of all, various types of the common-mode suppression filters are proposed to eliminate the common-mode noise [7]–[10]. Although the common-mode noise could efficiently be suppressed through using these common-mode suppression filters, large areas in the ground plane would be consumed and ground bounce problem would be induced.

In order to reduce the circuit size and prevent from the ground bounce problem, various compensation circuits are introduced at the turn of the differential trace [11]-[14]. However, by introducing the compensation circuit at the turn, the reflection of the differential trace would inevitably be increased. In order to reduce the reflection caused by the compensation circuit at the turn, a bended differential microstrip line using the timing-offset differential signal is proposed [15]. Since no compensation circuit is placed at the turn, the reflection of the differential trace would be reduced.

Although the bended differential microstrip line using the timing-offset differential signal could efficiently reduce the common-mode noise, the common-mode noise would be further reduced while the microstrip line is replaced with the stripline. As a result, a bended differential stripline using the timing-offset differential signal is proposed in this paper.

II. Bended Differential Microstrip Line Using Timing-Offset Differential Signal

The top view of the bended differential microstrip line is shown in Fig. 1(a) along with its cross-sectional view shown in Fig. 1(b). In order to investigate the common-mode noise of the bended differential microstrip line, a differential step function with an amplitude of 0.5 V and a rise time t_r= 40 ps is applied at the input ports 1 and 2 of the bended differential microstrip line. The common-mode noise is then acquired from the output ports 3 and 4 of the bended differential microstrip line. As can be seen from Fig. 1(a), the differential signal has an offset time Δt, which is used to compensate for the path difference between the inner and outer lines. The simulated and measured common-mode noises for the bended differential microstrip line with or without timing-offset differential signal are shown in Fig. 2. As can be seen from Fig. 2, the bended differential microstrip line with offset time Δt=0 ps will result in a large common-mode noise of 0.068 V while the bended differential microstrip line with offset time Δt=24 ps has a small common-mode noise of 0.026 V. By applying the timing-offset differential signal, the common-mode noise can be greatly reduced from 0.068 V to 0.026 V, estimating to be 38% of 0.068 V.



(a) (b)

Fig. 1. Bended differential microstrip line. (a) Top view. (b) Cross-sectional view.

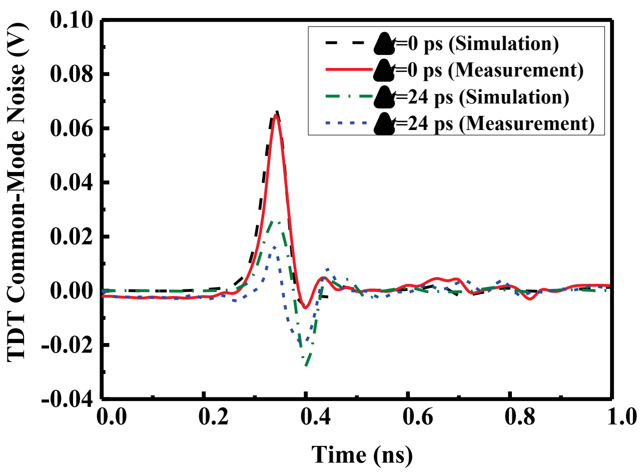


Fig. 2. Comparison between the simulated and measured results for the bended differential microstrip line with or without the timing-offset differential signal.

This work was supported by the Taiwan Ministry of Science and Technology Taiwan, under Grant MOST 108-2221-E-011-037.

III. Bended Differential Stripline Using Timing-Offset Differential Signal

In order to further reduce the common-mode noise of the bended differential microstrip line shown in Fig. 1, the microstrip line is replaced with the stripline. The top view of the bended differential stripline is shown in Fig. 3(a) along with its cross-sectional view shown in Fig. 3(b). The simulated and measured common-mode noises for the bended differential stripline with or without timing-offset differential signal are shown in Fig. 4. As can be seen from Fig. 4, the bended differential stripline with offset time Δt=0 ps will result in a common-mode noise of 0.051 V while the bended differential stripline with offset time Δt=40 ps has a common-mode noise of 0.013 V. By applying the timing-offset differential signal, the common-mode noise can be greatly reduced from 0.051 V to 0.013 V, estimating to be 25.4% of 0.051 V, which is much smaller than the common-mode noise caused by the bended differential microstrip line using the timing-offset differential signal.

IV. Conclusions

In this paper, to reduce the common-mode noise, a bended differential microstrip line using the timing-offset differential signal is proposed. No compensation circuit is needed at the turn to compensate for the path difference between the inner and outer lines of the bended differential microstrip line so that the reflection will be reduced. With the help of the timing-offset differential signal, the time-domain through common-mode noise could be greatly reduced from 0.068 V to 0.026 V, estimating to be 38% of 0.068 V. In order to further reduce the common-mode noise, the microstrip line is replaced with the stripline, forming a bended differential stripline using the timing-offset differential signal. With the help of the timing-offset differential signal, the time-domain through common-mode noise could be much more reduced from 0.051 V to 0.013 V, estimating to be 25.4% of 0.051 V, which is much smaller than the common-mode noise caused by the bended differential microstrip line using the timing-offset differential signal.

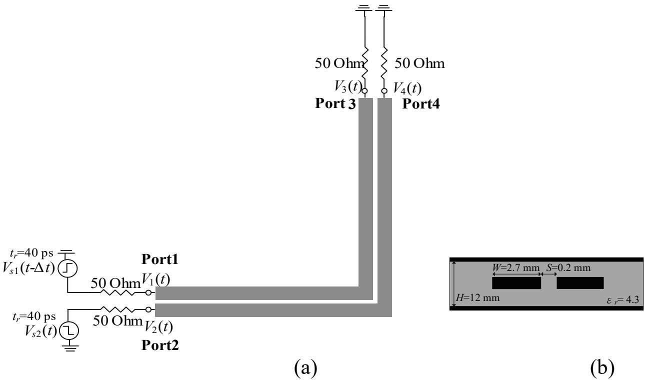


Fig. 3. Bended differential stripline. (a) Top view. (b) Cross-sectional view.

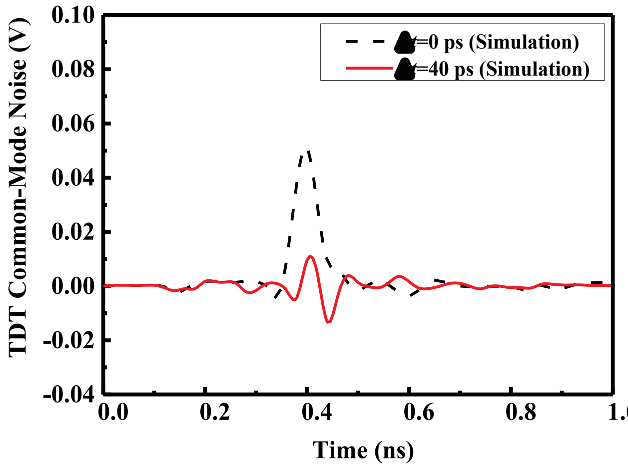


Fig. 4. Comparison between the simulated results for the bended differential stripline with or without the timing-offset differential signal.

References

[1] S. H. Hall, G. W. Hall, and J. A. McCall, *High-Speed Digital System Design.* New York: Wiley, 2000.

[2] P. E. Fornberg, M. Kanda, C. Lasek, M. Piket-May, and S. H. Stephen, "The impact of a nonideal return path on differential signal integrity," *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 11–15, Feb. 2002.

[3] G.-H. Shiue and R.-B. Wu, "Reduction in reflections and ground bounce for signal line through a split power by using differential coupled microstrip lines," in *Proc. Electrical Performance of Electronic Packaging*, Princeton, New Jersey, Oct. 2003, pp.107–110.

[4] Y. Massoud, J. Kawa, D. MacMillen, and J. White, "Modeling and analysis of differential signaling for minimizing inductive crosstalk," in *Proc. Design Automation Conf.*, 2001, pp. 804–809.

[5] E. P. Li, H. F. Jin, W. L. Yuan, and L. W. Li, "Parallelized computational technique for signal propagation analysis at very high speed differential transmission lines," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, 2003, pp. 850–854.

[6] W.-D. Guo, G.-H. Shiue, C.-M. Lin, and R.-B. Wu, "Comparisons between serpentine and flat spiral delay lines on transient reflection/transmission waveforms and eye diagrams," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 4, pp. 1379–1387, Apr. 2006.

[7] W.-T. Liu, C.-H. Tsai, T.-W. Han, and T.-L. Wu, "An embedded common-mode suppression filter for GHz differential signals using periodic defected ground plane," *IEEE Microw. Wireless Compon. Lett.*, vol. 18, no. 4, pp. 248–250, Apr. 2008.

[8] S. J. Wu, C. H, Tsai, T. L. Wu, and T. Itoh, "A novel wideband common-mode suppression filter for gigahertz differential signals using coupled patterned ground structure," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no.4, pp. 848–855, Apr. 2009.

[9] C.-H. Tsai and T.-L. Wu, "A broadband and miniaturized common-mode filter for gigahertz differential signals based on negative-permittivity metamaterials," *IEEE Trans. Microw. Theory Tech.*, vol. 58, no.1, pp. 195–202, Jan. 2010.

[10] M. A. Varner, F. de Paulis, A. Orlandi, S. Connor, M. Cracraft, B. Archambeault, M. H. Nisanci, D. Di Febo, "Removable EBG-based common-mode filter for high-speed signaling: experimental validation of prototype design," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 4, pp. 672–679, Aug. 2015.

[11] G.-H. Shiue, W.-D. Guo, C.-M. Lin, and R.-B. Wu, "Noise reduction using compensation capacitance for bend discontinuities of differential transmission lines," *IEEE Trans. Adv. Packag.*, vol. 29, no. 3, pp. 560–569, Aug. 2006.

[12] C.-H. Chang, R.-Y. Fang, and C.-L. Wang, "Bended differential transmission line using compensation inductance for common-mode noise suppression," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 2, no. 9, pp. 1518–1525, Sep. 2011.

[13] C. Gazda, D. V. Ginste, H. Rogier, R. B. Wu, and D. D. Zutter, "A wideband common-mode suppression filter for bend discontinuities in differential signaling using tightly coupled microstrips," *IEEE Trans. Adv. Packag.*, vol. 33, no. 4., pp. 969–978, Nov. 2010.

[14] B.-R. Huang, C.-H. Chang, R.-Y. Fang, and C.-L. Wang, "Common-mode noise reduction using asymmetric coupled line with SMD capacitor," *IEEE Trans. Components, Packaging, and Manufacturing Technology*, vol. 4, no. 6, pp. 1082–1089, Jun. 2014.

[15] C.-C. Yeh, B.-R. Huang, K.-C. Chen, R. Y. Fang, and C. L. Wang, "Reduction of common-mode and differential-mode noises using timing-offset differential signal," *IEEE Trans. Components, Packaging, and Manufacturing Technology*, vol. 5, no. 12, pp. 1818–1827, Dec. 2015.

A Compact Triple Passband Bandpass Filter Using Stub Load - Uniform impedance Resonators

1st Kuan-Jen Lin
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan
ren94521@gmail.com

2nd Zong-Fu Li
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan
confused0314@gmail.com

3rd Tzu-Chun Tai
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
yuhshooulove84@gmail.com

4th Hung-Wei Wu
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan
qqq25q@gmail.com

5th Yeong-Her Wang
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
yhw@ee.ncku.edu.tw

***Abstract*—In this paper, a compact triple passband bandpass filter (BPF) using stub load - uniform impedance resonators (SL-UIRs) is proposed. The resonators can provide the multipath propagation to enhance the filter performance and compact circuit size. The multipath SL-UIRs is designed to have two resonant paths at 2.4, 3.5 and 5.2GHz for WLAN and 5G. The resonant frequencies can be easily controlled by tuning the length of the uniform impedance resonators. The proposed triple passband filter is showing a simple configuration, an effective design method and small circuit size. The measured results are in good agreement with the full-wave electromagnetic simulation results.**

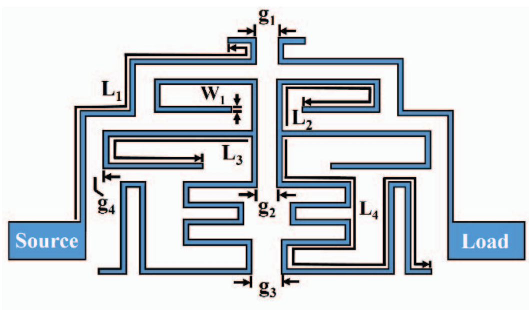
Keywords—Bandpass filter(BPF), multipath, stub-load uniform impedance resonators(SL-UIRs), triple passband.

I. Introduction

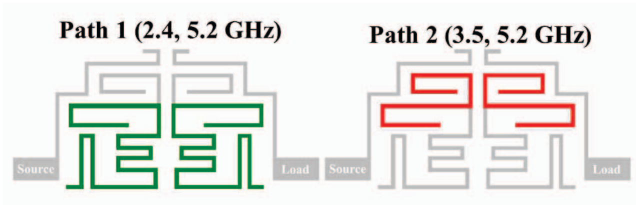
Recently, multi-passband bandpass filters become important component in front-end radio frequency application in the multi-band wireless communication systems. To design a multi-band filter with closed passbands, low insertion loss, compact size, and good passband selectivity is a challenge for the circuit designers.

Till now, some methods for triple passband bandpass filters (BPFs) were reported [1]-[2]. The tri-band bandpass filter is designed based on a stub-loaded step-impedance resonator (SL-SIR) [1]. The 0° feed structure is used to provide at least one transmission zero near passband edge of each passband, resulting in high selectivity. The dual-and triple-passband filter using alternately cascaded multiband resonators was proposed in [2]. The design concept is to add some extra coupled resonator sections in a single-circuit filter to increase the degrees of freedom in extracting coupling coefficients. However, the circuit size should be improved. This idea is good and inspired us to further study this issue.

In this study, we proposed compact triple passband bandpass filter by using stub load - uniform impedance resonators (SL-UIRs). The filter is designed to have two resonant paths (path 1 for 2.4 GHz and 5.2 GHz , path 2 for 3.5 GHz and 5.2 GHz). The path 1 and path 2 can be individually controlled for high design freedom. The filter is only using two coupling resonators to form three passbands with low insertion loss, high passband selectivity and compact size. The filter realized on a printed circuit with an area of only 0.23 × 0.11 = 190 mm^2. The measured results are in good agreement with the full-wave electromagnetic(EM) simulation results [3].



(a)



(b)

Fig. 1. **(a)** Configuration and **(b)** each resonant path of the proposed filter. (L_1 = 16.18, L_2 = 15.77, L_3 =16.47, L_4 = 37.875, W_1 = 0.4, g_1 = 0.2, g_2 = 0.15, g_3 = 0.5, g_4 = 0.2) All are in mm.

II. Design Of Triple Passband Filter

The filter is designed and fabricated on the substrate of Duroid 5880 with dielectric constant ε_r = 2.2, loss tangent δ = 0.0009 and thickness h = 0.787 mm. Fig. 1 (a) shows the configuration of proposed filter. The resonator includes two resonant paths at 2.4, 3.5 and 5.2 GHz, The filter is not only using two couple resonators to generate triple passband but also producing the transmission zeros at each passband skirt. The multipath stub loaded resonator includes two resonant paths as shows in Fig. 1 (b). Path 1 (indicated by green), path 2 (indicated by red) are designed at 2.4, 3.5, and 5.2 GHz. Specification of the filter was set at the center frequencies f_{01} = 2.4 GHz, f_{02} = 3.5 GHz and f_{03} = 5.2 GHz, with 3dB fractional bandwidth f_{01} = 4.22%, f_{02} = 3.98% and f_{03} = 3.43%. Fig. 2 shows coupling coefficients at 1st, 2nd and 3rd passband of the proposed filter. The lumped circuit element values of the low-pass prototype filter are found to be g_0 = 1, g_1 = 0.94982. g_2 = 1.35473, J_1 = -0.12333, J_2 =

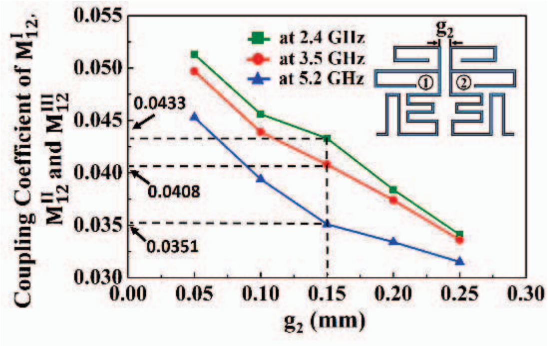


Fig. 2. Coupling coefficients at 1st, 2nd and 3rd passband of the proposed filter.

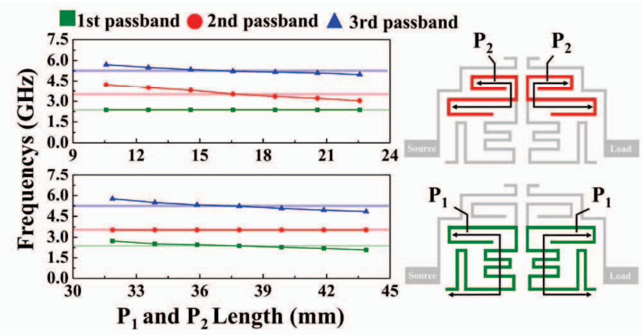


Fig. 3. Simulated frequency shifts ability of different length (P_1 and P_2) for each resonant paths.

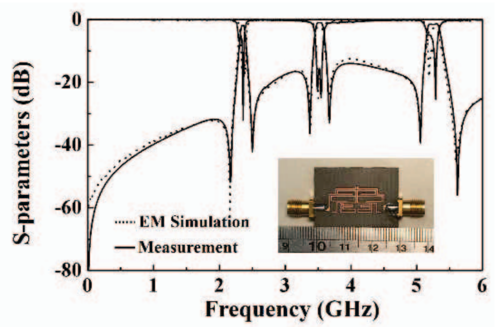


Fig. 4. Measured results and fabricated circuit of the proposed filter.

1.0181. The required coupling coefficients were found to be $M^{I}_{12} = M^{I}_{21} = 0.0433$ at 2.4 GHz, $M^{II}_{12} = M^{II}_{21} = 0.0408$ at 3.5 GHz and $M^{III}_{12} = M^{III}_{21} = 0.0351$ at 5.2 GHz [4]. Fig. 3 shows simulated frequency shift ability of different lengths (P_1 and P_2) for each resonant path. To simplify the design, the filter changes the lengths of P_1 and P_2 for evaluating the effects of three passbands. For an example as path 1, 1st passband (2.4GHz) and 3rd passband (5.2GHz) are shifted to lower frequency when P_1 is increased. Similarly, the resonant frequency of path 2, 2nd passband (3.5 GHz) and 3rd passband (5.2GHz) are shifted to lower frequency when P_2 is increased.

III. Result

The proposed filter was fabricated and then measured by an HP8510C Network Analysis. Fig. 4 shows the measured results and fabricated circuit of the proposed filter. The fabricated triple passband bandpass filter occupies a small size; around 9.5 mm × 20 mm, i.e., approximately by 0.23 λ_g × 0.11 λ_g. The filter has measured center frequency at 2.4,

TABLE I.
COMPARISONS WITH OTHER PROPOSED FILTER (λ_g IS THE GUIDED WAVELENGTH OF THE 1ST CENTER PASSBAND FREQUENCY)

	Ref. [1]	Ref. [2]	Ref. [5]	**Proposed filter**
Substrate height(mm) /ε_r	0.787 / 2.2	0.508 / 3.38	0.787 / 2.2	**0.787 / 2.2**
Passbands (GHz)	1.575 / 2.4 / 3.5	2.3 / 3.7 / 5.3	1.575 / 2.4 /3.5	**2.4 / 3.5 / 5.2**
$\lvert S_{11}\rvert$ (dB)	9 / 18.9 / 13.5	17 / 17 / 17	24 / 20 / 20	**17.3 /25.3 / 19.8**
$\lvert S_{21}\rvert$ (dB)	1.6 / 1.5 / 2.3	2.5 / 1.9 / 2.9	0.8 / 0.9 / 0.9	**1.76 / 0.95 / 1.5**
FBW (%)	5.2 / 3.8 / 4.6	3.8 / 6.8 / 5	15.87/3.87 / 2.57	**4.22 / 3.98 /3.43**
Circuit Size (mm^2) (λ_g* λ_g)	2793 (0.5904)	1238 (0.1904)	690 (0.0416)	**190 (0.0253)**

3.5 and 5.2 GHz. The measured passbands have insertion losses of 1.76 dB, 0.95dB and 1.5 dB, return losses of 17.3 dB, 25.3 dB and 19.8 dB and 3dB fractional bandwidth (FBW) of 4.22%, 3.98% and 3.43%, corresponding to 2.4, 3.5 and 5.2 GHz, respectively. The comparison of the filter with other reported work is summarized in Table I.

IV. Conclusion

In this paper, we proposed the multipath triple passband bandpass filter with high passband selectivity by using stub load - uniform impedance resonators (SL-UIRs). The resonators can generate two resonant paths to form triple closed passbands and no parasitic interactions producing between each passband. The resonators feature a simple configuration, effective design method, high passband selectivity and good band-to-band isolation. The proposed filter is suited for advanced multi-band wireless communication applications.

Acknowledgment

This work was supported by the Ministry of Science and Technology under Grant No. MOST 107-2221-E-168-012–MY2.

References

[1] W. Y. Chen, M. H. Weng and S. J. Chang, "A new tri-band bandpass filter base on stub-loaded step-load step-impedance reonator," *IEEE Mirow. Wireless Compon. Lett.*, vol. 22, no. 4, pp. 179-181, Apr. 2012.

[2] C. F. Chen, T. Y. Huang and R. B. Wu, "Design of dual- and triple-passband filters using alternately cascaded multiband resonators," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 9, Sep. 2006.

[3] IE3D Simulator, Zeland Software, Inc., 2002.

[4] J. S. Hong, *Microstrip Filters for RF/Microwave Application*, 2nd Edition, John Wiley & Sons, Inc., 2011.

[5] C. W. Tang and P. H. Wu, "Design of a planar dual-band bandpass filter," *IEEE Microw.Wireless Compon. Lett.,* vol. 21, no. 7, pp. 362-364, Jul. 2011.

A New Transparent Planar Reflector Antennas for Satellite DTV Applications

1st Yu-Ming Lin
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
corn0924@gmail.com

2nd Zong-Fu Li
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan
confused0314@gmail.com

3rd Tzu-Chun Tai
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
yuhshooulove84@gmail.com

4th Hung-Wei Wu
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan
qqq25q@gmail.com

5th Shoou-Jinn Chang
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
changsj@mail.ncku.edu.tw

6th Yeong-Her Wang
Department of Photonics
National Cheng Kung University
Tainan, Taiwan
yhw@ee.ncku.edu.tw

***Abstract*—This paper presents the new thick film coating technology for transparent conducting film (TCF). The coated thin film is having the microwave characteristic of high transparency and low resistivity. Design of traditional dipole antenna consisted of iron wire, which can be manufactured via the simple bending process as that for the reflector antennas. On the surface is designed, the planar reflecting layer of using our invented activation process is used to rearrange the direction of liquid nano-material the thin film from disorder to ordered arrangements. The optical transmittance of reflector layer can be achieved in 82%. The overall dimension size of the is 210 x 290 mm^2. The transparent reflective layer is prototyped can apply on the Ku-band for the satellite DTV.**

Keywords—Transparency, Planar reflector antennas, Satellite DTV, thin film

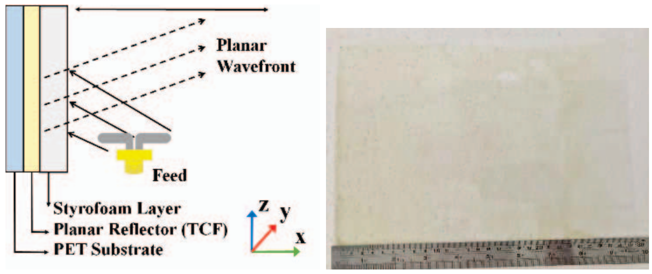
I. Introduction

In recent years, parabolic reflector antennas are in broadband wireless communication services received more attention as the applications to satellite DTV and Direct Broadcast media [1-3]. Conventional parabolic reflector antennas are efficient radiators, however they have to limited from their volume size, high cost, and large weight due to the curved reflector surface. The curved parabolic reflector is difficult to be manufactured, particularly at higher microwave frequencies. In the past, many researchers report about transparent conducting films, such as Indium tin oxide (ITO)、Gallium doped zinc oxide (GZO), and Aluminum doped ZnO (AZO) can be deposition on transparent dielectric substrate in place of metals for achieve transparent thin film [4-6].

This study we propose a new conducted liquid nano-material the thin film coating technique to make the reflector planar and transparency. The activation process can effectively improve the conducted liquid nano-material arrangement inside the thin film from disorder to order in around 3 mins treatment. The proposed ordered TCF has a low cost, high resistivity, and decrease reflector antennas weight. All analysis has been carried out using High Frequency Structure Simulator (HFSS), ANSYS [7].

II. Design of the Planar Reflector Antenna

Fig 1. shows the configuration the planar reflector and photograph of the proposed order TCF. The planar reflector is designed and fabricated on the PET substrate (Polyethylene ; dielectric constant $\varepsilon_r = 2.25$, loss tangent δ = 0.001 and thickness h = 0.01 mm). The developed thin film material has very low sheet resistance (Rs = 4.4 Ω/sq) after using our invented activation process. Fig 2. (a) is having disordered conducted liquid nano-material arrangement before using the activation process. Fig 2. (b) is having ordered liquid nano-material with using activation process. It is clearly observed that the arrangement of nano-structure is in ordered after activation process.



(a) (b)

Fig 1. (a) Configuration of the planar reflector and (b) photograph of the proposed ordered TCF.

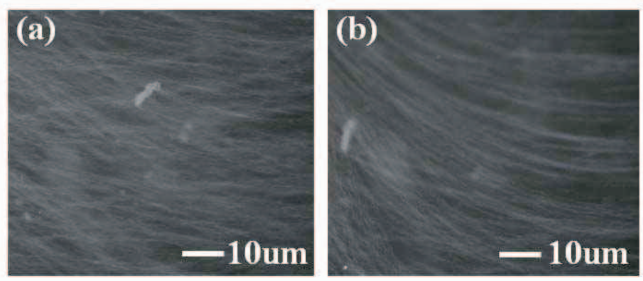


Fig 2. Optical microscope image of transparent thin film. (a) liquid nano-material arrangement without using activation process and (b) ordered liquid nano-material with using activation process.

Fig 3. shows the optical transmittance spectra of the transparent planar reflector with different case. The optical transmittance of PET substrate is around 91% over the wavelength from 400 – 800 nm. In ordered TCF case, the invented activation process is used to transfer the arrangement of nano-structure from disordered to ordered so as to obtain the performance of 82% transmittance at 550nm and low sheet resistance (Rs = 4.4 Ω/sq).

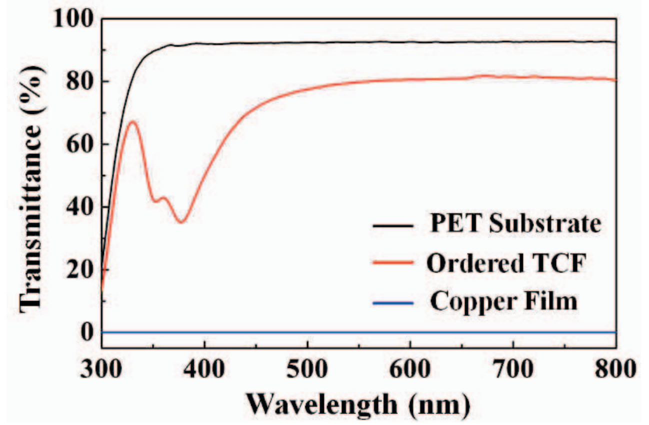


Fig 3. Optical transmittance spectra of the transparent planar reflector with different case.

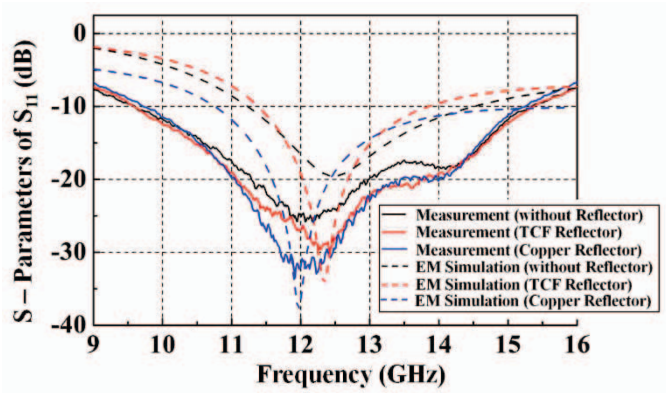


Fig 4. Comparison of simulation and measured results of different reflector case.

III. Results

The antenna was measured by an Agilent N5230A. Fig 4 shows the comparison of simulation and measured results of different reflector case. Good measured results include the center frequencies of 12.2 GHz with without reflector antennas $|S_{11}|$ of 26 dB and using ordered TCF reflector $|S_{11}|$ of 29.7 dB using copper film reflector $|S_{11}|$ of 33 dB. Fig. 5 and Fig 6. shows simulated and measured 2D radiation pattern of the proposed antennas at 12.2GHz. Some discrepancy can be observed between the measured results and simulated data, which are mainly contributed by two factors including: 1) the fabrication discrepancy and 2) thin air gap between the radiation layer boundary conditions. Table I summarized the comparison with other proposed reflector antenna.

IV. Conclusion

This paper presents a reflector antenna with high transmittance. The used conducted liquid nano-material is based on our invented activation process to transfer the disordered to ordered nano-material in the thin film, therefore, the high transmittance and low resistivity can be achieved. The measured return loss shows the fractional bandwidth of 46%. The proposed antenna with an overall compact size of 210 x 290 mm^2. The transparent reflector antenna has very promising to apply on the applications of Ku-band for the satellite DTV, and Smart cars.

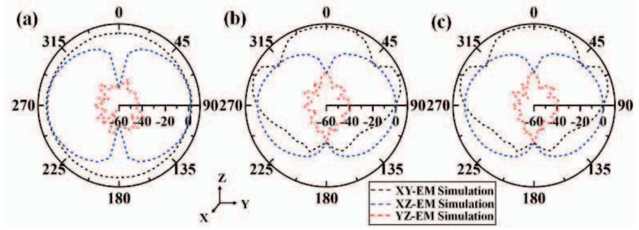


Fig 5. EM Simulation 2D radiation pattern at 12.2 GHz for (a) XY–plane, (b) XZ-plane, (c) YZ-plane of the antenna.

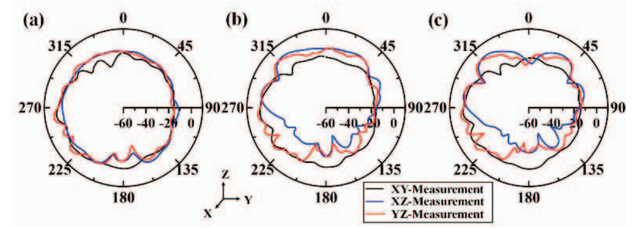


Fig 6. Measured 2D radiation pattern at 12.2 GHz for (a) XY–plane, (b) XZ-plane, (c) YZ-plane of the antenna.

TABLE I.
Comparisons With Other Proposed Reflector Antenna.

Ref.	Reflector substrate	f_0 (GHz)	$\lvert S_{11} \rvert$ (dB)	Bandwidth
[1]	Metal	12.2	x	x
[2]	Metal mesh	10.5	21	40%
[3]	Metal mesh	x	x	x
This study	Polyethylene	**12.2**	**29.7**	**46%**

Acknowledgment

This work was supported by the Ministry of Science and Technology under Grant No. MOST 107-2221-E-168-012-MY2.

References

[1] H.T. Chou, C.Y. Lin and M.H. Wu, "A high efficient reflectarray antenna consisted of periodic all-metallic elements for the Ku-band DTV applications," *IEEE Antennas Wireless Propag. Lett.*, vol.14, pp.1542-1545, 2015.

[2] M. Sanad and N. Hassan, "A low wind load lightweight dual cylindrical reflector antenna with a novel feed for direct broadcast satellite TV reception," *2010 Loughborough Antennas & Propagation Conference*, Nov., 2010, pp. 285-288.

[3] Alexander D, Henderson P and Turner G, "Advancements in large mesh reflector technology for multi-beam antenna applications," *European Conference on Antennas and Propagation,* pp. 410-412, 2014.

[4] D.P. Tran, H.I. Lu and C.K. Lin, "Conductive characteristics of Indium Tin Oxide thin film on polymeric substrate under Long-Term static deformation," *Coatings,* vol.8, issue 6, pp. 212, 2018.

[5] E. Muchuweni, T. S. Sathiaraj and H. Nyakotyo, "Effect of gallium doping on the structural, optical and electrical properties of zinc oxide thin films prepared by spray pyrolysis," *Ceram. Int.,* vol. 42, no. 8, pp. 10066–10070, Jun. 2016.

[6] S.F. Tseng, "Investigation of post-annealing aluminum-doped zinc oxide (AZO) thin films by a graphene-based heater," *Appl. Surf. Sci.,* vol. 448, pp. 163-167, Aug. 2018.

[7] High Frequency Structure Simulator *(HFSS)*, *Ansoft Corporation*, 2011.

Overflow-free realizations for LTI digital filters

Shuichi Ohno
Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima, Japan
Email: o.shuichi@ieee.org

Shenjian Wang
Hiroshima University, Japan

***Abstract*—This paper presents a design of overflow-free realizations for linear time-invariant digital filters. Under a sufficient condition for no-overflow, the noise gain from the quantization error to the filter output is minimized to obtain an overflow-free realization. A numerical example is provided to see the advantage of our design over the conventional design.**

***Index Terms*—digital filter, quantization, overflow**

I. Introduction

Digital filters are still often implemented by fixed-point arithmetic digital signal processors (DSP), since they require less implementation cost and smaller power consumption than floating point DSPs. Since real numbers are quantized to binary numbers with a finite precision, finite word-length (FWL) effects cannot be negligible when a sufficient number of bits cannot be assigned to each number.

There are two major effects resulted from the finite number of bits for quantization; overflow and roundoff error. The overflow occurs when the real number to be quantized does not lie in the dynamic range of the binary number. On the other hand, the roundoff error is the difference between the original real number and the rounded quantized binary number due to the finite precision of the binary number. There is a trade-off between the overflow and the roundoff error, since for a fixed number of bits, the increase of the dynamic range also increases the roundoff error. It is not that easy to balance the overflow and the roundoff error.

The digital filter is implemented according to its state-space realization. For a given transfer function of a digital filter, there exists an infinite number of state-space realizations. It is known that state-space realizations for an identical transfer function have different characteristic to quantization. To mitigate the overflow, state variables of state-space realization have to be scaled based on the impulse responses from the input of the digital filter to the state variables. The scaling by the l_2 norm of the impulse response has been utilized to assign the bits to minimize the variance of the roundoff noise [1]. However, not the l_2 norm scaling but the l_1 norm scaling guarantees no-overflow.

This paper presents overflow-free state-space realizations, based on the l_1 norm scaling. Since the l_1 norm is difficult to be evaluated, a sufficient condition for no-overflow is utilized as in [2]. Under this condition, the noise gain is minimized to obtain an overflow-free realization. A numerical example is provided to show that our design has smaller noise gains than the conventional design by the l_2 norm scaling.

II. System realization and quantization

Let us consider realizations of a single-input/single-output discrete-time linear time-invariant (LTI) digital filter, whose transfer function is a strictly proper rational function $H[\mathrm{z}]$ of order n in z^{-1}. The filter is realized according to space state matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, and $C \in \mathbb{R}^{1\times n}$ for equations:

$$x_{k+1} = Ax_k + Bw_k, \quad y_k = Cx_k, \tag{1}$$

where $w_k \in \mathbb{R}$ is the filter input, $x_k \in \mathbb{R}^{n\times 1}$ is the state vector, and $y_k \in \mathbb{R}$ is the filter output. The filter is assumed to be stable, i.e., all of the eigenvalues of A lie in the unit circle, reachable and observable.

The transfer function $H[\mathrm{z}]$ is expressed by the state-space expression (A, B, C) as

$$H[\mathrm{z}] = C(\mathrm{z}I - A)^{-1}B. \tag{2}$$

Let T be an $n \times n$ non-singular matrix. It is well-known that the change of variable defined as

$$\tilde{x}_k = T^{-1}x_k \tag{3}$$

leads to another state-space expression $(T^{-1}AT, T^{-1}B, CT)$. There exist an infinite number of realizations for a given transfer function.

When the filter is implemented as a real digital circuit, only a finite number of bits is assigned to describe a real number. If a sufficient number of bits cannot be assigned, roundoff errors cannot be neglected.

Let us consider the quantization of the state vector when a same type of binary expression is used for each entry of the state vector. Since the precision of each value is identical, we assume that roundoff errors have zero mean and an identical variance. The transfer function from the roundoff error to the filter output is given by $G[\mathrm{z}] = C(\mathrm{z}I - A)^{-1}$. Then, the effect of the roundoff error on the filter output can be evaluated by the system norm of $G[\mathrm{z}]$, which is called the noise gain.

Overflow is another major disadvantage of quantization. which occurs if the value to be quantized does not fall into the fixed given dynamic range. The overflow can cause a significant effect on the filter output, since it is generally unbounded. Our objective is to design an overflow-free realization that minimizes the noise gain for a given transfer function.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

III. Realizations by the l_1 norm scaling

Without loss of generality, we assume that the input w_k of the system is bounded such that

$$|w_k| \leq 1, \quad \forall k. \tag{4}$$

We utilize the H_2 norm $\|G[\mathrm{z}]\|_2$ to measure the noise gain. For given (A, C) matrices, the H_2 norm is given by

$$\|G[\mathrm{z}]\|_2^2 = \operatorname{trace} W_o \tag{5}$$

where W_o is the observability Gramian that satisfies the Riccati equation: $W_o = A^T W_o A + C^T C$.

We denote the transfer function of the input to the ith entry of the state vector $\tilde{x}_k$ as $\tilde{F}_i[\mathrm{z}]$, which is given by the the ith entry of $\tilde{F}[\mathrm{z}] = T^{-1}(\mathrm{z}I - A)^{-1}B$. Let the impulse response of the ith entry of $\tilde{F}[\mathrm{z}]$ as $\{\tilde{f}_i(k)\}_{k=0,1,2,\ldots}$. The l_1 norm of the impulse response is defined as

$$\|\tilde{f}_i\|_1 = \sum_{k=0}^{\infty} |\tilde{f}_i(k)|. \tag{6}$$

The absolute value for the ith state variable is bounded by $\|[\tilde{f}]_i\|_1$ such that

$$\sup_k |[\tilde{x}_k]_i| \leq \|\tilde{f}_i\|_1 \sup_k |w_k| = \|\tilde{f}_i\|_1 \tag{7}$$

where we denote the ith entry of $\tilde{x}_k$ as $[\tilde{x}_k]_i$. Let the maximum of the dynamic range for the state variables be γ_x. Then, if

$$\sup_k |[\tilde{x}_k]_i| \leq \gamma_x \tag{8}$$

for $i \in [1, n]$, any overflow never occurs. It follows from (7) and (8) that if

$$\|\tilde{f}_i\|_1 \leq \gamma_x \tag{9}$$

for every $i \in [1, n]$, then there is no overflow. The condition on the l_1 norm of the impulse response is called the l_1 norm scaling. On the other hand, the l_2 norm of the impulse response is defined as $\|\tilde{f}_i\|_2 = \left(\sum_{k=0}^{\infty} |\tilde{f}_i(k)|^2)\right)^{\frac{1}{2}}$ and the scaling by the l_2 norm is called the l_2 norm scaling.

In this paper, we use a sufficient condition for the l_1 norm scaling to avoid overflow. Let $\mathcal{E}(P)$ be the ellipsoid defined by an $n \times n$ real symmetric positive definite matrix $P \succ \mathbf{0}$ as $\mathcal{E}(P) = \{x \in \mathbb{R}^n : x^T P x \leq 1\}$. When $T^{-1} = \gamma_x P^{\frac{1}{2}}$, it follows from $x_k^T P x_k = x_k^T T^{-T} T^{-1} x_k / \gamma_x^2 = \tilde{x}_k^T \tilde{x}_k / \gamma_x^2$ that if $x_k^T P x_k \leq 1$, then $\tilde{x}_k^T \tilde{x}_k \leq \gamma_x^2$. This implies that if $x_k^T P x_k \leq 1$, then $\sup_k |[x_k]_i| \leq \gamma_x$ and hence any overflow never happens. Using this, we would like to find a non-negative matrix P that minimizes the noise gain $\operatorname{trace}(T^T W_o T) = \operatorname{trace}(T T^T W_o) = \operatorname{trace}(P^{-1} W_o)$ subject to $x_k^T P x_k \leq 1$.

We denote the spectrum radius of A as $\rho(A)$. Let $P(\alpha)$ is a solution of the Riccati equation given by

$$\frac{1}{1-\alpha} A P^{-1}(\alpha) A^T - P^{-1}(\alpha) + \frac{1}{\alpha} B B^T = \mathbf{0} \tag{10}$$

for $\alpha \in (0, 1 - \rho^2(A))$. Then, using the results in [3], we can show that our problem is equivalent to the following optimization problem:

$$\min_{\alpha \in (0, 1-\rho^2(A))} \operatorname{trace}(P(\alpha)^{-1} W_o). \tag{11}$$

TABLE I
Noise gain for proposed/conventional realization

	full re-scaling	proportional re-scaling
l_1 norm scaling	6.285	7.992
l_2 norm scaling	11.362	14.896

For a given α, we can obtain the solution of the Riccati equation (10). Thus, we can find the optimal α using a numerical search for $\alpha \in (0, 1 - \rho^2(A))$.

IV. Numerical Example

We compare our realization with the realization by the l_2 norm scaling. Our realization by the l_1 norm scaling is sub-optimal, that is, the l_1 norms of the obtained realization are below 1, whereas the realization by the l_2 norm scaling does not satisfy the l_1 norm scaling condition in general. We need re-scaling for both realizations to prevent overflows.

We have tested two kinds of re-scaling. Let $\hat{T}$ be the transformation matrix of an obtained realization and $\hat{f}_i(k)$ be the impulse response of the ith entry of $\hat{F}[\mathrm{z}] = \hat{T}(\mathrm{z}I - A)^{-1}B$.

One re-scaling is $T = \Lambda \hat{T}$, where Λ is a diagonal matrix given by

$$\Lambda = \operatorname{diag}(\|\hat{f}_1\|_1, \|\hat{f}_1\|_2, \ldots, \|\hat{f}_1\|_n). \tag{12}$$

For this re-scaling, the equality holds in (8) for every $i \in [1, n]$. The other re-scaling is $T = \max_{i \in [1,n]} \|\hat{f}_i\|_1 \hat{T}$ with which the equality holds in (8) for one $i \in [1, n]$ and the inequality holds for the remaining i. For convenience, we refer to the former re-scaling as the full re-scaling and the latter as the proportional re-scaling.

We consider a 6th-order lowpass digital Butterworth filter with normalized cutoff frequency 0.5. Table I shows the noise gains of the l_1 and the l_2 norm scaling with the full and the proportional re-scaling. For each re-scaling, the realization by the l_1 norm scaling exhibits smaller noise gains than the realization by the l_2 norm scaling.

V. Conclusion

We have developed a design of overflow-free realizations for digital filters. A numerical example has been provided to demonstrate its effectiveness.

Acknowledgment

This work was partially supported by JSPS KAKENHI Grant Number 16H02921.

References

[1] C. Mullis and R. Roberts, “Synthesis of minimum roundoff noise fixed point digital filters,” *IEEE Transactions on Circuits and Systems*, vol. 23, no. 9, pp. 551–562, Sep 1976.

[2] S. Ohno and Y. Yoshimura, “State space realizations robust to overloading for discrete-time LTI systems,” *Signal Processing*, vol. 156, pp. 12–20, 2019.

[3] H. Shingin and Y. Ohta, “Optimal invariant sets for discrete-time systems: Approximation of reachable sets for bounded inputs,” in *10th IFAC/IFORS/IMACS/IFIP Symposium on LSS, 2004, pp. 401–406.*

Proposal of DTW distance calculator using neuron CMOS inverter

Takanori Kurano
The Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
8bjnm009@mail.u-tokai.ac.jp

Kaiki Yamada
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619 ,Japan

Ryo Yamamoto
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan

Hayato Yagi
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan

Nao Onji
The Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan

Masaaki Fukuhara
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
fukuhara@tokai.ac.jp

***Abstract*—In recent years, various data are used with the development of IoT (Internet of Things). In similarity search, DTW (Dynamic Time Warping) distance is often used. However, the calculation of the DTW distance has a problem that the amount of calculation becomes enormous. In order to solve this problem, many approaches of algorithm improvement have been proposed. We consider that the DTW distance can be calculated more quickly by using dedicated circuits for various improved algorithms. In this paper, we propose a circuit to calculate the DTW distance. In addition, we perform HSPICE simulation of the proposed circuit.**

***Index Terms*—DTW distance, neuron CMOS inverter, similarity search**

I. INTRODUCTION

In recent years, big data has been used in various fields with the development of IoT (Internet of Things). In order to stack such data efficiently, a system is necessary to analyze continuously the stream data (sequence data) sent from some sensors. Similarity search using DTW (Dynamic Time Warping) distance is one of the methods used to analyze stream data. However, the calculation of the DTW distance has a problem that the amount of calculation increases if the amount of data is vast. Many algorithm improvements have been proposed to solve this problem [1,2]. We consider that the amount of calculation of the DTW distance can be reduced by applying the improved algorithm with dedicated hardware. However, few papers have been proposed for hardware to execute the DTW algorithms. In this paper, we design a circuit to calculate the DTW distance between two sequence data. From the simulation results by HSPICE, it is confirmed that the proposed circuit can obtain the same result as the DTW definition shown in Section II.

II. PROPOSED CIRCUIT TO CALCULATE DTW DISTANCE

The definition of DTW distance described in [1] follows below. For a sequence X=(x_1,x_2,x_3,. . .,x_n) of length n and a sequence Y=(y_1,y_2,y_3,. . .,y_m) of length m, the DTW distance $D(X,Y)$ is defined by

$$D(X,Y) = f(n,m)$$

$$f(t,i) = |x_t - y_i| + Min\begin{cases} f(t,i-1) \\ f(t-1,i) \\ f(t-1,i-1) \end{cases} \quad (1)$$

$$f(0,0) = 0, f(t,0) = f(0,i) = \infty$$
$$(t = 1,\ldots,n; i = 1,\ldots,m).$$

The circuit which performs the equation (1) is shown in Fig. 1. The 8-bit SRAM$_i$(i = 1, 2, 3,. . . ,m) and the address decoder are circuits for storing the sequence Y. The difference absolute circuit calculates a difference absolute value between the sequence Y and x_t of the sequence X which is stored to the register A_1 to A_8. The ADL (After Difference Latch) is stored the calculated value. The value of ADL and the value of $Min.\ circuit$ are added, and the results of adder is outputted through the registers of AAL (After Adder Latch) and DRL (DTW Result Latch).

The $Min.\ circuit$ compares three values expressed as equation (1). For that reason, the $Min.\ circuit$ is constructed with neuron CMOS inverters, which can take appropriate data out of multiple data at one time [3]. The neuron CMOS inverter can calculate the weighted sum of all input signals at the gate level, and controls the inversion behavior based on the result of such weighted sum calculation.

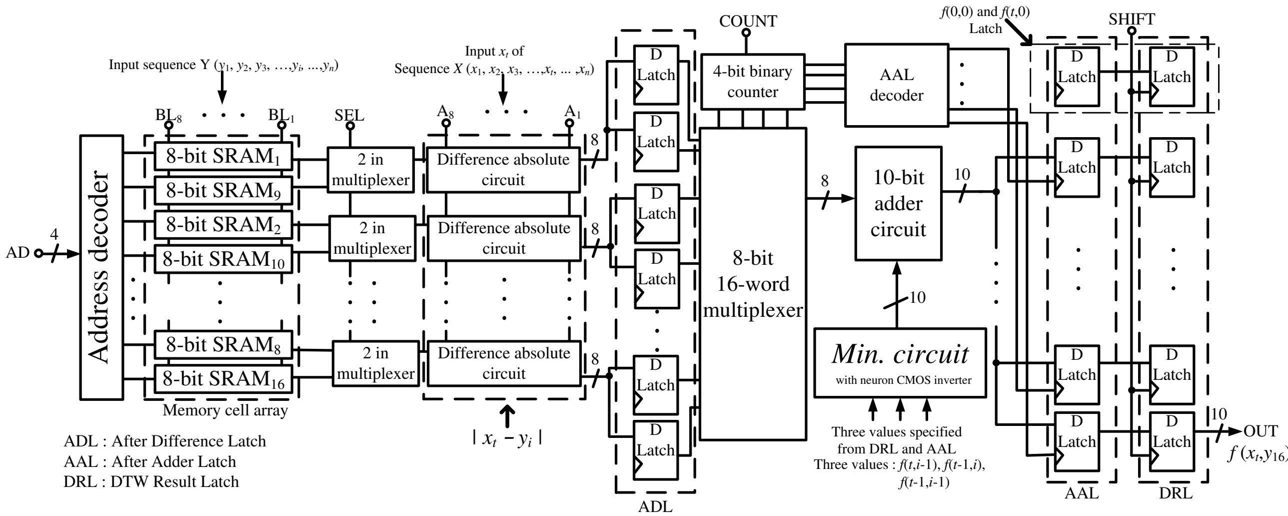


Fig. 1. Block diagram of proposed circuit to calculate DTW distance between sequence X and sequence Y

Next, each control signals are described. AD is for the address decoder. SEL specifies data to be subtracted, and is a signal for writing data of $|x_t - y_i|$ in DLatch of the ADL part. COUNT is for operating i of $f(t, i)$ to be calculated the difference absolute value. SHIFT is for writing distance data to the DRL after writing of AAL data is completed.

The proposed circuit stores the data of sequence Y in 8-bit SRAM$_i$. Thereafter, the data of sequence X is inputted in order. Then, the proposed circuit outputs DTW distance $f(x_t, y_{16})$ of sequence X and sequence Y.

III. Simulation Result and Conclusion

The proposed circuit was simulated by HSPICE using ROHM 0.18 μm CMOS process to design neuron CMOS inverter. In the simulation, $V_{DD} = 1.8$ [V] and $GND = 0$ [V]. Sequence X and sequence Y are input data according to TABLE I. Since the values shown in Fig. 2 and the values shown in TABLE I are the same, it can be confirmed that the circuit operation can be sufficiently performed. Further, even if the number is larger than the sequence X, the calculation can be performed similarly.

We proposed a circuit that can calculate DTW distance. From simulation results, we confirmed that the proposed circuit is the same as the solution obtained by equation (1).

In the future, we would like to confirm the usefulness by prototyping this circuit and comparing with software.

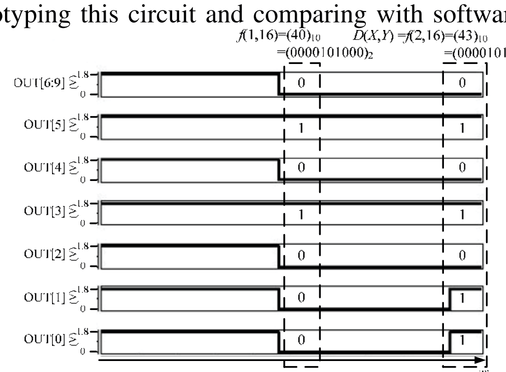


Fig. 2. HSPICE Simulated result of proposed calculator for DTW distance

TABLE I
DTW DISTANCE CALCULATION TABLE

sequence Y	16	y_{16}= 3	∞	40	$\boxed{43}$
	15	y_{15}= 1	∞	38	40
	14	y_{14}= 7	∞	38	33
	13	y_{13}= 3	∞	32	35
	12	y_{12}= 1	∞	30	32
	11	y_{11}= 7	∞	30	25
	10	y_{10}= 3	∞	24	27
	9	y_9= 1	∞	22	29
	8	y_8= 1	∞	22	24
	7	y_7= 7	∞	22	17
	6	y_6= 3	∞	16	19
	5	y_5= 1	∞	14	16
	4	y_4= 7	∞	14	9
	3	y_3= 3	∞	8	11
	2	y_2= 1	∞	6	13
	1	y_1= 7	∞	6	7
	0	y_i ↑	0	∞	∞
	i ↑	x_t →		x_1=1	x_2=8
		t →	0	1	2
$\boxed{}$	: $f(n,m) = D(X,Y)$			sequence X	
(shaded)	: $f(t,i)$				

Acknowledgment

This work is supported by VLSI Design and Education Center(VDEC), the University of Tokyo in collaboration with Synopsys, Inc., Cadence Design Systems, Inc., and Rohm Corporation. This work is supported by Tokai University General Research Organization Grant.

References

[1] Y. Sakurai, C. Faloutos, and M. Yamamuro, ”Stream Processing under the Dynamic Time Warping Distance”,The IEICE transactions on information and systems (Japanese edition) 92(3), pp.338-350, March 2009

[2] M. Toyoda, Y. Sakurai, and Y. Ishikawa, ”An efficient method for pattern discovery in data streams”, IFAT,2012-IFAT-107(9), pp.1-8, 2012

[3] M. Fukuhara, N. Onji, T. Kurano, T. Nakajima, Y. Harada, K. Fujimoto, and M. Yoshida, ”Short-Circuit-Current Reduction by using a Clocked Neuron CMOS Inverter in a Time-Domain Data Coincidence Detector,” ICIC Express Letters, Part B, Applications: An International Journal of Research and Surveys, Vol. 9, No. 6, pp. 477-484, June 2018

A reduction of current consumption of a Hamming distance detector by improvement of current mirror circuit

Nao Onji
Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
8bjnm007@mail.u-tokai.ac.jp

Takanori Kurano
Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
8bjnm009@mail.u-tokai.ac.jp

Toyoki Saho
Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
9bjnm015@mail.u-tokai.ac.jp

Riku Ohtsuka
Graduate School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
9bjnm004@mail.u-tokai.ac.jp

Kaiki Yamada
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan

Masaaki Fukuhara
Department of Embedded Technology, School of Information and Telecommunication Engineering
Tokai University
Tokyo 108-8619, Japan
fukuhara@tokai.ac.jp

***Abstract*— A Hamming distance detector using a neuron CMOS inverter and a current mirror circuit is capable of data match search. The constant current of the current mirror circuit in the detector occupies almost the current consumption of the detector. In this paper, we reduce the constant current by adding an NMOS switch to the current mirror circuit. By performing HSPICE simulation of the detector with the NMOS switch type current mirror circuit, we confirm that the current consumption of the detector is reduced compared with a conventional detector.**

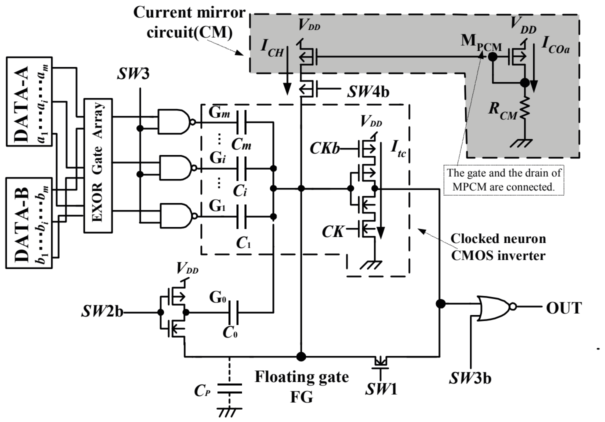
***Keywords*— constant current, Hamming distance detector, current mirror**

I. Introduction

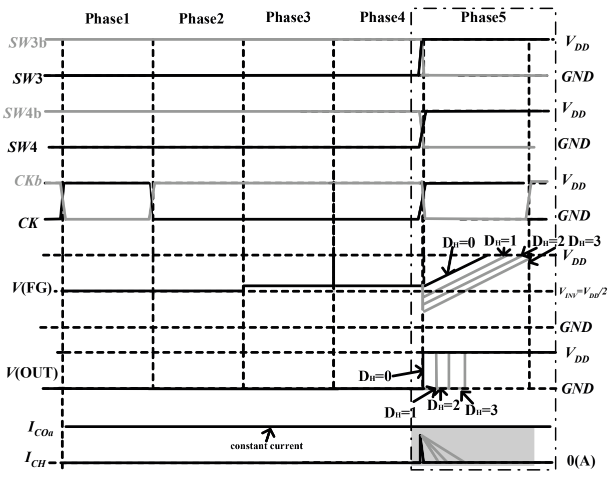
Similarity search can be performed by applying Hamming distance detector using a neuron CMOS inverter [1] and a current mirror (CM). The detector has problems with the through current of the neuron CMOS inverter and the constant current of the CM. The through current was reduced by using a clocked neuron CMOS inverter in the detector [2,3]. On the other hand, the constant current of the CM is large because the PMOS used in the CM is always in a saturation region. In this paper, we analyze the constant current of the CM by hand calculation and propose an NMOS switch type CM. Finally, we confirm from the HSPICE simulation results that the current consumption of the proposed detector with NMOS switch type CM can be reduced and compare the detector that is not used NMOS switch [3].

II. Hamming Distance Detector

Fig. 1 illustrates the circuit configuration and timing chart of the Hamming distance detector using CM. This detector outputs a voltage V(OUT) according to the Hamming distance D_H between DATA-A (a_1, a_2, ..., a_m) and DATA-B (b_1, b_2, ...b_m) in the time domain. As shown in Fig. 1(a), the gate and the drain of the PMOS(M_{PCM}) in the CM are connected. As shown in Fig. 1(b), the charging current I_{CH} flows to the floating gate during Phase5, but the constant current I_{COa} always flows from V_{DD} to GND.



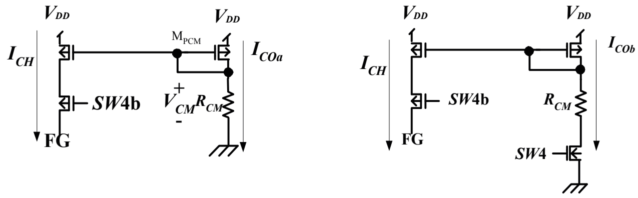
(a) Schematic of circuit



Note: The timing of a_i,b_i (i=1,2,...,m), SW1 and SW2b are shown in Ref. [3].

(b) Timing chart

Fig. 1. Hamming distance detector



(a) Conventional (b) NMOS switch type

Fig. 2. Schematic of Current mirror (CM)

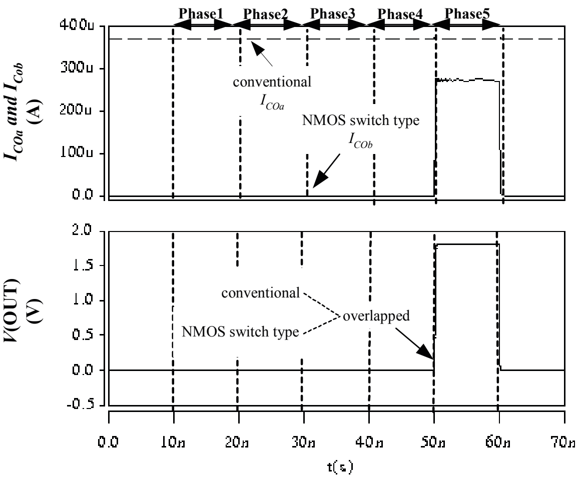


Fig. 3. Simulation waveforms

III. Analysis of Current Mirror

Fig. 2(a) shows conventional current mirror. We analyze I_{COa} in Fig. 2(a) by hand calculation. When we assume that the M_{PCM} is in the state of "$V_{CM}<V_{DD}-|V_{THPCM}|$", M_{PCM} is operating in ON condition, where V_{CM} is a voltage applied to R_{CM}, and V_{THPCM} is threshold voltage of M_{PCM}.

As the gate and the drain of M_{PCM} are connected, M_{PCM} is in the saturation region and the I_{COa} of M_{PCM} is expressed as follows:

$$\boldsymbol{I_{COa} = \frac{\beta_p}{2}(V_{DD} - V_{CM} - |V_{THPCM}|)^2}, \qquad (1)$$

where β_p is the gain factor for M_{PCM} . From Ohm's law I_{COa} is expressed as follows:

$$\boldsymbol{I_{COa} = \frac{V_{CM}}{R_{CM}}}\ , \qquad (2)$$

where R_{CM} is the resistance. Equation (3) can be obtained by equalizing equations (1) and (2).

$$\boldsymbol{V_{CM} = V_{DD} - |V_{THPCM}| + \frac{1}{\beta_p R_{CM}} - \sqrt{\frac{2}{\beta_p R_{CM}}(V_{DD} - |V_{THPCM}|) + \left(\frac{1}{\beta_p R_{CM}}\right)^2}} \qquad (3)$$

I_{COa} is obtained by inserting equation (3) into equation (2). The theoretical equation of the I_{COa} shows how much current flows. For reducing the $I_{COa,}$ we propose Hamming distance detector with NMOS switch type CM as shown in Fig. 2(b). I_{COb} is constant current of Fig. 2(b). The timing of the NMOS switch is the timing of $SW4$ in Fig. 1(b).

TABLE I shows the parameters used in hand calculation and the simulation. We performed hand calculation and HSPICE simulation with Rohm 0.18 µm CMOS process design rule. The simulation times in TABLEs II and III are from Phase1 to Phase5 in Fig. 3.

TABLE I. Device parametr

symbol	value	symbol	value
V_{DD}	1.8(V)	W_n / L_n	1.0(µm) / 180.0(nm)
GND	0(V)	W_p / L_p	3.0(µm) / 180.0(nm)
R_{CM}	1(kΩ)		

TABLE II. Comparison of constant current I_{COa}

Calculation value	Simmulation value	Relative error
467(µA)	369(µA)	21%

TABLE III. Comparison about average current (I_{COa} and I_{COb}) and average current consumption of all of detector

	Conventional	NMOS switch type	reduction rate
Average current	I_{COa}=369(µA)	I_{COb} = 53.6(µA)	85%
Average current consumption of all of detector	386(uA)	71.6(µA)	81%

TABLE Ⅱ compares the calculated values obtained by substituting TABLE I and equation (3) into equation (2) with the simulated values of the conventional CM. The 21% error is due to the MOS model difference between hand calculation and simulation.

TABLE III summarizes the average current (I_{COa} and I_{COb}) and the average current consumption of all of the detector under the conventional CM and the NMOS switch type CM, respectively. Fig. 3 shows the simulated waveforms of the average current (I_{COa} and I_{COb}), and V(OUT) when D_H is 1. From TABLE III and Fig. 3, the constant current of CM and consumption current of all of the Hamming distance detector using the neuron CMOS inverter and the CM has been significantly reduced. From V(OUT) of Fig. 3, the operation of the Hamming distance detector is not affected by using the NMOS switch type CM.

IV. Conclusion

We studied to further reduce of current consumption of a Hamming distance detector using neuron CMOS inverter with an NMOS switch type CM and. In the future, we theoretically analyze the effect of the Hamming distance detector using this NMOS switch CM.

Acknowledgment

This work is supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Synopsys, Inc., Cadence Design Systems, Inc., and Rohm Corporation. This work is supported by Tokai University General Research Organization Grant.

References

[1] K. Kotani, T. Shibata, M. Imai, and T. Ohmi, "Clocked-Neuron-MOS Logic circuits Employing Auto-Threshold-Adjustment," in IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers, pp. 320-321, San Francisco, Feb. 1995.

[2] Y. Harada, K. Kuniaki, M. Fukuhara, and M. Yoshida, "A Minimum Hamming Distance Search Associative Memory Using Neuron CMOS Inverters," Electronics and Communications in Japan, Vol. 100, No. 3, pp. 10-18, Mar. 2017.

[3] M. Fukuhara, N. Onji, T. Kurano, T. Nakajima, Y. Harada, K. Fujimoto, and M. Yoshida, "Short-Circuit-Current Reduction by using a Clocked Neuron CMOS Inverter in a Time-Domain Data Coincidence Detector," ICIC Express Letters, Part B, Applications: An International Journal of Research and Surveys, Vol. 9, No. 6, pp. 477-484, June 2018.

Study and explore on the energy harvesting of the solar cell with DC/DC converter PWM system

Jian-Chiun Liou*[1], and Zhen-Xi Chen[1]
[1]School of Biomedical Engineering, Taipei Medical University

ABSTRACT

Energy harvesting is generally defined as the collection of energy sources such as light, heat, vibration, and electromagnetic waves, and converting them into electricity to provide the equipment itself with enough energy to maintain normal operation. This research uses sunlight to convert into energy output. It controls the voltage of the output by controlling the duty of the PWM technology. The system requires a DC/DC design and the output has a large capacitance. And the DC/DC post-stage filtering must be good. CIGS solar cells (efficiency of about 20%) are composed of Cu, In, Ga, and Se in a specific ratio, and have a wider wavelength range and absorption coefficient than tantalum materials.

INTRODUCTION

Since a wide variety of energy can be collected from the natural environment, and these resources can be converted into the electrical energy we need in order to be applied to a wide variety of electrical products. However, due to the difference in power generation technology, it is necessary to collect energy according to individual environmental design, and design a cost-effective power system to achieve maximum energy collection efficiency. The energy obtained through the natural "phenomena" of these surroundings is very small, and the output power that can be provided is quite limited. Therefore, in most applications, energy harvesting technology is combined with sensors and wireless components to create a The acquired power supply components are used by themselves or transmitted to other portable devices that can use these tiny powers. In terms of technical form, energy harvesting can be divided into four key technologies: solar energy, electric (Electrodynamos), piezoelectric (piezoelectricity) and thermoelectric (Thermoelectrics). According to IDTechEx research, the market size of the overall energy harvesting equipment is expected to exceed US$5 billion in 2022. Solar energy will continue to dominate consumer applications, but other energy harvesting technologies will find a broader application market in the industrial sector[1-4].

Solar energy can also be converted to electricity in one of two ways:

(1) Solar photovoltaic: Solar photovoltaic uses technologies such as photovoltaic (PV) panels on the roof and PV concentrating systems to convert sunlight directly into electrical energy.

(2) Solar thermal energy: Solar thermal energy uses a lens and a reflector to collect sunlight, thereby generating electricity using a steam turbine.

ARCHITECTURE

The process of converting solar energy into electrical energy and thermal energy (such as hot water, steam, or direct heat production) has many characteristics that make solar energy different from other energy sources. Unlike most other energy sources, solar energy can use photovoltaic materials to generate electricity directly without the need for mechanical conversion. From single solar panels on the house to large-scale power plants that power hundreds of thousands of homes, solar energy is the only energy source of all sizes.

When light strikes the pn junction, energy is transferred to the electrons of the helium atom. Some of the electrons thus have enough energy to leave the atom and become free electrons. When an electron leaves an atom, a "vacancy" appears in the position of the electron in the atom, called a "hole." The so-called cavity is actually just an atom that lacks an electron, so it is positively charged as a whole. The intrinsic electric field of the pn junction applies the free electrons to the side of the n-type crucible. The intrinsic electric field also moves the electrons of adjacent atoms, continuing to fill the "vacancy" of electrons in adjacent atoms, so the holes appear to move toward the side of the p-type enthalpy. The free electrons and holes move in opposite directions, separating positive and negative charges, creating a potential difference across the pn junction. Therefore, when the pn junction is connected to the external circuit through the upper and lower metal contact layers, the circuit provides a path for electrons to pass through and recombine with the holes at the other end of the pn junction. Current is generated in the circuit.

Compound semiconductor solar cells Such solar cells are mainly made of compound semiconductors formed by three or five or six groups, such as cadmium telluride, copper indium selenide, copper indium gallium selenide, gallium arsenide, and the like. Compared with silicon materials, these compound semiconductors are usually direct energy gaps, which can directly absorb photon energy and have a high absorption coefficient. Therefore, solar cells can be made thinner to reduce the chance of recombination of electrons. Better conversion efficiency, such as solar cells made of gallium arsenide, has exceeded 28% efficiency, but the process cost is relatively high. With tandem solar cells consisting of two or three different energy gap materials, the efficiency has even exceeded 40%. In this study, CIGS solar

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

cells (efficiency of about 20%) are composed of Cu, In, Ga, and Se in a specific ratio, and have a wider wavelength range and absorption coefficient than tantalum materials, so that solar cells can be made thinner and reduced. Defects cause non-composite luminescence, which makes this solar cell perform better even on cloudy days, and it has a longer service life than other thin-film solar cells.

The power supply commonly used in single-chip microcomputers is DC5V or DC3.3V or DC1.8V. Usually when designing products, the nominal voltage of the products is DC9V, DC12V, DC24V, etc. The voltage required to convert the input nominal voltage into the circuit is the process of the power supply design. We are constantly exposed to AC/DC and DC/DC. AC/DC converts AC power to DC power. For example, a mobile phone charger converts AC220V to DC5V. DC/DC converts DC to DC, for example, converting DC24V to DC3.3V. For the case where the output and input voltage difference is not large, the LDO solution can be used. For example, if DC12V is converted to DC5V, the LDO chip can be used. The advantages of the LDO are: 1) the peripheral circuit is simple, and 2) the text wave is easy to control. For the case of a large differential pressure, a DC/DC wafer scheme can be used. This article focuses on the DC/DC solution. There are two types of DC/DC power supplies as shown in Fig.1, one is isolation and the other is non-isolated. The GND of the isolated DC/DC output is unrelated to the input GND, also known as the floating supply. Most common DC/DC chips are non-isolated. First, let's talk about the non-isolated DC/DC principle. This type of power supply is divided into boost/buck, which is the boost/buck mode.

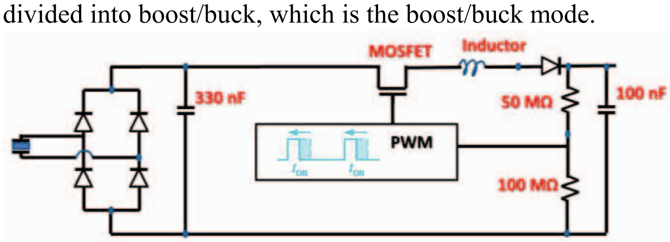


Fig.1 DC/DC power supply with PWM system

RESULTS

We analyze the working principle. When the power MOS (hereinafter referred to as the switch) is closed, the power supply supplies power to the load through the inductor, and stores the energy in the inductor L and the output capacitor. Due to the self-inductance of the inductor L, when the switch is closed, the current increases more slowly. Its output does not immediately reach the voltage value of the power supply. After a certain period of time, the switch is turned off. Due to the self-inductance of the inductor L (which can be visually assumed to have an inertial action in the inductor), the current in the circuit will remain unchanged and the current will continue to flow from left to right. The current flows through the load, returns from the ground. It remains to the Schottky diode collection, and returns to the left end of the inductor L through the diode, thereby forming a loop. The output voltage can be controlled by controlling the duty cycle of the PWM. When the switch closes the device, the inductor stores energy and releases energy during the disconnection. Therefore, the inductor L is called the storage inductor, and the diode is responsible for supplying the current path to the L during the disconnection of the switch, so the diode is called a freewheeling diode. When the switch is closed, the voltage is small, so the heating power U*I will be small, which is the reason for the high efficiency of the switching power supply. The solar cell with DC/DC Converter PWM system is shown in Fig.2. This solar energy harvesting is used to drive the fan.

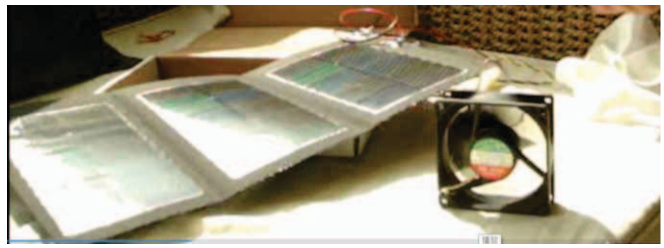


Fig.2 Solar cell with DC/DC Converter PWM System

CONCLUSIONS

The system platform includes a DC/DC converter, energy harvesting. Through this principle, we know why in the DC / DC design, the output must have a large capacitance, why the diode and inductor must be close to the IC. Moreover, the DC/DC post-stage filtering must be good because there is a switching frequency inside and the noise is very large.

ACKNOWLEDGMENTS

This work was financially supported of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (DP2-108-21121-01-O-05-04).

REFERENCES

[1]Hong-Yi Huang ; Shao-Zu Yen ; Jhen-Hong Chen ; Hao-Chiao Hong ; Kuo-Hsing Cheng,"Low-voltage indoor energy harvesting using photovoltaic cell",2016 IEEE 19th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS),pp.1-4(2016).

[2]Moeen Hassanalieragh ; Tolga Soyata ; Andrew Nadeau ; Gaurav Sharma,"UR-SolarCap: An Open Source Intelligent Auto-Wakeup Solar Energy Harvesting System for Supercapacitor-Based Energy Buffering",IEEE Access,Volume: 4,pp.542-557(2016).

[3]Saswat Kumar Ram ; Sauvagya Ranjan Sahoo ; Sudeendra K ; Kamalakanta Mahapatra,"Energy Efficient Ultra Low Power Solar Harvesting System Design with MPPT for IOT Edge Node Devices",2018 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS),pp.1-4(2018).

[4]Jian-Chiun Liou ; Te-Jen Su ; Wei-Jie Wen ; Wen-Chieh Lin,"A novel printhead multiplexer data registration chip system",2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE),pp.1-2(2016).

Inkjet technology addressing and precise control of DNA liquid

Jian-Chiun Liou*[1], and Zhen-Xi Chen[1]
[1]School of Biomedical Engineering, Taipei Medical University

ABSTRACT

The ceramic piezo printing system mechanism can print almost any type of ink, including water-soluble materials, solvent inks, and the above-mentioned heated solid material sprays. Liquid material for ceramic piezoelectric printing, with wide liquid viscosity and tension characteristics (viscosity can be adjusted between 1-18 mPa.s. Tension is between 35-45mN/m). Ceramic piezoelectric inkjet technology has a broader development prospect than thermal bubble printing technology. The study shows inkjet technology addressing and precise control of the amount of droplets. Logic multiplexed data processing integrates high voltage drive transducer array.

INTRODUCTION

The main working methods of liquid inkjet printers are divided into thermal foaming and ceramic piezoelectric. The technology of heating inkjet to maintain stable viscosity and surface tension has been applied to liquid inkjet in the past decade to adapt to different (regional) printing environments, such as tropical/temperate/cold belt, season/day and night. The change in temperature difference ensures the quality of the print and the reproducibility of the color of the print. However, due to its unique printing characteristics, that is, the solid ink maintains a stable temperature during the process of melting into a liquid state, so the viscosity and surface tension of the liquid ink remain relatively stable, and such characteristics can be varied in various ways. Non-absorbent materials (with different surface energies) are printed, that is, sprayed dry, and the selectivity to the printing material is not critical. In China, ceramic inkjet printing technology has set off a wave of frenzy, and inkjet printing technology for refrigerators and washing machine casings is also in production. Inkjet printing technology will greatly change the clarity and freedom of the pattern, which is a typical representative of high-tech. It could be said that ceramic piezoelectric digital printing technology is only in juvenile period, and thermal bubble printing technology has entered the old age.

For foaming technology, by heating the nozzle, the liquid is caused to generate thermal bubbles, which are sprayed onto the printing carrier medium. It is a high frequency, high-resolution printing technology. The working principle is using a thin film high-resistance resistance element (30Ω/□). The ink of less than 5 pico-liter is instantaneously heated to above 300 °C in the ink ejection area, forming numerous tiny bubbles, and the bubbles are at a very fast speed (less than 10 micrometers seconds) gather into large bubbles and extend the extrusion to force the droplets to eject from the nozzle opening. After the heating resistor is actuated to hold the bubble for 3-5 micro second, the liquid will be back-filled around the cavity without pressure. After the bubble disappears, the liquid ink of the nozzle shrinks back. Then the surface tension will generate suction, and the new ink will be added to the ink ejection area chamber to prepare for the next loop printing[1-3]. The ink close to the nozzle portion is heated and cooled by the heating resistors. The temperature of the thermal resistance array accumulated in the entire wafer is continuously increased (to 30-50°C), and thus the ink circulation cooling of the ink crucible cooling circulation structure is required. But in the system operation column in the printing, the liquid in the entire storage liquid body will remain at about 40-50°C. Working under such severe conditions, the ink jet used in the thermal bubble jet machine requires extremely high stability, resulting in high cost, and the service life of the micro-resistance heater is relatively shortened, due to the long-term alternating heat and cold work. It causes inevitable corrosion and is scrapped. Since thermal bubble printing is performed at higher temperature conditions, the liquid of its ink jet architecture must be designed for low viscosity characteristics (modulation of approximately < 1.5 mPa.s) and high tensile characteristics (modulation of > 40 mN/ m), to ensure long-term operation and continuous high-frequency high-speed printing.

ARCHITECTURE

For micro piezoelectric print head, it has stronger control ability for ink droplets. The ink sprayed by this technology is only 1/3 of the normal, the minimum can be 2pl, and the shape of the droplets is more regular, the positioning is more accurate, and the printing resolution is improved. High-precision printing, and no need to heat the micro-piezoelectric inkjet, the ink will not be chemically changed due to heat, which reduces the requirement for ink. Because the temperature is continuous and stable, the ink is continuously printed for a long time without interruption. The viscosity and surface tension remain relatively stable, ensuring consistency before and after the quality of the print, and the uniformity of the reproduction of the customer's reprint (replace order). Water-soluble inkjets made by anti-corrosion technology are used for printing on hot bubble and ceramic piezoelectric printers. The life of thermal bubble nozzles is about 8,000-10,000 ml, which needs to be replaced. For ceramic piezoelectric nozzles, the average life expectancy is about 50,000-100,000 ml. From the above data, we can see that the two printing methods have great economic benefits. Due to the high patented technology fees, ink jet printheads made with piezoelectric technology are relatively expensive, but because the pressure in the micropiezoelectric ink head is small, the nozzles are actually replaced by piezoelectric ceramic wafers. In the micro-piezoelectric ink head, the nozzle and the ink cartridge are of a discrete structure, and one nozzle can support a plurality of ink cartridges. When the ink is replaced, the ink cartridge is replaced without replacing the printhead. Of course, after the piezoelectric inkjet printer's ink head is clogged, it has to waste a lot of ink to clean the ink head. Every time you change the ink head, you must first waste 30% of the ink to do the ink head cleaning. It takes 2 to 3 minutes to clean each time you turn it on, and you need to waste some ink. The gas meter requires the user to replace 17% to 25% of the ink when replacing the ink cartridge. Fig.1 is shown inkjet technology addressing and precise control of the amount of droplets. Logic multiplexed data processing integrates high voltage drive transducer array.

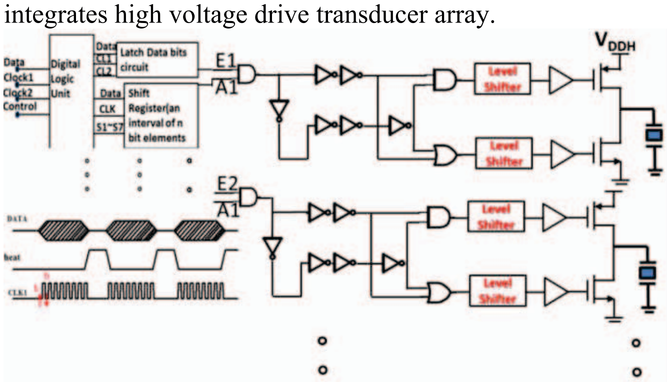


Fig.1 Inkjet technology addressing and precise control of the amount of droplets

The main use of piezoelectric inkjet technology is medical field and industry's inkjet printer. Most of the special industrial printers use ceramic piezoelectric inkjet technology.

RESULTS

It belongs to normal temperature and atmospheric printing technology. This technology places many tiny piezoelectric ceramics around the nozzle of the liquid-jet multilayer structure. The piezoelectric ceramic structure has the characteristics of bending deformation under the voltage change at both ends. When the liquid bead is addressed, the information voltage electrode is supplied to the piezoelectric ceramic. In the structure, the vibration deformation of the telescopic body of the piezoelectric ceramic structure will change with the change of the information voltage of the address data, and the liquid in the ink head can be uniformly and accurately under the normal state of normal temperature and pressure. Spray the liquid beads. Compared with foaming technology, piezoelectric inkjet technology has the rules of ink dot shape, no sputtering, controllable dot size, controllable jet velocity, accurate positioning, more inks with different chemical compositions, less corrosion opportunities, The advantages of the life of the nozzle are prolonged, and pigment ink can also be used to prevent discoloration and fading caused by ultraviolet radiation. Fig.2 is shown the droplets distribution profile.

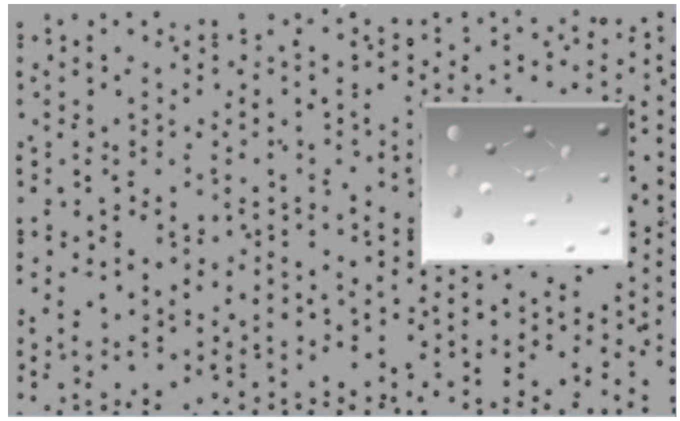


Fig.2 The droplets distribution profile

The liquid specially developed for the piezoelectric technology spray nozzle has the advantages of solubility characteristics, rapid liquid discharge, and fast drying characteristics, and can be quickly placed on the carrier. It is not burr, does not affect the liquid bead of the neighboring neighbors, does not smudge, and the printing quality is very good compared to the hot bubble form.

CONCLUSIONS

Foaming technology and piezoelectric technology have their own advantages and disadvantages. Most of the inkjet devices on the market are based on these two technologies, but to improve the printing effect and printing speed of the inkjet printer, it is impossible to work on the inkjet method alone. Still - make a fuss about the ink drops. The development of technology to improve the quality of color inkjet printing has gone through several stages: the first stage is to pursue resolution; the second stage is to increase the level of color in the image after the resolution reaches a certain level. The number of drops is used to improve print quality; the third stage is to use tiny droplets.

ACKNOWLEDGMENTS

This work was financially supported of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (DP2-108-21121-01-O-05-04).

REFERENCES

[1]Shuichi Yamaguchi ; Ryanto The ; Akira Ueno ; Yoshitake Akiyama ; Keisuke Morishima,"Ejection of a single cell in a single droplet using piezoelectric inkjet head", 2012 International Symposium on Micro-NanoMechatronics and Human Science (MHS),pp.(2012).

[2]Jian-Chiun Liou ; Fan-Gang Tseng,"An Intelligent High-Speed 3D Data Registration Integrated Circuit Applied to Large Array Format Inkjet Printhead",2006 1st IEEE International Conference on Nano/Micro Engineered and Molecular Systems, pp.368-372(2006).

[3]Jian-Chiun Liou ; Ting-Yu Su,"Investigation of the DNA droplet CMOS/MEMS chip microfluidic channels geometry",2018 7th International Symposium on Next Generation Electronics (ISNE),pp.1-2(2018).

A semi-supervised learning model based on convolutional autoencoder and convolutional neural network for image classification

Yu-Xuan Li
school of big data managment
Soochow university
Taipei, Taiwan
ssy1011tw@gmail.com

Hsiang-Yuan Yeh
school of big data managment
Soochow university
Taipei, Taiwan
hyyeh.richard@gmail.com

***Abstract*— Deep learning has achieved the state-of-the-art performance in image classification. But, the model with supervised learning approach should be trained with large parameters and completely labeled datasets. Therefore, we proposed a semi-supervised learning model based on a convolutional auto-encoder and a complementary convolutional neural network to assist image classification. Experimental results show that in the proposed model, the number of labelled data can be reduced by more than half, and the classification accuracy continues to have the same performance. The results show that the effectiveness and feasibility of our model with a limited number of labeled data.**

***Keywords*—convolutional autoencoder, convolutional neural network, semi-supervised, image classification**

I. INTRODUCTION

Recently, deep learning has achieved the state-of-the-art performance in image classification. The model can learn features automatically depend on the training set with superior discriminatory power for image representation while comparing to hand-crafted image descriptors. However, the supervised learning method requires human effort in the selection of millions of the images in the training phase. It also spends a lot of money and time for experts to annotate images. The other concern is that the intra-class variation and inter-class similarity poses even greater challenges for image classification [1]. On the other hand, deep learning with enough hardware and large data would fit models well. For example, training a deep learning model such as VGG-16, using a batch size of 128 will require about 14GB of global memory. Another concept for feature representation is called the unsupervised learning with unlabeled data. It provides a cheaper way to prepare the dataset, and the ability of the unlabeled data holds significant promise in expanding the applicability of learning methods. Therefore, we face several challenges in image classifications as follows: 1) labeled samples with correct annotations are hard to obtain and also very time-consuming. 2) deep learning with multiple layers are hard to turn parameters without a lot of labeled dataset. 3) The parameters turning processes of the deep learning model is time consuming and need enough hardware and memory to deal with it. To address challenges mentioned above, semi-supervised learning methods have been proposed to leverage the abundance of unlabeled data and provide higher generalizability [2, 3]. Semi-supervised learning from both labeled and unlabeled data can thus potentially provide better predictions, compared to supervised learning using only labeled data.

II. METHOD

Among supervised methods, deep learning is less dependent on hand-crafted features and has strong ability of image classification tasks. Unfortunately, data in our real-life typically with few labeled samples and are also highly varying. We study how to train accurate and generalizable with limited labeled data and large scale unlabeled data. We propose a semi-supervised learning based on convolutional autoencoder (CAE) and convolutional neural network (CNN) to assist image classification. In pre-trained unsupervised learning stage, we apply data augmentation to increase both the amount and diversity of data. The common augmentations include translating the image by scaling, or flipping processes. Our proposed approach learns features automatically through pre-trained CAE which reduces the extremely large feature embedding that are extracted from deep layers. After that, the feature embedding extracted by CAE model replace the original image data as the input for CNN, then keep on training image classification model. The workflow of our proposed method is shown in Figure 1:

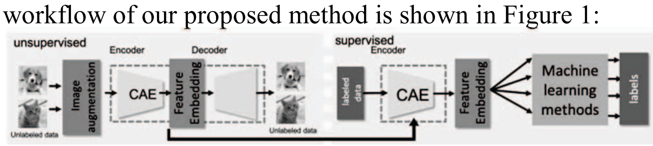


Figure 1: The workflow of our proposed methods

A. Convolutional autoencoder

Traditional autoencoder used fully connected layer as encoder and decoder whereas CAE can capture the 2d structure of the image. CAE models include convolutional, pooling, fully-connected, deconvolution, un-pooling and loss layers. The workflow of the unsupervised feature learning using CAE is shown in Figure 2. The convolution layer uses filter kernels moves from left to right and hops down to next row with same stride until the end of image and extracts the latent feature maps. The pooling layer is reducing the space of the feature maps after convolution layer. In addition, it is usage for extracting superior features, hence keeping the effectively process of training model. In the encoder part, we learn features maps using convolution layer and pooling layer. Deconvolution means the inverse process of convolution and un-pooling is denoted reverse max pooling. The un-pooling layer restores the max-pooled feature into the correct place and fill zero in the other positions. We reconstruct image by deconvolution and up-sampling processes and make it as close as possible to the input data. To optimize CAE, we use *RMSProp* computes a dimension-wise learning rate that adapts the rate of gradients of all previous updates. It is widely used because of its strong theoretical guarantee of convergence and empirical successes.

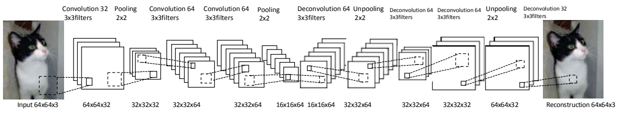


Figure 2: The structure of CAE architecture

B. Convolutional neural network

CNN is a specialized category of neural network for processing image data that has a grid-like structure. A regular CNN architecture contains three main types of layers: convolutional layer, pooling layer, and fully-connected layer in Figure 3. The input of CNN is an image which represents as an array of pixel value. Convolutional layer utilizes the filter kernels to capture the specific structures. The pooling layer effectively reduces the number of parameters and retains the important information. The output layer represents the high-level feature maps of the original image. We then flatten them into the fully connected layer with *softmax* function and use these features to classify the input data into several categories.

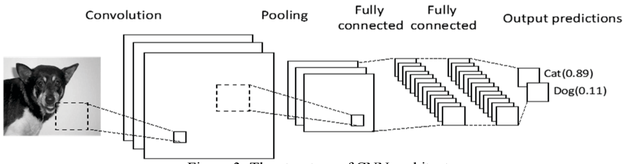


Figure 3: The structure of CNN architecture

III. EXPERIMENTS

We demonstrate the performance of our proposed methods on real world image datasets. We applied different kinds of the feature representation learning algorithms including linear approaches (auto-encoder, principle component analysis (PCA)) and nonlinear approaches (CAE, autoencoder, independent component analysis (ICA)) and several machine learning methods (random forest (RF), logistic regression (LR), CNN) in Figure 4.

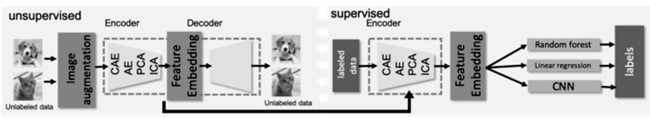


Figure 4: experiments workflow

A. Datasets

We evaluate our methods on image datasets: Dogs vs. Cats from Kaggle (https://www.kaggle.com/c/dogs-vs-cats). It consists of 12,500 cat images and 12,500 dog images. Our problem is a binary classification problem, and we use *crossentropy* as loss function. We split 80% for training and 20% for validation, and apply *adam* optimizer to train the image classification model.

B. The performance of our proposed models

In Figure 5, we use 25,000 images without label information as input dataset to train CAE model and get smaller reconstruction loss (loss=0.03) between original image and reconstructed image. With 1,400 labeled image dataset that we randomly choose from dataset, the accuracy goes 68% to 75%. The results show that the number of labeled data can be reduced from 10,000 to 1,400, and the accuracy is still no less than 75%. It shows that the effectiveness of our model because of the ability to improve performance with a limited number of labeled data supported.

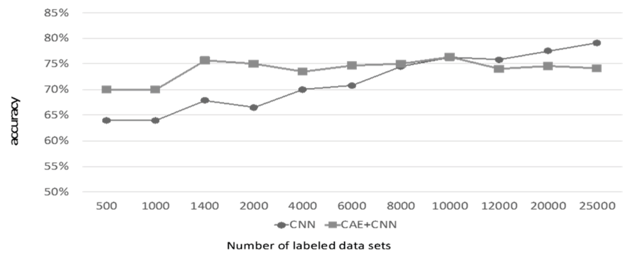


Figure 5: The accuracy of the proposed model

C. The Performances of feature learning

We compress the images to 768 pixels using PCA, ICA and autoencoder and CAE keep two dimensions of the images. In Figure 6, it shows that convolutional layers in CAE model allow better feature representaions in terms of capture the 2d structure of the image well. Based on the accuracy comparison of different feature representation and classifiers, the performance of CAE with CNN is 0.75 which is the highest one and the other methods such as LR (0.71) and RF (0.66) get lower accuracy of the image classification task.

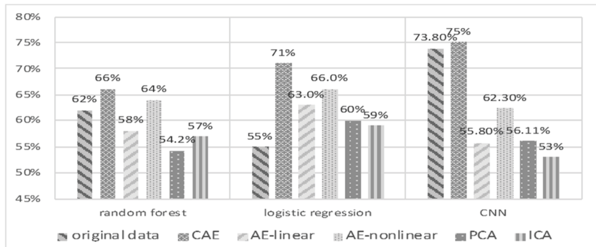


Figure 6: The performance of the feature learning and classifiers

IV. CONCLUSION

Although deep learning-based approaches outperform the state-of-the-art in image analysis tasks, substantial challenges still remain which needs a large amount of labeled data supported. Semi-supervised learning methods are able to learn from limited number of labeled data with the help of a large number of unlabeled data points. We proposed a semi-supervised learning method combining CAE and CNN model. Our model is enable to learn useful representations of image data which can capture the 2d structure of the image. The experiments are shown that the number of labeled data in the training stage is reduced by more than half, all while ensuring a classification accuracy continues to have the same performance.

[1] Y. Song, W. Cai, H. Huang, Y. Zhou, D.D. Feng, W. Yue, M.J. Fulham, and M. Chen, "Large margin local estimate with applications to medical image classification," IEEE transactions on medical imaging, vol. 34, no. 6, pp. 1362-1377, 2015.

[2] R. Marc'Aurelio, F.J. Huang, Y.L. Boureau, and Y. LeCun. "Unsupervised learning of invariant feature hierarchies with applications to object recognition," in Proc. Computer Vision and Pattern Recognition Conference, IEEE Press, 2007.

[3] D. Erhan, Y. Bengio, A. Courville, P.A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" The Journal of Machine Learning Research, vol. 11, pp. 625-660, 2010.

[4] P. Mettes, D.C. Koelma, and C.G. Snoek. "The imagenet shuffle: Reorganized pre-training for video event detection," in Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. 2016.

Reversible Data Embedding For Fingerprint Minutiae Templates Authentication

Yen-Lung Lai, MingJie Lee, Zhe Jin, KokSheik Wong
School of Information Technology
Monash University Malaysia, Jalan Lagoon Selatan,
Bandar Sunway, 47500 Subang Jaya, Selangor
Email: {yenlung.lai, ming.lee, jin.zhe, wong.koksheik}@monash.edu

***Abstract*—In the digital information era, biometric recognition systems have been widely deployed for identity authentication. Among several biometrics, human fingerprint has became one of the most popular biometrics traits. However, the manipulation of the fingerprint data rises the urgent needs of fingerprint data authentication to protect the digital content and ensure authenticity. In this paper, we propose a protection scheme to provide authenticity as well as integrity for individual fingerprint data. A well known reversible data embedding technique named different expansion (DE) is adopted to allow the exact recovery of the hashed fingerprint data with personal secret for authentication. Depending on the size of the data to be embedded, the proposed scheme is able to maintain low distortion over the source fingerprint data, hence maintaining its visual quality and original data structure.**

I. Introduction

Today, biometrics has been widely utilized for personal authentication. In order for the system to function, biometric data of the registered users need to be stored in a database for subsequent authentication. However, the stored data is likely to be compromised and modified by an adversary. Due to the fact that biometric is permanently associated to individual's identity, a compromised database lead to the loss of individual privacy and identity. As a solution for this, biometric protection schemes such as biometric cryptosystem [1], [2], and cancellable biometric [3], [4] are introduced.

Despite the relative high security could be offered by biometric cryptosystem and cancellable biometric, both approaches inevitably destroyed the structure of the original biometric data. Such distortion applied can pique the interest of potential attacker, suggesting the existance of a secret data [5], [6]. In view of this issue, maintaining the quality of the original biometric data, at least visually, is preferable. Biometric data hiding [7] [8], conceals the secret in the biometric template imperceptibly, becomes our resort for such goal.

In this paper, we focus on protecting fingerprint data stored in a database by providing authenticity and integrity check for individual enrolled templates. Our main contribution is to allow exact recovery of the original fingerprint data from its distorted version once the user authenticity is claimed.

A. Fingerprint Minutiae Templates

The fingerprint minutiae templates capture the detailed features of a fingerprint called minutia point. A minutia point can be represented in various way, such as the ridge bifurcation or end. A fingerprint minutiae template can be expressed as a set of n minutia points, $\{m_1, \ldots, m_N\}$, where each point $m_i = (x_i, y_i, \theta_i)$ consists of the x-coordinate, y-coordinate, and its orientation information θ, respectively.

B. Reversible Data Embedding using Difference Expansion

The differences expansion first proposed by Tian [9] is a reversible data embedding technique for data embedding. Given a pair of integer values $\alpha = 206$, $\beta = 201$, denote $d = |\alpha - \beta|$ be the distance between α and β. Suppose we wanted to embed one bit of data as $s \in \{0, 1\}$. Given $s = 1$, then the embedding start from appending s to the binary representation of $d = 206 - 201 = 5 = 101_2$, which becomes $101s_2 = 1011_2 = 11$, refers to the new distance between α and β.

To achieve reversibility, one simply transform α and β to their new coordinate α' and β' follows Eq. 1 and Eq. 2 respectively, yielding $\alpha' = 209$ and $\beta' = 198$.

$$\alpha' = \alpha + \lfloor (d + s + 1)/2 \rfloor \quad (1)$$

$$\beta' = \beta - \lfloor (d + s)/2 \rfloor \quad (2)$$

From the new transformed pair (α', β'), we can extract the embedded bit data $s \in \{0, 1\}$ by first computing $s' = |\alpha' - \beta'| = 11 = 1011_2$, then refer to its least significant bit, which is 1. With that, the original (α, β) pair can be reverted via Eq. 1 and Eq. 2 in its reverse manner.

II. Proposed Method

Our proposed method embeds additional data in each elements of the fingerprint minutiae templates. The embedded information can later be retrieved *exactly* by the provider with a helper data generated during the embedding process. For the purpose of this work, we adopt the DE method due to its reversible property. In particular, we applied DE in an unorthodox way by generating two perturbed fingerprint minutiae templates. One of these perturbed template will be stored and the other serves as a *helper data* to facilitate the extraction of the embedded data for template authentication.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

TABLE I: Example of embedding data with our strategy

s	$M = M'$ $(Original)$	M $(Perturbed)$	M' $(Perturbed)$
3, 2, 3	50, 75, 224	52, 76, 226	49, 74, 223
3, 0, 1	130, 95, 135	132, 95, 136	129, 95, 135
0, 1, 2	88, 101, 149	88, 102, 150	88, 101, 148

A. Data Embedding for Fingerprint Minutiae Templates

Suppose a minutia template $M = \{m_1, m_2, \ldots, m_N\}$ consists of N minutia points, where $m_i = (x_i, y_i, \theta_i)$ is the i-th minutia point. We first denote a string of all zeros 0^n of n bits, viewed as the distance between the minutia points defined as:

$$d(m_i, m_j) = (|x_i - x_j|, |y_i - y_j|, |\theta_i - \theta_j|) \tag{3}$$

We exploit $d(m_i, m_j)$ to embed our data. As we shall see, for $i = j$, then $d(m_i, m_j) = (0^n, 0^n, 0^n)$. The value of n indicates the number of bits to be embedded. Such representation implies that at most $3 \times n$ bits of data can be embedded in each minutia point. To do so, we first generate an exact copy of the original minutia template, denoted as $M' = M$, then apply DE by introducing a pairwise distance $d(m_i, m'_i)$ equal to the three desired values $s_{i,1}, s_{i,2}$, $s_{i,3}$ to be embedded for x, y, and θ, respectively.

For instance, to embed a string $s_i = 001011$, let $n = 2$, then we can view $s_i = (0, 3, 4)$. By Eq. 3, we can express $d(m_i, m_j) = (s_{i,1}, s_{i,2}, s_{i,3}) = (|x_i - x_j|, |y_i - y_j|, |\theta_i - \theta_j|) = (0, 3, 4)$. We let $x_i = \alpha = \beta$ and $d = 0$ to compute (α', β') using Eq. 1 and Eq. 2, then repeat such computation by replacing x_i to y_i and θ_i subsequently, yielding three different perturbed pairs (α', β') as $(\hat{x}_i, \check{x}_i)$, $(\hat{y}_i, \check{y}_i)$, and $(\hat{\theta}_i, \check{\theta}_i)$ at the end (for $i = 1, \ldots, N$). Define $M(Perturbed)$ $=\{(\hat{x}_1, \hat{y}_1, \hat{\theta}_1), \ldots\}$ and $M'(Perturbed)$ $=\{(\check{x}_1, \check{y}_1, \check{\theta}_1), \ldots\}$, Table I depicted an example of proposed embedding strategy over an minutia template M of $N = 3$ minutia points.

III. Discussion

This section discuss the authenticity, template integrity, data embedding capacity and visual distortion of our proposal.

A. Authenticity and Integrity of Minutiae templates

During the embedding process, a cryptographic hash function H can be used to hash a secret w with $M(Perturbed)$. Therefore, the embedding data can be expressed as $s = H(w, M(Perturbed))$. Then, $M(perturbed)$ will be stored instead of $M(Original)$. At the same point, $M'(perturbed)$ serves as the helper data to be kept by the provider. For authentication, suppose the provider processes the secret w' and the helper data, then $H(w, M(Perturbed))$ can be extracted and authentication can be done successfully if $H(w, M(Perturbed))$=$H(w', M(Perturbed))$. In view of this, any provider with the helper data $M'(Perturbed)$ must know the secret w in order to get authenticated. With such argument, authenticity of the minutia template can be claimed.

On the other hand, we provide the second argument to claim the integrity of the minutia templates. Suppose H is a collision resistant cryptographic hash function, where any tampering over the stored minutia templates $M(Perturbed)$ must lead to different hash output. If $H(w, M(Perturbed)) \neq H(w', M(Perturbed))$ (by assuming $w = w'$), then the stored $M(Perturbed)$ must be tempered with, thus offering template integrity check.

B. Embedding Capacity vs Template Distortion

The number of bits to be embedded n comes with a price of greater perturbation applied over the original template $M = M'(Original)$. To embed more data, greater distance has to be introduced and thus leading to greater distortion to the original template M. Given n bits of embedding, it creates a perturbation of at most 2^{n-1} over individual elements of $M(Original)$. Note that such perturbation would increase exponentially with the increment of n, lead to trade-off in between the data embedding capacity and the template distortion. Especially for data hiding and stenography communities, such perturbation over the original template is always not a desired outcome for suppressing the attention to the embedded data as an object of scrutiny. In our case, we suggested to use $n = 1, 2, 3$, or 4 for reasonable embedding capacity and template distortion. With such setting, one can easily embed $N \times 3 \times n$ bits of data while keeping the distortion in between 2 to 8.

IV. Conclusion

In this work, we proposed a reversible data embedding strategy for fingerprint minutia template authentication. The proposed strategy supporting low template distortion and enable the exact recovery of the original minutia template for error-free authentication (given the helper data is presented). We also show that the template authenticity and integrity can be checked by using a collision resistance hash function H with a given secret w.

References

[1] A. B. J. Teoh and J. Kim, "Secure biometric template protection in fuzzy commitment scheme," *IEICE Electronics Express*, vol. 4, no. 23, pp. 724–730, 2007.

[2] U. Uludag, S. Pankanti, and A. K. Jain, "Fuzzy vault for fingerprints," in *International Conference on Audio-and Video-Based Biometric Person Authentication*. Springer, 2005, pp. 310–319.

[3] A. B. Teoh, Y. W. Kuan, and S. Lee, "Cancellable biometrics and annotations on biohash," *Pattern recognition*, vol. 41, no. 6, pp. 2034–2044, 2008.

[4] M. J. Lee, Z. Jin, and A. B. J. Teoh, "One-factor cancellable scheme for fingerprint template protection: Extended feature vector (efv) hashing," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*. IEEE, 2018, pp. 1–7.

[5] N. K. Ratha, M. A. Figueroa-Villanueva, J. H. Connell, and R. M. Bolle, "A secure protocol for data hiding in compressed fingerprint images," in *International Workshop on Biometric Authentication*. Springer, 2004, pp. 205–216.

[6] A. Ross and A. Othman, "Visual cryptography for biometric privacy," *IEEE transactions on information forensics and security*, vol. 6, no. 1, pp. 70–81, 2010.

[7] A. K. Jain and U. Uludag, "Hiding biometric data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 25, no. 11, pp. 1494–1498, 2003.

[8] S. Li and A. C. Kot, "Privacy protection of fingerprint database," *IEEE Signal Processing Letters*, vol. 18, no. 2, pp. 115–118, 2010.

[9] J. Tian, "Reversible data embedding using a difference expansion," *IEEE transactions on circuits and systems for video technology*, vol. 13, no. 8, pp. 890–896, 2003.

Practice of a Two-Stage Model Using Support Vector Regression and Black-Litterman for ETF Portfolio Selection

Jung-Bin Li
Department of Statistics and Information Science
New Taipei City, Taiwan
071635@gapp.fju.edu.tw

Chuan-Yin Chen
Department of Statistics and Information Science
New Taipei City, Taiwan
406145080@mail.fju.edu.tw

***Abstract*—Robo-advisor is a hot topic in the field of financial technology (FinTech) in recent years. This study proposes a programmatic ETF portfolio configuration that combines SVR and Black-Litterman's two-stage model. The results of the study showed that the MSE of the first stage of the model was 2.7970 with good performance. According to the prediction result of the first stage, the parameter setting of the Black-Litterman is performed, and then the model is constructed to further adjust the configuration. The final results show that under the same risk value, the SVM+BL two-stage model proposed by this study has a higher return rate than historical return and implied return. Therefore, it can be provided as a reference for the development of an ETF investment strategy.**

Keywords—support vector regression, Black-Litterman, exchange traded fund, robo-advisor, fund portfolio

I. INTRODUCTION

Until the end of July 2018, there were about 2000 exchange traded funds (ETF) issued and approved for public transaction in Taiwan. It is a difficult job for investors to choose funds that suit their personal needs and goals. This study mainly proposes a two-stage model combining Support Vector Regression (SVR) and Black-Litterman model. Through programmatic calculation and adjustment of various fund investment configurations, our model is able to suggest a portfolio suitable for ordinary investors [1][2].

II. RELATED STUDIES

A. Support Vector Machine / Support Vector Regression

There have been many studies on applying SVM to investment forecasting. Thissena, Brakela, Weijerb, Melssena & Buydens used ARMA, SVM, and Elman recurrent neural networks to compare prediction accuracy [3]. Kim applied SVM to predict stock price index [4]. Lu et.al proposes a two-stage modeling approach, using Independent Components Analysis to generate predictable variables with independent components, and then using these predictive variables with only less noise as variables for SVR modeling [5]. Henrique, Sobreiro & Kimura proposed using the SVR model to predict three different capitalized stock markets [6]. It turns out that the SVR model has a better prediction ability under the periodic update model.

B. Black-Litterman Model

Markowitz proposed the Mean-Variance (MV) Portfolio Model, which laid the foundation for modern portfolio theory [7]. Black & Litterman [8] then developed a new portfolio model based on the Markowitz model.

Silva, Pinheiro & Poggi pointed out that although the BL model improved the shortcomings of the MV model, investors used the subjective view to evaluate the future return rate of the commodity and perform asset allocation [9]. But how to construct these subjective views is confusing. This approach allows for the configuration of individual investment portfolios without the help of an expert. The results of the study show that such an approach can mitigate the impact of no subjective view on future trends.

III. THE PROPOSED MODEL

A. Model Framework

The goal of Robo-Advisor is to use programmatic operations to achieve a reasonable allocation of investment assets. This study integrates the SVR and BL model computation processes to calculate the investment weight. The model framework is shown in Fig.1 below.

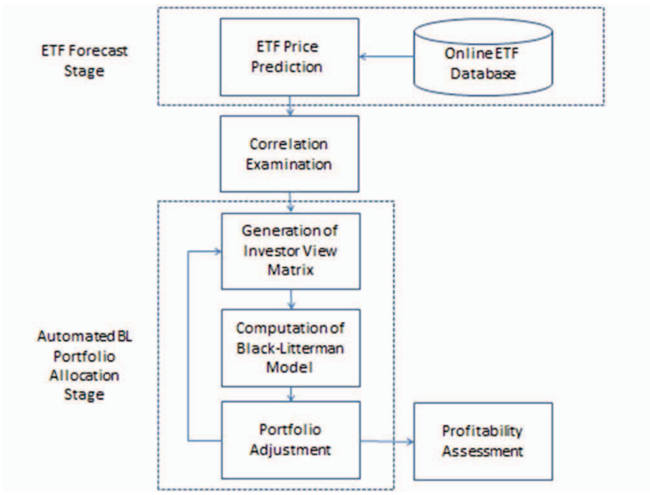


Fig. 1. Model framework.

In the BL Portfolio Allocation Stage, risk aversion coefficient (λ), covariance matrix of historical data (Σ), Market capitalzation weights (w_mkt), implied equilibrium return vector (π), etc., can be estimated from historical data. However, the investor view (Q) is an investor's view of the future return of investment commodities and cannot be estimated using historical data. Therefore, this study first uses SVR to model historical data in the ETF Forecast Stage, predicts the future trend of investment return, and further analyzes the correlation between commodities as input parameters of investors' views.

B. Source of data

The source of this research is the Taiwan Economic Journal (TEJ), and the research objects are the ETF funds issued in Taiwan. We refer to the three major robo-advising websites in the United States, and take 10 funds as our asset allocation portfolio. By removing funds which exists less than one year and the liquidated ones, we choose 10 ETF from the remaining 30. The chosen funds are 0050, 00632R, 00637L, 00679B, 00724B, 07723B, 00663L, 00687B, 00722B, and 00677U. ETF data is collected from February 2017 to March 2019, so there are 520 data entries for each fund. We take 500 entries for the model training and the remaining 20 for testing.

C. Experimental results

In this study, the data from February 2017 to February 2019 was used for training, and the data in March 2019 was for testing. After comparison of parameter combinations, the model is built by cost=1 and epsilon=0.2 to predict price. The average return per day for each ETF fund for March 2019 is shown in Table I. It can be seen from this table that 00632R, 00677U and 00687B have negative returns. Therefore, the weight of each ETF investment is adjusted.

TABLE I. AVERAGE RATE OF RETURN

ETF ID	Average return per day	Average rate of return (ARR)
0050	0.0003	0.0680
00632R	-0.0010	-0.2150
00637L	0.0024	0.8069
00663L	0.0014	0.4270
00677U	-0.0043	-0.6577
00679B	0.0003	0.0756
00687B	-0.0001	-0.0293
00722B	0.0007	0.1908
00723B	0.0006	0.1742
00724B	0.0013	0.3688

To understand the relationship between these three ETF funds with negative returns and other funds, we calculate the correlation coefficient of all chosen funds. It can be found that 00632R has a high negative correlation with 00637L, 00677U has a high negative correlation with 00637L, and 00687B has a negative correlation with 00632R. Based on our observation, three investor views can be summarized:

- V1: 00637L overweight by 0.6577, 00677U underweight by 0.6577
- V2: 00637L overweight by 0.215, 00632R underweight by 0.215
- V3: 00632R overweight by 0.0293, 00687B underweight by 0.0293

Considering of the interaction among investment assets, possible combinations of three investor views are also examined and compared.

Based on our experiments, V1 has the highest ARR among three investor views, and the combination of V1+V3 can further increase the Sharpe Ratio, which means that it is the best asset allocation solution. After re-adjusted the proportion of V1, this study increases the weight from 0.6577 to 1. After such adjustment, the Sharpe Ratio increases significantly from 1.3352 to 1.4312. The efficiency frontier of implied returns with adjusted views is shown in Fig.2.

Taking σ=0.3 as an example, the implied return with adjusted views of the proposed model is around 0.6, which is better than historical returns and implied returns. Hence this two-stage model outperforms the model without view adjustment.

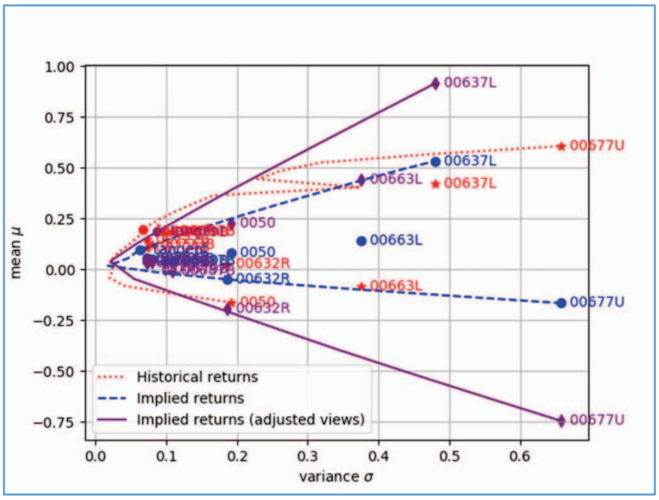


Fig. 2. Efficiency frontiers of ETFs

IV. CONCLUSION

This study attempts to use a fully programmatic approach to combine the SVR prediction with the Black-Litterman investment allocation model to achieve an automated re-allocation of investment assets through a two-stage process. In the first stage of machine learning, the annual return rate derived from the SVR prediction results in this study is within a reasonable range, so the SVR prediction results are used as the source of the second stage model. In the second stage, the correlation coefficient between each investment asset is used as the basis for each asset to be overweighted and underweighted, and the annual return rate of the predicted result is used as the value of the addition and subtraction code. And use this as a parameter for investors in the Black-Litterman model. Through the proposed two-stage model, this study can achieve the goal of better asset allocation by reducing risks and retaining investment return.

REFERENCES

[1] J. Y. Park, J.P. Ryu, & H.J. Shin, "How to Manage Portfolio by Robo-Advisor", Information, vol.20, issue 5B, pp. 3463-3470, 2016.

[2] D. Jung, V. Dorner, C. Weinhardt & H. Pusmaz, "Designing a robo-advisor for risk-averse, low-budget consumers", Electronic Markets, vol.28, issue 3, no.8, pp. 367-380, 2018.

[3] U. Thissena, R. van Brakela, A.P. de Weijer, W.J. Melssen & L.M.C. Buydens, "Using support vector machines for time series prediction", Chemometrics and Intelligent Laboratory Systems, vol. 69, issue 1, pp.35-49, 2003.

[4] K.-j. Kim, "Financial time series forecasting using support vector machines", Neurocomputing, vol.55, issue 1, pp.307-319, 2003.

[5] C.-J. Lu, T.-S. Lee & C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression", Decision Support Systems, vol.47 issue 2, pp. 115-125, 2009.

[6] B.M. Henrique, V.A. Sobreiro & H. Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices", The Journal of Finance and Data Science, vol.4, issue 3, pp.183-201, 2018.

[7] H. Markowitz, "Portfolio Selection", The Journal of Finance, vol.7, issue 1, pp.77-91, 1952.

[8] F. Black, R. Litterman, "Global Portfolio Optimization", Financial Analysts Journal, vol.48, issue 5, pp.28-43, 1992.

[9] T. Silva, P.R. Pinheiro & M. Poggi, "A more human-like portfolio optimization approach", European Journal of Operational Research, vol.256, issue 1, pp.252-260, 2017.

A Continuous Facial Expression Recognition Model based on Deep Learning Method

Szu-Yin Lin
Department of Computer Science and Information Engineering
National Ilan University
Yilan county, Taiwan
szuyinlin@gmail.com

Yi-Wen Tseng
Department of Information Management
Chung Yuan Christian University
Taoyuan County, Taiwan
ian543850204@gmail.com

Chang-Rong Wu
Department of Information Management
Chung Yuan Christian University
Taoyuan County, Taiwan
ujhg21408@gmail.com

Yun-Ching Kung
Department of Information Management
Chung Yuan Christian University
Taoyuan County, Taiwan
lovelycat3312@gmail.com

Yi-Zhen Chen
Department of Information Management
Chung Yuan Christian University
Taoyuan County, Taiwan
yizhen1223@gmail.com

Chao-Ming Wu
Department of Information Management
Chung Yuan Christian University
Taoyuan County, Taiwan
mislighter@gmail.com

Abstract*— Effective facial expression recognition is a very important part of perceiving the user's emotions and designing a successful human-computer interaction system in all fields. Many studies also applied computer vision, machine learning, and deep learning methods on affective computing. However, they mainly focus on the facial expression identification by a single static image, but didn't consider the continuous facial expression of human emotions maybe is for a period of times. Therefore, to recognize the pattern of continuous facial expression and mood changes for further application is still a challenging task. In this study, we propose a continuous facial expression recognition model based on deep learning approach, which combined convolutional neural networks (CNN) and recurrent neural networks (RNN) to analyze and identify continuous facial expressions in a period of time to improve the traditional image recognition method.

***Keywords*—affective computing, facial expression recognition, emotion recognition, deep learning**

I. INTRODUCTION

In recent years, artificial intelligence (AI) has received great attention in the world, especially in the field of computer vision. The goal of computer vision is to make machine use the camera for image recognition and object tracking like human eyes, and let the computer do image analysis and processing tasks. In the past, many convolutional neural networks (CNN) related studies have been proposed in the field of image recognition, such as in face recognition [1] or face detection [2]. Many studies have focused on affective computing, which uses images, sounds, words, or videos as training data to allow computers to recognize human emotions, known as emotion recognition [3]. The emotion recognition is mainly based on the recognition of facial expressions. The facial expressions is based on the FACS (Facial Action Coding System) proposed by P. Ekman and W. Friesen (1978). FACS defines the basic action units (AUS) of human face and then uses these units to define the facial expressions [4].

Although there are many past studies that use CNN in emotion recognition, most of them mainly focus on the analysis and identification of a single static facial expression image. They didn't consider the continuous facial expression of human emotions maybe is for a period of times. Only few image recognition studies are conducted on continuous facial expressions [5]. Therefore, to recognize the pattern of continuous facial expression and mood changes for further application such as service encounter or adaptive learning is still a challenging task. This study establishes a continuous facial expression recognition model for emotion pattern recognition. In addition to using convolutional neural networks (CNN) for image recognition, also combines an improved model of recurrent neural networks (RNN) with time series properties: long short-term memory (LSTM) [6], for the continuous facial expressions over a period of time, makes the predictions more accurate.

II. SYSTEM ARCHITECTURE

The proposed continuous facial expression recognition training model in this study is shown in the Figure 1. It is divided into two phases, which combines CNN, RNN, and facial emotion recognition methods.

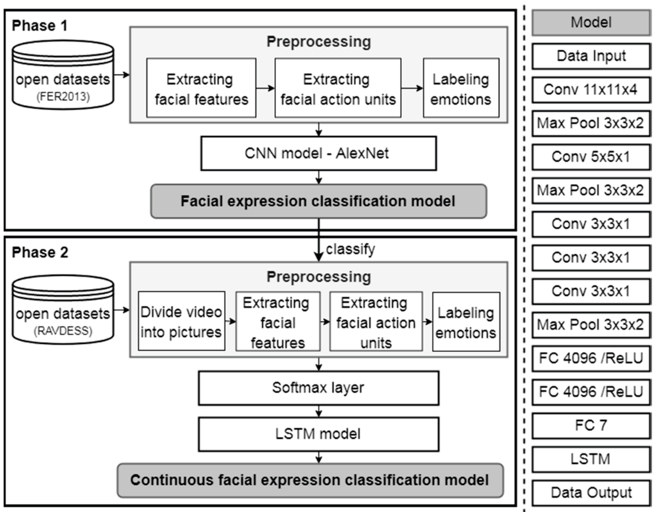


FIGURE 1. CONTINUOUS FACIAL EXPRESSION RECOGNITION MODEL

A. Phase 1

The public emotion dataset FER2013 [7] from Kaggle challenges is used here for training the facial expression classification model in phase 1. The training dataset consists of 28,709 examples. In this phase, the preprocessing steps include (1) extracting facial features, (2) extracting facial action units, and (3) labeling emotions. Firstly, we recognize the region of the human face in the image. Secondly, to extract the action units from the human face, and then finally to define and label emotion tags for each pictures by FACS's definition. After preprocessing, the training data will be put into a CNN AlexNet model [8] for training to a basic facial expression classification model. The parameters of the CNN model refers to the setting of AlexNet model [8]. At the testing stage of the model, this phase will output the vector of each facial expression image in the video sequentially.

B. Phase 2

The public emotion dataset RAVDESS [9] from Zenodo is used here for training the continuous facial expression classification model in phase 2. In this phase, the preprocessing steps include (1) dividing video into images, (2) extracting facial features, (3) extracting facial action units, and (4) labeling emotions. The only different part of preprocessing is the first step. In this step, each emotional video must be first segmented to several image frames according to the setting time interval in the preprocessing. Other steps of preprocessing are the same with the phase 1. After preprocessing, the training data will be put into the first phase's facial expression classification model for generating vectors of all facial expression image frames through the softmax layer. Then, the output vectors will be the input data of LSTM model in the second phase. The target of this phase is to apply LSTM model to discover the pattern of emotional changes within a period of time, and then to achieve the effect of continuous facial expression recognition.

III. Experimental Design

A. Data Collection

Two public datasets for emotion recognition are used in this study for training models. They are FER2013 and RAVDESS.

1) Public Emotions Dataset- FER2013:

There are about 40,000 emotional images in the dataset, including 7 kinds of emotional categories: nature, surprised, fearful, disgust, happy, sad, and angry. The dataset is used to train the first phase's model: facial expression classification model (CNN model).

2) Public Emotions Dataset-RAVDESS:

The ryerson audio-visual database of emotional speech and song (RAVDESS) contains 828 emotional videos. It includes 5 different expressions, such as nature, happy, sad, angry, and fearful. The emotional video is segmented to every second, and then the system saves every frame as an image. Because each video is amount five seconds long, there are total 4,140 emotional images.

B. Data Preprocessing

The facial features and action units of the images are extracted, and save them as the preprocessed files. With the help of data preprocessing, the recognition result of the facial expression recognition model can be more accurate.

C. Experimental Results

We compare the proposed model with the traditional CNN model in 5 emotion classifications. The 5 emotion categories are natural, happy, sad, angry, and fearful. The proposed model's accuracy can reach to 74.07%, which is higher than just use CNNs' 43.17% (Table I).

TABLE I. METHOD COMPARISON

Condition	Method	Accuracy
With five emotion classifications	Facial expression classification model	43.17%
	Continuous facial expression classification model	**74.07%**

In addition, we use 108 videos in RAVDESS's emotional videos as the testing data for the continuous facial expression classification model, and use the confusion matrix to present its results (Table II).

TABLE II. CONTINUOUS FACIAL EXPRESSION CLASSIFICATION MODEL CONFUSION MATRIX

Confusion matrix		Actual				
		natural	happy	sad	angry	fearful
Predicted	natural	**10**	2	3	4	4
	happy	2	**20**	0	0	0
	sad	0	1	**17**	1	1
	angry	0	1	1	**15**	1
	fearful	0	0	3	4	**18**

In the end, all 5 kinds of emotion classifications are mostly accurate, which proves that this continuous facial expression recognition model has its accuracy for various emotion classifications. However, as we know in the past experiences, the natural expression is the most difficult recognition target, because it's neutrality and non-emotional. This model has the accuracy of up to 82.4% without considering the natural category.

IV. Conclusion

This study proposed a continuous facial expression recognition model based on deep learning approach, which is integrated with convolutional neural network and recurrent neural network for emotion recognition. To compare the accuracy with the past CNN method. The proposed model gets more accurate results from the experiment. Because this proposed model analyzes and identifies the pattern of continuous facial expressions and mood changes in a period of time to improve the traditional image recognition method.

Acknowledgment

This research work was supported by the Ministry of Science and Technology, Taiwan under the grant 108-2410-H-197 -004 and 107-2410-H-197 -004.

References

[1] S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.

[2] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.

[3] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98–125, Sep. 2017.

[4] P. Ekman, W. V. Freisen, and S. Ancoli, "Facial signs of emotional experience.," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1125–1134, 1980.

[5] S. E. Kahou *et al.*, "EmoNets: Multimodal deep learning approaches for emotion recognition in video," *J Multimodal User Interfaces*, vol. 10, no. 2, pp. 99–111, Jun. 2016.

[6] H. Sak, A. Senior, and F. Beaufays, "Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition," *arXiv:1402.1128 [cs, stat]*, Feb. 2014.

[7] "fer2013." [Online]. Available:https://kaggle.com/deadskull7/fer2013. [Accessed: 13-Jul-2019].

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, USA, 2012, pp. 1097–1105.

[9] "RAVDESS Emotional speech audio." [Online]. Available: https://kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio. [Accessed: 13-Jul-2019].

Mixing Binary Face and Fingerprint based on Extended Feature Vector (EFV) Hashing

Ming Jie Lee, Zhe Jin
School of Information Technology
Monash University Malaysia
{ming.lee, jin.zhe}@monash.edu

Minyi Li
School of Science
RMIT University, Australia
minyi.li@rmit.edu.au

Daniel Bo-Wei Chen
Department of Electrical Engineering
National Sun Yat-Sen University
bo-wei.chen@mail.nsysu.edu.tw

***Abstract*—Multimodal biometric template protection (BTP) is gaining increasing attention as it overcomes the drawback of unimodal BTP. In this work, a token-less cancellable biometric scheme, namely Extended Feature Vector (EFV) Hashing, is applied to fuse the multimodal binary face and fingerprint template at a feature-level fusion. In advance to the unimodal EFV hashing, this work reorganizes the feature extension mechanism in EFV transformation and introduces the random sampling mechanism to increase the difficulty for reverse processing of the cancellable template. In addition, two fusion options are developed for producing the cancellable template. Experiments are conducted on the well-known Fingerprint FVC2004 and Face LFW datasets. The results demonstrate that the propose approach achieves a satisfactory verification rate with EER ≈ 0.1%.**

***Keywords*—cancellable biometrics, feature-level fusion, multi-biometrics**

I. Introduction

Biometric identity management (IdM) system refers to the technology for managing and authenticating one's identity based on one's physiological or biological traits, such as, fingerprint. In our modern life, biometric system has become an indispensable part of our life, for instance, biometric authentication in smart devices. However, this leads to significant increase of pre-stored biometric templates, which implies a higher chance that the biometric templates are stolen for harmful activities, e.g., privacy invasion and identity theft [1]. In this unfortunate event, the user cannot reuse the same biometric traits due to the irrevocable traits of biometrics [1].

Biometric template protection (BTP) is the crux in tackling this issue. BTP is commonly categorised as biometric cryptosystems (BCS) and cancellable biometrics (CB) [2]. In this work, we focus on CB (readers could refer to [2] for more details on BCS). Until now, most of existing CB schemes, e.g., [3], [4] are designed as a two-factor authentication scheme as it requires a user to input a *token* (an external factor for storing transformation key) and along with his/her *biometric feature(s)* for authentication. It is impractical when the user has to manage the token especially the token can be easily stolen or forgotten [5]. Moreover, once the key is compromised, the adversary can utilize the stolen-key for launching a zero-effort false acceptance attack to gain illegal access [6]. In addition, the user has to re-issue the cancellable template whenever the token is accidentally lost or forgotten. Thus, token-less (or one-factor) approach, e.g., EFV hashing [7] or one-factor IFO [8] is more desirable for a BTP scheme.

Lately, multimodal biometrics is gaining increasing interests from the research community as it overcomes the disadvantages of the unimodal biometrics, e.g., performance degradation when the BTP is applied [6], [9]. In addition, feature-level fusion-based CB scheme won popularity because of the small template storage size and high security compared to score-level and decision-level approaches [3], [4]. In this work, we explore feature-level fusion and apply the EFV hashing [7] on a multimodal face and fingerprint based biometric system. The contributions of this work are briefly highlighted as: 1) Application of EFV hashing [7] onto a multimodal biometric system to realize *token-less* and *multimodal* biometric template protection, and 2) Inspired from [3], two feature-level fusion options are developed to transform the *binary* face and fingerprint templates into a cancellable template.

II. Methodology

A. Random sampling enabled EFV hashing

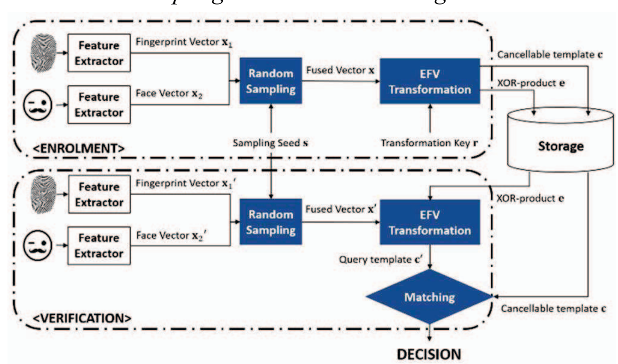


Fig. 1. System diagram of proposed framework (fusion option-1)

EFV hashing [7] is a unimodal cancellable biometric scheme which was originally proposed for fingerprint biometrics. In this work, the plain feature extension of the EFV hashing is replaced by a random sampling process (refer to step-1 below) to increase the randomness of the input concatenated face and fingerprint vectors. Given a biometric input $\mathbf{x} \in [0,1]^p$, the process to generate a cancellable template $\mathbf{c} \in [0,1]^{pn}$ is outlined as below:

1) Given a random sampling seed $\mathbf{s} \in [1, d]^{pn}$, permute and extend $\mathbf{x}$ based on $\mathbf{s}$ to yield the sampled biometric vector $\bar{\mathbf{x}} \in [0,1]^{pn}$.

2) For each $\bar{x}_i \in \bar{\mathbf{x}}$, form a *sub-bits block* by appending $\bar{x}_i$ with its following elements in $\bar{\mathbf{x}}$ until the sub-bits block satisfying k-bits. This process generates pn numbers of sub-bit blocks.

3) Convert each sub-bit block to an integer value $\hat{x}_i$ where i is the number of sub-bit block. The $\hat{x}_i$ is then transformed in accordance to $\hat{x}_i = \left((\hat{x}_i \times i) \mathrm{mod}(pn+1)\right) + 1$. Repeat this process until pn sub-bit blocks are transformed. After that, an integer biometric vector $\hat{\mathbf{x}} \in [1, pn]^{pn}$ is formed.

4) Finally, permute the transformation key $\mathbf{r} \in [0,1]^{pn}$ according to $\hat{\mathbf{x}} \in [1, pn]^{pn}$ to yield the cancellable template $\mathbf{c} \in [0,1]^{pn}$ where $c = r_{\hat{x}_i}$.

During enrolment stage, the transformation key $\mathbf{r} \in [0,1]^{pn}$ is generated via a pseudo-random number generator. After $\mathbf{c}$ is generated, $\mathbf{r}$ will be transformed to an XOR product $\mathbf{e} \in [0,1]^{pn}$ where $\mathbf{e} = \mathbf{r} \oplus \mathbf{x}$. The $\bar{\mathbf{x}}$ and $\mathbf{r}$ are then be deleted. During verification stage, the $\mathbf{r}'$ is generated as $\mathbf{r}' = \mathbf{e} \oplus \bar{\mathbf{x}}'$ where $\bar{\mathbf{x}}'$ is transformed from the query biometric input $\mathbf{x}'$. On the other hand, since $\mathbf{s}$ is merely a sampling seed to extend the input $\mathbf{x}$ to $\bar{\mathbf{x}}$, $\mathbf{s}$ can be stored in database publicly.

B. *Two feature-level fusion options*

Given a face vector $\mathbf{x}_1 \in [0,1]^a$ and a fingerprint vector $\mathbf{x}_2 \in [0,1]^b$, the following two feature-level fusion options:

1) $\mathbf{x}_1$ and $\mathbf{x}_2$ are concatenated as $\mathbf{x} = \mathbf{x}_1 | \mathbf{x}_2$. Then, $\mathbf{x}$ is transformed to a cancellable template $\mathbf{c} \in [0,1]^{(a+b)n}$.

2) $\mathbf{x}_1$ and $\mathbf{x}_2$ are transformed to $\mathbf{c}_1$ and $\mathbf{c}_2$ separately. After that, we concatenate $\mathbf{c}_1$ and $\mathbf{c}_2$ to yield the cancellable template $\mathbf{c} \in [0,1]^{an+bn}$.

III. Experiment and Discussions

We conduct experiments on the two commonly used datasets, i.e., fingerprint FVC2004 and face LFW dataset to evaluate the performance of the proposed approach in terms of *verification rate*. In this experiment, fingerprint vectors are extracted via a learning-based method reported in [10], where the procedures include: i) MCC template extraction, ii) KPCA transformation and iii) Binarization. Besides that, face vectors are extracted via the FaceNet [11] with an additional binarization process. This face vector extraction involves: i) Face image alignment and cropping, ii) Real-valued vector extraction and iii) Mean rule binarization. Both fingerprint and face templates are extracted as binary vector with length of 256-bits. Reader may refer to [10], [11] for the details.

A. *Verification rate*

TABLE I. Verification rate of binary unprotected system

FVC2004 DB1	FVC2004 DB2	FVC2004 DB3	LFW
1.58	4.39	1.74	1.01

TABLE II. Verification rate of protected system

Dataset	Fusion Option	EER (%)		
		Unimodal Protected		*Multimodal Protected*
		Face	*fingerprint*	
FVC2004 DB1 + LFW	1	1.02 ± 0.002	1.62 ± 0.003	0.13 ± 0.001
	2			0.17 ± 0.001
FVC2004 DB2 + LFW	1		5.00 ± 0.001	0.09 ± 0.01
	2			0.10 ± 0.00
FVC2004 DB3 + LFW	1		1.84 ± 0.003	0.18 ± 0.001
	2			0.07 ± 0.001

We follow the full FVC protocol presented in [12] to examine the verification rate with different fusion options, and in unimodal and multimodal protection, respectively. In each experiment, we generate 1000 *genuine* and 4950 *impostor* matching scores for computing the Equal Error Rate (EER). 5 experiments (with 5 set of $\mathbf{s}$ and $\mathbf{r}$) are conducted on each dataset and the mean EER is calculated. Towards cancellable template generation, we follow the best tuned parameter $\{n = 1000, k = 3\}$ reported in [7]. Note that the similarity score S between the pre-stored and query templates is calculated as $S = (1 - \text{Hamming Distance})$.

From Table II, we observe that the proposed scheme can works well in unimodal face or fingerprint system. On the other hand, the performance degradation issue is further resolved when the proposed scheme is operated in multimodal mode, where the EER is significantly lower than that in unimodal mode. For instance, on average, it achieves low EER with 0.13% (option-1), 0.11% (option-2) in multimodal mode. On the other hand, with unimodal fingerprint template (resp. face template), EER is always above 1%.

IV. Conclusion and Future work

In this work, we develop a token-less cancellable multimodal face and fingerprint based biometric system. We introduce a random sampling mechanism in EFV transformation to do feature-level fusion and develop two fusion options to handle multimodal face and fingerprint biometric features. The proposed scheme can be generalised to other biometric modalities with binary representation as well as unimodal/ multimodal mode. For future work, we will extend the current framework to fuse the real-valued biometric inputs, where different biometric modalities could hold different value ranges.

Acknowledgement

This work is partly supported by the Ministry of Science and Technology, Taiwan (107-2218-E-110-013-MY3), NVIDIA Titan Xp GPU grant donation, FRGS (FRGS/1/2018/ICT02/MUSM/03/3).

References

[1] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric Template Security," *EURASIP J. Adv. Signal Process*, vol. 2008, pp. 113:1--113:17, Jan. 2008.

[2] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP J. Inf. Secur.*, vol. 2011, no. 1, p. 3, Sep. 2011.

[3] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "A fingerprint and finger-vein based cancelable multi-biometric system," *Pattern Recognit.*, vol. 78, pp. 242–251, 2018.

[4] M. Gomez-Barrero, C. Rathgeb, G. Li, R. Ramachandra, J. Galbally, and C. Busch, "Multi-biometric template protection based on bloom filters," *Inf. Fusion*, vol. 42, pp. 37–50, 2018.

[5] O. Ouda, N. Tsumura, and T. Nakaguchi, "Tokenless Cancelable Biometrics Scheme for Protecting Iris Codes," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 882–885.

[6] K. Nandakumar and A. K. Jain, "Biometric Template Protection: Bridging the performance gap between theory and practice," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 88–100, Sep. 2015.

[7] M. J. Lee, Z. Jin, and A. B. J. Teoh, "One-factor Cancellable Scheme for Fingerprint Template Protection: Extended Feature Vector (EFV) Hashing," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2018, pp. 1–7.

[8] J. Kim and A. B. J. Teoh, "One-factor Cancellable Biometrics based on Indexing-First-Order Hashing for Fingerprint Authentication," in *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018, pp. 3108–3113.

[9] A. K. Jain, A. A. Ross, and K. Nandakumar, "Multibiometrics," in *Introduction to Biometrics*, Boston, MA: Springer US, 2011, pp. 209–258.

[10] Z. Jin, M. H. Lim, A. B. J. Teoh, B. M. Goi, and Y. H. Tay, "Generating Fixed-Length Representation From Minutiae Using Kernel Methods for Fingerprint Authentication," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 46, no. 10, pp. 1415–1428, Oct. 2016.

[11] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: {A} Unified Embedding for Face Recognition and Clustering," *CoRR*, vol. abs/1503.0, 2015.

[12] R. Cappelli, D. Maio, D. Maltoni, J. L. Wayman, and A. K. Jain, "Performance evaluation of fingerprint verification systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 3–18, Jan. 2006.

Analysis of chroma pixel value prediction using luma pixel values

[1] Yan-Jhu Wang , [1] You-Sheng Guo[1]* , [1]Changsheng Deng, [2]Ting-Lan Lin

[1]*Department of Electronic Engineering, Chung Yuan Christian University, Chungli, Taiwan (ROC)*
[2]*Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan (ROC)*

***Abstract*—In this paper, we propose a new mathematical method that can improve the accuracy of reconstruction luma to be able to choose the correct method of restoration for screen content images (SCIs). After receiving the decoded luma image and subsampled chroma image from the decoder, SSE (the sum of squares due to error) minimization is proposed to identify the used chroma subsampling scheme before compression. By using our method, subsampling reconstructed can be predicted more precisely through our way.**

***Keywords*— screen content images (SCIs), reconstruction luma, the sum of squares due to error (SSE).**

I. INTRODUCTION

Screen content images (SCIs) are computer generated images for screen display and one SCI usually contains a few different objects, such as symbols, characters, graphics, and/or natural subimage(s). Figure 1 in the traditional compressed system, to begin with, the RGB image is transformed into the YUV image. Then the chroma UV image is subsampled by the adopted subsampling scheme. Furthermore, the image is encoded. When the encoding values are received by the decoder, we can inverse the steps above. Finally, we obtain the reconstructed compressed image.

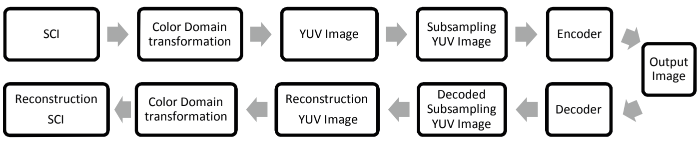


Figure 1. traditional compressed system

We use five existing chroma subsampling schemes, **4:2:0(A), 4:2:0(L), 4:2:0(R), 4:2:0(DIRECT), 4:2:0(MPEG-B)**. The 4:2:0(A) determines averaging all the U and V components of the 2×2 UV block. The 4:2:0(L) and 4:2:0(R) determine by averaging the U and V components in the left and right columns of the 2x2 UV block, respectively. The 4:2:0(DIRECT) selects the upper-left U and V components of the block. The 4:2:0(MPEG-B) performs the 13-tap filter with mask [2, 0, −4, −3, 5, 19, 26, 19, 5, −3, −4, 0, 2]/64 on the upperleft U and V components of the block.

For reconstructing the chroma image, we have to know the subsampling method to optimize the restore image. In the existing paper, they used the linear regression to solve coefficients a and b, which are the slope and intercept of the fitting line $U'' = aY + b$, comparing with subsampling U’ and predicting U’’ to predict the subsampling method for reconstruction of the chroma image.

Using linear regression is not the best choice to reconstruction Luma. Therefore, we proposed the new method “quadratic method” $U'' = aY^2 + bY + c$. It can increase the accuracy of method which subsampling used.

II. THE SUM OF SQUARES DUE TO ERROR

In this subsection, we want to improve the accuracy of reconstructed luma to be able to choose the correct method of restoration. When receiving the decoded luma image Y and the subsampling chroma image $U'^{(cs)}$ from the decoder, our goal is to efficiently identify the used subsampling scheme CS. We have 5×5 subsampling luma blocks by running the five subsampling schemes. Since the proposed identification strategy is a block-based approach, the configuration of the subsampled pixels in one 5×5 subsampled luma and chroma block-pair have nine subsampled pixel-pairs.

We take Figure 2 as a test example. This image has size 682×672. We use 4:2:0(L) to subsample for example. In the existing paper, the correlation for the block-pair$(U'^{(L)}, Y'^{(Is)})$ can be determined by solving the following overdetermined system:

$$\begin{bmatrix} Y'_1 & 1 \\ Y'_2 & 1 \\ \vdots & \vdots \\ Y'_9 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} U'_1 \\ U'_2 \\ \vdots \\ U'_9 \end{bmatrix} \tag{1}$$

The correlation parameter-pair (a, b) can be solved by linear regression. Compare the subsampling U' and the U'' $(U'' = aY' + b)$, we predict the subsampling method.

We proposed two methods. In the first one, the block-pair$(U'^{(L)}, Y'^{(Is)})$ is determined by exponential function. It can be solved by following overdetermined system:

$$a \times \exp\left(b \times \begin{bmatrix} Y'_1 \\ Y'_2 \\ \vdots \\ Y'_9 \end{bmatrix} \right) = \begin{bmatrix} U'_1 \\ U'_2 \\ \vdots \\ U'_9 \end{bmatrix} \tag{2}$$

*Corresponding author. E-mail address: s10526306@cycu.edu.tw

Figure 2. Test SCI

The (a, b) can be solved by matlab function polyfit and exponential. Comparing the subsampling U' and the U'' ($\mathrm{U}'' = \mathrm{a} \times e^{bY'}$) to determine the subsampling method.

In the second method, the block-pair($U'^{(L)}, Y'^{(Is)}$) is determined by quadratic equation in the following overdetermined system:

$$\begin{bmatrix} (\mathrm{Y}'_1)^2 & \mathrm{Y}'_1 & 1 \\ (\mathrm{Y}'_2)^2 & \mathrm{Y}'_2 & 1 \\ \vdots & \vdots & \vdots \\ (\mathrm{Y}'_9)^2 & \mathrm{Y}'_9 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \end{bmatrix} = \begin{bmatrix} \mathrm{U}'_1 \\ \mathrm{U}'_2 \\ \vdots \\ \mathrm{U}'_9 \end{bmatrix} \tag{3}$$

The (a, b, c) can be solve by matlab function polyfit. Comparing the subsampling U' and the U'' ($\mathrm{U}'' = \mathrm{aY}'^2 + \mathrm{bY}' + \mathrm{c}$), we the subsampling method.

To acquire minimum error between original image and reconstructed image, we apply SSE to compare with differences between U' and U'' (V' and V'') in the five subsampling methods. The best reconstructed method is the one with the smallest SSE. The mathematical for SSE is:

$$\begin{aligned} \mathrm{SSE} &= (\mathrm{U}'' - \mathrm{U}')^2 \\ &= (\mathrm{aY}^2 + \mathrm{bY} + \mathrm{c} - \mathrm{U}')^2 \end{aligned} \tag{3}$$

We compute the fitting line for every 5×5 image size. It takes a total of 452903 times. For every 5×5 block we can find a predicted method that is closest to the original chroma image with smallest SSE value. We compute the accuracy of the methods predicting the original methods. For example for U pixels, when the subsampling method is All, the linear method can achieve 39.5%, exponential can achieve 58.2% and quadratic can achieve 53.8%; this shows that the proposed methods can increase the hit rate of the true subsampling method. The results are shown in table 1, along with the test results of other subsampling methods. The similar test is done for U pixels in table 2. Both tables show that the proposed methods can improve the hit rate in almost all subsampling methods other than MPEG-B, and this is the subject of our future work.

*Corresponding author. E-mail address: s10526306@cycu.edu.tw

Table 1. Predicting the upsampling method by the percentage (U layer)

subsampling method	linear	exponential	quadratic
All	0.395	0.582	0.538
Left	0.015	0.088	0.127
Right	0.009	0.122	0.141
DIRECT	0.021	0.085	0.107
MPEG-B	0.541	0.122	0.138

Table 2. Predicting the upsampling method by the percentage (V layer)

subsampling method	linear	exponential	quadratic
All	0.393	0.561	0.539
Left	0.014	0.083	0.127
Right	0.008	0.157	0.132
DIRECT	0.021	0.088	0.108
MPEG-B	0.541	0.105	0.132

III. Conclusion

In this paper, we improve the accuracy of reconstructed luma to be able to be chosen for the correct method of restoration. The results show that the quadratic method can improve the accuracy compared with existing methods.

References

[1]. S. Wang, K. Gu, S. Ma, and W. Gao, “Joint chroma downsampling and upsampling for screen content image,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 9, pp. 1595–1609, Sep. 2016.

[2]. Y. Zhang, D. Zhao, J. Zhang, R. Xiong, and W. Gao, “Interpolation dependent image downsampling,” IEEE Trans. Image Process. vol. 20, no. 11, pp. 3291–3296, Nov. 2011.

[3]. S. Minaee and Y. Wang, “Screen content image segmentation using least absolute deviation fitting,” in Proc. IEEE Int. Conf. Image Process., Sep. 2015.

[4]. J. Korhonen, “Improving image fidelity by luma-assisted chroma subsampling,” in Proc. IEEE Int. Conf. Multimedia Expo, pp. 1-6, Aug. 2015.

[5]. K.- L. Chung, C.-C. Huang, and T.-C. Hsu, “Adaptive chroma subsampling-binding and luma-guided chroma reconstruction method for screen content images,” IEEE Transactions on Image Processing. vol. 26, no.12, pp. 6034–6045, Dec. 2017.

Mapping and Permutation Set Design for Spatial Permutation Modulation (SPM)

Che-Wei Lee, Hsu-Hsuan Tu, I-Wei Lai
Department of Electrical Engineering, National Taiwan Normal University, Taipei, Taiwan
E-mail: 60775015h@ntnu.edu.tw

***Abstract*—Spatial permutation modulation (SPM) has been proposed and demonstrated as an efficiency multiple-input multiple-output (MIMO) technique which generalizes spatial modulation (SM) by exploiting the time coordinate for further transmit diversity gain. In particular, SPM activates the transmit antennas at successive time instants by the permutation vector which brings the information bits. In this work, we elaborate on the design of the permutation set of SPM and the mapping between the permutation vector and the information bits. A MIMO system with 6 transmit antennas and 16 permutation vectors in the adopted set is applied as an example for these two design issues. Numerical results demonstrate that, by using our mapping rules and designed permutation sets, the error rate of the SPM system is greatly improved.**

***Index Terms*—Multiple-input multiple-output (MIMO), spatial modulation (SM), spatial permutation modulation (SPM), permutation, mapping.**

I. Introduction

Multiple-input multiple-output (MIMO) technology has been applied in numerous wireless standards. The technique called spatial modulation (SM) [1] is considered one of the promising MIMO technique, which actives one transmit antenna at a time according to the informaion data. The resulting *spatial symbol* is used to bring the information, together with the transmitted QAM symbols.

Based on SM, the spatial permutation modulation (SPM) has been proposed [2] by exploiting the time coordinate with the aid of the permutation vector. In this way, the transmit diversity is improved and more flexible trade-off between the reliability and rate is achieved. Specifically, a permutation vector is selected from the set according to the information data. The transmit antennas are then successive activated with the indices equal to the entries of the selected permutation vector. As can be seen, the design of the permutation set for SPM and the mapping between the permutation vector to the information bits are essential for SPM. In this work, we exemplarily explain these two issues by considering a MIMO system with 6 transmit antennas and 4 bits of the spaital symbol, i.e., 16 permutation vectors of the adopted set. Note that this permutation can also be utilized to the system with 4 transmit antennas with 2 activate transmit antennas simultaneously.

The rest paragraphs of this paper are as follows: Section II revisits the SPM-MIMO system. Section III introduces the mapping and permutation set design. Section IV shows the bit error rate (BER) performance of the proposed SPM. The conclusions are drawn in Section V.

II. System Model

Define $\tilde{C}_{N_t,T}$ as the set of different ordered arrangements of a T-element subset of an N_t-set. For example,

$$\tilde{C}_{3,3} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\}. \tag{1}$$

We depict the permutation set with defining Hamming distance matrix $\mathbf{D}$ whose(i,j)th entry $d_{i,j}$ represents the Hamming distance between the ith and jth permutation vectors. For example, the set $\{[1,3,2]^\top,[2,1,3]^\top,[2,3,1]^\top,[3,1,2]^\top\} \subset \tilde{C}_{3,3}$ its hamming distance matrix is

$$\mathbf{D} = \begin{bmatrix} 0 & 3 & 2 & 2 \\ 3 & 0 & 2 & 2 \\ 2 & 2 & 0 & 3 \\ 2 & 2 & 3 & 0 \end{bmatrix}. \tag{2}$$

The minimum Hamming distance of a permutation set i.e., smallest off-diagonal component in the matrix $\mathbf{D}$ can be described by

$$d_{\min} = \min_{i,j \ \ i\neq j} d_{i,j}. \tag{3}$$

We denote the subset $C_{N_t,T}(K,d_{\min}) \subseteq \tilde{C}_{N_t,T}$ which selects K permutation vectors with the minimum Hamming distance d_{min} from $\tilde{C}_{N_t,T}$. The SPM transmitter modulates $\lfloor \log_2 K \rfloor$ bits by using $C_{N_t,T}(K,d)$. The error rate performance and throughput of SPM can be exchanged by adjusting K and $d_{\min}$ of $C_{N_t,T}(K,d_{\min})$. For the SPM transmission, SPM modulates partial data bits to a permutation $\mathbf{p} = [p_1,...,p_T]^\top \in C_{N_t,T}(K,d)$ in order that the p_tth transmit antenna is activated at the t_{th} time instant. The received signal vector at the t_{th} time instant is given by

$$\mathbf{y}_t = \mathbf{h}_{p_t} s_{\lceil \frac{t}{M} \rceil} + \mathbf{n}_t, \tag{4}$$

where $\mathbf{h}_{pt}$ is the p_tth column of the channel matrix $\mathbf{H}$, and s is the QAM symbol, which is selected from the QAM constellation set χ. The parameter M represents the number of repetitive transmission of the QAM symbol, and $\lceil \cdot \rceil$ is the ceiling operation. That is, we have the QAM vector

$$\mathbf{s} = [\underbrace{s_1,\ldots,s_1}_{M}, \underbrace{s_2,\ldots,s_2}_{M}, \ldots, \underbrace{s_{\frac{T}{M}},\ldots,s_{\frac{T}{M}}}_{M}]^\top, \tag{5}$$

which is transmited at one SPM transmission. By cascading $\mathbf{y}_t$ at all time instants, i.e., $t = 1,...,T$, into a received matrix $\mathbf{Y} = [\mathbf{y}_1,\ldots,\mathbf{y}_T] \in \mathbb{C}^{N_t \times T}$, we get

$$\mathbf{Y} = \mathbf{H}(\mathbf{p})\text{diag}(\mathbf{s}) + \mathbf{V}, \tag{6}$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

where $\mathbf{H}(\mathbf{p})$ is the permuted channel matrix, and diag($\mathbf{s}$) is a diagonal matrix with diagonal entries being $[s_1, \ldots, s_{\frac{T}{M}}]$ and V = $[\mathbf{n}_1 \ \ldots, \ \mathbf{n}_T]$ is the noise matrix.

III. Mapping and Permutation Set Design

SPM is sensitive to the design of the permutation set for SPM and the mapping between the permutation vector to the information bits. For design of the permutation set in a general situation, we consider about spatially-correlated channels, the spatial correlation matrix model is adopted [3] , [4]

$$\mathbf{H}_{\text{sc}} = \mathbf{R}_{\text{T}}^{1/2}\mathbf{H}\mathbf{R}_{\text{R}}^{1/2}, \tag{7}$$

where $\mathbf{R}_{\text{T}}^{1/2} \in \mathbb{C}^{N_t \times N_t}$ and $\mathbf{R}_{\text{R}}^{1/2} \in \mathbb{C}^{N_r \times N_r}$ with $R_{\text{T}_{i,j}} = R_{\text{R}_{i,j}} = \rho^{-|i-j|}$ correlation matrices which used to model the correlation at the both transmitter and the receiver sides.

In the part of design of the permutation set, we consider the $d_{\min}$ of permutation set and spatial correlation which can easily affect the performance. Take $(N_t, T, K) = (4, 2, 4)$ as an example, frist we select the permutation sets which have maximum $d_{\min}$ as candidates. Then, due to the spatial correlations, the neighbor antennas have similar fading gains and thus should be avoided even they are different antennas. For example, the permutation pair between $[1, 2]^\top$ and $[1, 3]^\top$ is more likely to be erroneously detected than the pair between $[1, 2]^\top$ and $[1, 4]^\top$. As a result, we get $\{[1, 3]^\top, [2, 4]^\top, [3, 1]^\top, [4, 2]^\top\}$ which has maximum $d_{\min} = 2$ and the lowest spatial correlation.

SPM can flexibly adjust the permutation set $C_{N_t,T}(K, d_{\min})$ to achieve the desired balance between the error rate performance and throughput. In order to increase throughput, the transmitter equips more antennas to increase the number of usable permutation vectors. Take for the case of $(N_t, T, K) = (6, 2, 16)$ SPM transmission for example which shows in Fig. 1, there are $\frac{N_t!}{(N_t-T)!}$ candidate permutation vectors and the $d_{\min} = 1$. We remove the grey permutation vectors because of spatial correlation which causes similar channels in different time instants, this will make more bit error in process of decoding.

For the mapping rule, since the permutation pairs with larger Hamming distances result in lower pair-wise error probabilities, we map those pairs to the bit pairs with larger Hamming distances as well. In this way, the cases with large error bits are less likely to occur.

IV. Simulation Results

In this section, the performance of transmission in fast-fading channels with $N_t = 6$ and $N_r = 2$ is evaluated by means of Monte Carlo simulation. We compare SPM which applies our proposed method and the one that randomly generates the permutation set with random mapping rule. Different throughputs and different number of receive antennas are considered. We set $\rho = 0.9$ to simulate the highly-correlated scenario.

Fig. 2 shows the BER comparisons with parameters $(T, M, d_{\min}, K) = (2, 1, 1, 16)$. In modulation we adopt QPSK and 16-QAM, resulting in throughput 4 bps/Hz and 6 bps/Hz, respectively. As can be seen, the SPM with the proposed permutation set and mapping rule delivers better error rate performance, which is up to 5 dB SNR gain when the target BER is 10^{-4} in 4 bps/Hz.

T=2 \ T=1	1	2	3	4	5	6
1		2,1	3,1	4,1	5,1	6,1
2	1,2		3,2	4,2	5,2	6,2
3	1,3	2,3		4,3	5,3	6,3
4	1,4	2,4	3,4		5,4	6,4
5	1,5	2,5	3,5	4,5		6,5
6	1,6	2,6	3,6	4,6	5,6	

Fig. 1. Example of $K = 16$ permutation set design. Frist row and column represent the transmit antenna which is activated at t_{th} time instant. After design process, we remove grey permutation vectors.

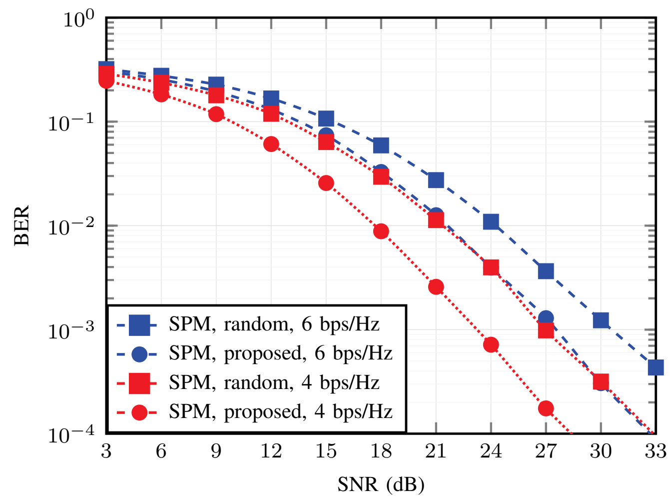


Fig. 2. BER comparisons of SPM with different permutation set design in spatially-correlated fast-fading channel and $(N_t, N_r) = (6, 2)$, under various rate constraints.

V. Conclusions

In this work, we propose the mapping and permutation set design for SPM. The permutation set is designed to maximize the minimum Hamming distance and reduces the effect of the spatial correlation, while the mapping rule is designed to protect the those bit pairs with large Hamming distance. Numerical results demonstrate the novelty of such design and mapping guideline.

References

[1] R. Y. Mesleh, H. Haas, S. Sinaović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[2] I.-W. Lai, J.-W. Shih, C.-W. Lee, H.-H. Tu, J.-C. Chi, J.-C. Wu, and J.-C. Huang, "Spatial permutation modulation for multiple-input multiple-output (mimo) systems," *IEEE Access*, vol. 7, pp. 68 206–68 218, 2019.

[3] E. Başar, Ümit Aygölü, E. Panayırcı, and H. V. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823–832, Mar. 2011.

[4] S. L. Loyka, "Channel capacity of mimo architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

Hybrid Multiple Access Using Simultaneously NOMA and OMA

Hirofumi SUGANUMA, Hiroaki SUENAGA, and Fumiaki MAEHARA
Graduate School of Fundamental Science and Engineering, Waseda University
3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169-8555, Japan
E-mail: fumiaki_m@waseda.jp

***Abstract*—This paper proposes a hybrid multiple access scheme using both non-orthogonal multiple access (NOMA) and orthogonal multiple access (OMA) to overcome the inherent problems of NOMA. In NOMA, as multiple access is performed within the power domain, a small difference in the channel gain between users or an increase in the number of simultaneous users reduces the effectiveness. To cope with these problems, the proposed scheme applies a novel combination of NOMA and OMA as well as NOMA for multiple access patterns, and then determines the best pattern based on the system capacity. The effectiveness of the proposed scheme is demonstrated in comparison with conventional hybrid multiple access using NOMA and OMA through computer simulations.**

I. Introduction

The fifth-generation mobile communication systems (5G), to be realized after 2020, require higher speed, greater capacity, and lower latency. One promising approach being considered is non-orthogonal multiple access (NOMA) [1]–[3], employing successive interference cancellation (SIC). This is because NOMA is a multiple access technique to allocate radio resources within the power domain, thus making it possible to offer more bandwidth per user and to achieve better spectral efficiency than orthogonal multiple access (OMA).

However, NOMA suffers from the inherent problem that if the difference in the channel gain between users is small, then resource allocation becomes insufficient owing to inter-user interference. This results in a lower system capacity than that of OMA [1]–[3]. Moreover, it has also been reported that an increase in the number of simultaneously connected users reduces the benefits of NOMA [2]. To cope with this problem, hybrid multiple access [4], [5], which adaptively switches between NOMA and OMA, in the time domain has been proposed, and its effectiveness has been confirmed. However, as this approach adopts either NOMA or OMA using the entire bandwidth at any time, the improvement of the system capacity is limited.

With this background, we propose a hybrid multiple access scheme using both NOMA and OMA, to further improve the system capacity of NOMA. The proposed approach involves applying a combination of NOMA and OMA as well as NOMA to multiple access patterns, and determining the best multiple access pattern based on the system capacity. In the proposed scheme, as a novel combination of NOMA and OMA is introduced in the same bandwidth, the problems of NOMA under small differences in the channel gain and increases in the number of users can be effectively resolved. The effectiveness of the proposed scheme is demonstrated in comparison with conventional hybrid multiple access in terms of the system capacity through computer simulations.

II. Proposed Scheme

NOMA shares the entire bandwidth with all simultaneous users, allowing inter-user interference to occur, while OMA orthogonally allocates resources for all users. To cope with such inter-user interference, SIC is applied on the receiving side to extract transmitted signals with superposition coding. Figure 1 depicts the system configuration of NOMA. As shown in Fig. 1, NOMA allocates more transmit power for users with poor channel gains, while less transmit power is allocated for users with satisfactory channel gains. Moreover, considering that the users with less transmit power face severe interference from other simultaneous users, such inter-user interference is first decoded and then canceled out using SIC. However, this approach suffers from the inherent problem that when the difference in the channel gain is small SIC fails to decode inter-user interference, which leads to a system performance degradation. In such cases, NOMA provides a lower system performance than OMA [1]–[3], and therefore countermeasures to this problem are expected to be considered. With this background, we propose a hybrid multiple access scheme, simultaneously using NOMA and OMA. In the proposed scheme, as the combination of NOMA and OMA is introduced in the same bandwidth, the resource allocation flexibility can be increased, thereby alleviating the problems occurring for small differences. Moreover, the coexistence of OMA and NOMA in the entire bandwidth can reduce the actual number of simultaneous users in NOMA, which enhances the effectiveness of NOMA.

Figure 2 illustrates the concept of the proposed scheme, where the number of existing users is set to three. As shown in Fig. 2, the three different multiple access patterns of NOMA, OMA, and a combination of NOMA and OMA are adaptively employed. Considering the possible combinations of users, there are five types of resource pattern in total. In the proposed scheme, the instantaneous system capacities of possible multiple access patterns are first calculated using the channel state information (CSI), and then the access pattern achieving the maximum system capacity is determined as the best multiple access pattern. Using the channel gain $|h_k|^2$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

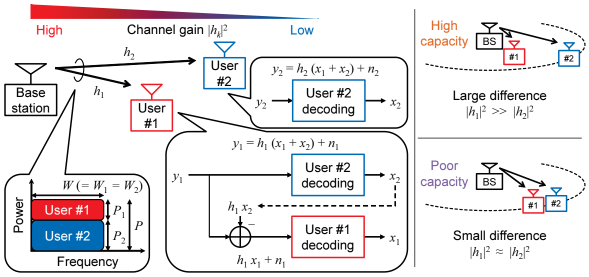


Fig. 1. System configuration of NOMA.

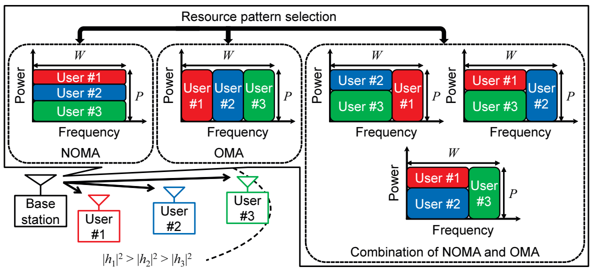
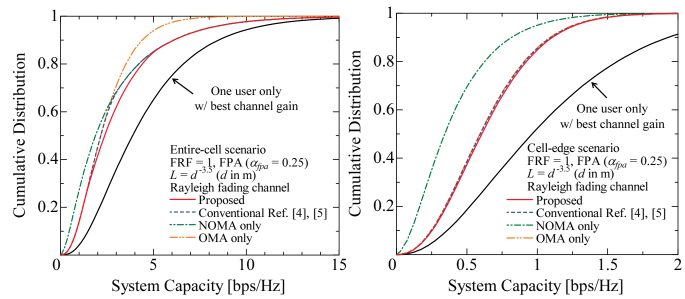


Fig. 2. Concept of the proposed scheme.

TABLE I
SIMULATION PARAMETERS

Number of existing users per cell	3
Total bandwitdh / Transmit power	20 [MHz] / 12 [W]
Cell radius / Path loss model	1000 [m] / $L = d^{-3.5}$ (d in m)
Channel model	Rayleigh fading channel
Thermal noise density / Noise figure	−174 [dBm/Hz] / 10 [dB]



(a) Entire-cell scenario (b) Cell-edge scenario

Fig. 3. Cumulative distribution of system capacity for the proposed scheme.

obtained from the CSI, the instantaneous channel capacity of each multiple access is given by

$$R_{k,\mathrm{NOMA}} = W_k \log_2(1 + P_k|h_k|^2/(\textstyle\sum_{j,|h_k|^2<|h_j|^2} P_j|h_k|^2 + N_0)), \quad (1)$$

$$R_{k,\mathrm{OMA}} = W_k \log_2(1 + P_k|h_k|^2/N_0), \quad (2)$$

where W_k and P_k denote the bandwidth and transmit power of the k-th user, respectively. N_0 is the noise power.

III. Numerical Results

In this section, we verify the effectiveness of the proposed scheme by comparison with NOMA and OMA only in terms of the system capacity. Table I presents the simulation parameters. In our performance evaluation, the number of cells is set to 37, with a frequency reuse factor of 1. Moreover, we assumed two types of calls: calls from an entire cell and calls made only from the cell edge, reflecting the worst case with a low SINR and only small differences in the channel gain between users. Fixed power allocation (FPA) is adopted for the transmit power allocation in NOMA, and the fixed power ratio α_{fpa} $(0 < \alpha_{fpa} < 1)$ is set to 0.25, which defines the difference in the transmit power between simultaneous users.

Figure 3 shows the cumulative distribution of the system capacity for the proposed scheme, where the number of existing users is set to three. For our comparison, the performances of NOMA only, OMA only, and the conventional hybrid multiple access scheme [4], [5] are presented for reference. From Fig. 3(a), it can be seen that NOMA only achieves a high system capacity and OMA only reduces the probability of a low system capacity for the entire-cell scenario. In contrast, it is found from Fig. 3(b) that NOMA dramatically reduces the system capacity in the cell-edge scenario owing to a small difference in the channel gain. These performances represent the general properties of NOMA and OMA. Moreover, it can be observed that the proposed scheme outperforms the conventional scheme, and yields the best system capacity performance regardless of the type of call. This is because the proposed scheme can effectively utilize a combination of NOMA and OMA in the same bandwidth.

IV. Conclusion

In this paper, we proposed a hybrid multiple access scheme using NOMA and OMA to further improve the system capacity of NOMA. The proposed scheme features the application of a combination of NOMA and OMA, as well as NOMA or OMA, for multiple access patterns. The numerical results demonstrated that because the adoption of OMA can improve the effectiveness of NOMA in the same bandwidth, the proposed scheme outperforms conventional hybrid multiple access approaches using NOMA and OMA in a typical multi-cell scenario.

Acknowledgment

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) Grant Number 19K04381.

References

[1] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE," *Proc. 2013 IEEE Globecom Workshops*, pp. 66–70, Dec. 2013.
[2] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Trans. Commun.*, vol. E98-B, no. 3, pp. 403–414, Mar. 2015.
[3] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 176–183, Oct. 2016.
[4] Y. Yuan *et al.*, "Non-orthogonal transmission technology in LTE evolution," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 68–74, July 2016.
[5] A. S. Marcano and H. L. Christiansen, "A novel method for improving the capacity in 5G mobile networks combining NOMA and OMA," *Proc. IEEE 85th Veh. Technol. Conf. (VTC 2017-Spring)*, pp. 1-5, June 2017.

Low-Complexity Maximum Likelihood (ML) Decoder for Space–Time Block Coded Spatial Permutation Modulation (STBC-SPM)

Hsu-Hsuan Tu, Che-Wei Lee and I-Wei Lai
Department of Electrical Engineering, National Taiwan Normal University, Taipei, Taiwan
e-mail:60675051H@ntnu.edu.tw

***Abstract*—Space–time block coded spatial permutation modulation (STBC-SPM) is a new MIMO technique that integraes the STBC-spatial modulation (STBC-SM) and the recently-proposed spatial permutation modulation (SPM). Specifically, SPM exploits the permutation vector to achieve diversity gain by dispersing the signal along the time coordinate. In this work, we examine the STBC-SPM in more details and propose a low-complexity decoder for STBC-SPM by utilizing the orthogonality of the STBC Alamouti code. The numerical simulations demonstrate the superior performance of the STBC-SPM from various aspects. Moreover, the proposed decoder structure achieves up to $90.6\%{\sim}99.9\%$ complexity saving without sacrificing the error rate performance, compared with the optimal maximum likelihood (ML) decoder implemented by the exhaustive search.**

***Index Terms*—multiple-input-multiple-output (MIMO), permutation, spatial modulation (SM), spatial permutation modulation (SPM), space-time block coded spatial modulation (STBC-SM), space-time block coded spatial permutation modulation (STBC-SPM), maximum likelihood (ML)**

I. Introduction

Multiple-input multiple-output (MIMO) systems are widely adopted in many wireless standards. Among numerous MIMO techniques, spatial modulation (SM) has recently gained increasing attention due to its efficiency [1]. Put simply, SM actives one transmit antenna at a time based on the transmitted binary data. In this way, the *spatial symbol* is constructed and transmitted in addition to the conventional QAM symbols. In [2], we have proposed the spatial permutation modulation (SPM) which further exploits the time coordinate by the permutation vector to gain the transmit diversity. Specifically, the transmitted binary data is used to select a permutation vector from a permutation set. Then, at successive time instants, the transmit antenna is sequentially activated according to the entries of the selected permutation vector. SPM is demonstrated to be more efficient than SM, and can be integrated with different SM-based techniques to further enhance their performance. As an example, the space-time block coded spatial permutation modulation (STBC-SPM) is the combination of STBC-SM and SPM [2]. In this work, we further elaborate on the STBC-SPM and propose an low-complexity optimal STBC-SPM decoder based on the maximum likelihood (ML) criterion. Numerical experiments demonstrate that, compared to the exhaustive search, the proposed low-complexity decoder reduces the complexity by two orders of magnitudes.

II. Review of STBC-SM

STBC-SM is a scheme that combines STBC and SM to improve error rate performance by using Alamouti code [3] in [4]. For example, define $\tilde{\mathrm{s}}_1 = [s_1 \;\; -s_2^*]^\top$ and $\tilde{\mathrm{s}}_2 = [s_2 \;\; s_1^*]^\top$ and assuming for transmit antennas, the STBC has the four codewords:

$$\tilde{\chi}_1(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [\tilde{\mathrm{s}}_1 \; \tilde{\mathrm{s}}_2 \; 0 \; 0]^\top, \tilde{\chi}_2(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [\; 0 \; 0 \; \tilde{\mathrm{s}}_1 \; \tilde{\mathrm{s}}_2]^\top,$$
$$\tilde{\chi}_3(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [0 \; \tilde{\mathrm{s}}_1 \; \tilde{\mathrm{s}}_2 \; 0]^\top e^{j\theta_1}, \tilde{\chi}_4(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [\tilde{\mathrm{s}}_2 \; 0 \; 0 \; \tilde{\mathrm{s}}_1]^\top e^{j\theta_1}, \tag{1}$$

where two transmit antennas are activated simultaneously at two time instants. In addition to two QAM symbols (s_1, s_2), each transmission also transmits two bits which are used to select codewords $\tilde{\chi}_i$. For instance, bits '00' and '01' express to utilize the codewords $\tilde{\chi}_1$ and $\tilde{\chi}_2$ respectively. In (1) rotation angle θ_1 is used to maximize coding gain and diversity [4].

III. System Model of STBC-SPM

In [2], STBC-SPM was proposed to improve the error rate performance of STBC-SM, where the codeword length is increased compared with STBC-SM. Taking $T = 2$ as an example

$$\chi(s_1, s_2, s_3, s_4) = [\tilde{\chi}_i(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) \quad \tilde{\chi}_{j,j\neq i}(\tilde{\mathrm{s}}_3,\tilde{\mathrm{s}}_4)], \tag{2}$$

where the selection of the successive STBC-SM codewords are based on the permutation vectors. For example, if permutation vector $[1, 2]^\top$ is used, the first and second STBC-SM codewords are selected to form the STBC-SPM. As can be seen, compared with the STBC-SM, the data are conveyed by the permutation vectors, rather than the index of the STBC-SM codeword. In fact, for STBC-SM with $N_t = 4$ we can have another two codewords in addition to those in (1)

$$\tilde{\chi}_5(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [\tilde{\mathrm{s}}_1 \; 0 \; \tilde{\mathrm{s}}_2 \; 0]^\top e^{j\theta_2}, \tilde{\chi}_6(\tilde{\mathrm{s}}_1,\tilde{\mathrm{s}}_2) = [0 \; \tilde{\mathrm{s}}_1 \; 0 \; \tilde{\mathrm{s}}_2]^\top e^{j\theta_2} \tag{3}$$

where N_t and θ_2 are number of transmit antennas and another rotation angle imposed for performance optimization, respectively. When used codewords change from 4 to 6, the number of combinations that can be used are also increased from $4 \cdot 3 = 12$ to $6 \cdot 5 = 30$. In other words, the number of bits that can be transmitted is also increased from 3 bits to 4 bits. In this case, numebr of codebooks n is 3. Accroding to [4],

$$\theta_k = \tfrac{(k-1)\pi}{2n}, \text{for QPSK}, 16\text{QAM} \tag{4}$$

for $1 \leq k \leq n$, we can get θ_1 and θ_2. In addition, STBC-SPM maximize coding gain and diversity by using rotation angle to further improve error rate performance like STBC-SM. We group the permutation vectors and give each ones a rotation angle. In this way, the minimum Hamming distance of the permutation set can be maximized by rotation. When two STBC-SM codewords, we can group the permutation vectors as the following

$$P_1 = \left\{ \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix} \right\}, P_2 = \left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\},$$
$$P_3 = \left\{ \begin{bmatrix} 6 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}, P_4 = \left\{ \begin{bmatrix} 6 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\}. \tag{5}$$

Then, according to the group index, rotation angle

$$\varphi_i = \frac{(i-1)\pi}{8} \tag{6}$$

is imposed respectively as stated in [3]. Following (3), the transmitted codeword of STBC-SPM is given by

$$\chi(s_1, s_2, s_3, s_4) = \left[\tilde{\chi}_{p_{l,1}}(s_1, s_2)e^{j\varphi_i} \quad \tilde{\chi}_{p_{l,2}}(s_3, s_4)e^{j\varphi_i}\right] \tag{7}$$

with $\mathbf{p}_l \in P_i$ and $p_{l,j} = 1, \ldots, 6$.

IV. Low-Complexity ML Decoder for STBC-SPM

In this section, we formulate the ML decoder for STBC-SPM system with N_t transmit and N_r receive antenna. The received $N_r \times$ 4 signal $\mathbf{Y}$ can be expressed as

$$\mathbf{Y} = \mathbf{HX} + \mathbf{N} \tag{8}$$

where $\mathbf{X} \in \chi$ is the $N_t \times 4$ STBC-SPM transmission signal. $\mathbf{H}$ and $\mathbf{N}$ represent the $N_r \times N_t$ quasi-static Rayleigh flat fading channel and $N_r \times 4$ complex Gaussian random variable with zero mean and unit variance noise, respectively. The ML detection chooses a codeword that minimizes the following decision metric:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X} \in \chi} \left\|\mathbf{Y} - \mathbf{HX}\right\|^2. \tag{9}$$

The minimization in (8) can be simplified by first reformulating the equivalent channel model from (8):

$$\mathbf{y}_1 = \mathcal{H}_{p,1}\begin{bmatrix} s_1 & s_2 \end{bmatrix}^\top + \mathbf{n}_1, \quad \mathbf{y}_2 = \mathcal{H}_{p,2}\begin{bmatrix} s_3 & s_4 \end{bmatrix}^\top + \mathbf{n}_2 \tag{10}$$

where $\mathcal{H}_{p,i}$ is the $2N_r \times 2$ equivalent channel matrix, p means permutation vector . $\mathbf{y}_i$ and $\mathbf{n}_i$ represent the $2N_r \times 1$ equivalent received signal and noise vectors, $i = 1, 2$, respectively.

Due to the orthogonality of Alamouti's STBC, the columns of $\mathcal{H}_\ell$ are orthogonal to each other. Consider the STBC-SPM transmission model, we have six different realizations for $\mathcal{H}_\ell$, which are given for N_r receive antennas as

$$\mathcal{H}_{i,\ell} = \begin{bmatrix} h_{1,\ell_a}\phi_i & h^*_{1,\ell_b}\phi^*_i & \cdots & h_{N_r,\ell_a}\phi_i & h^*_{N_r,\ell_b}\phi^*_i \\ h_{1,\ell_b}\phi_i & h^*_{1,\ell_a}\phi^*_i & \cdots & h_{N_r,\ell_b}\phi_i & h^*_{N_r,\ell_a}\phi^*_i \end{bmatrix}^\top \tag{11}$$

where h is the channel fading coefficient and $\phi_i = e^{j\varphi_i}$. In this case $i = 1, \ldots, 4$ indicates the rotated angles. The groups ℓ_a, ℓ_b denote transmit antenna pair.

$$\hat{s}_{i,\ell_j} = \arg\min_{s_i \in \gamma} \left\|\mathbf{y} - \mathbf{h}_{i,\ell_j} s_i\right\|^2 \qquad i = 1, 2, j = 1, 2 \tag{12}$$

where $\mathcal{H}_{i,\ell_1} = [\mathbf{h}_{1,\ell_1}\mathbf{h}_{2,\ell_1}]$ and $\mathcal{H}_{i,\ell_2} = [\mathbf{h}_{1,\ell_2}\mathbf{h}_{2,\ell_2}]$, $1 \le \ell_1, \ell_2 \le 6$, group $\ell_1, \ell_2 \in P$. $\mathbf{h}_{i,\ell_1}$ and $\mathbf{h}_{i,\ell_2}$, $j = 1, 2$, are $2N_r \times 1$ column vector. The associated minimum ML metrics m_{i,p_j} for s_i are

$$m_{i,\ell_j} = \min_{s_i \in \gamma} \left\|\mathbf{y} - \mathbf{h}_{i,\ell_j} s_i\right\|^2 \qquad i = 1, 2, j = 1, 2 \tag{13}$$

respectively. Their summation

$$m_p = m_{1,\ell_1} + m_{2,\ell_1} + m_{1,\ell_2} + m_{2,\ell_2}, 1 \le p \le 16. \tag{14}$$

By using the accumulated metrics, the decoder can optimally identify the transmitted permutation vector $\hat{p}$:

$$\hat{p} = \arg\min_p m_p. \tag{15}$$

V. Simulation Results and Comparisons

In this section, the performance of transmission in slow-fading channels with N_r =1 is evaluated by means of Monte Carlo simulation. Fig. 1 compares STBC-SM and STBC-SPM with N_t = 3 and 4. STBC-SPM shows better error performance than STBC-SM.

For the complexity comparison, the total number of ML metric calculations is listed in Table. 1. In particular, the complexity of exhaustive search is $C \cdot M^4$, where C and M represent the number of permutation vectors and number of constellation points, respectively. Meanwhile, the proposed low-complexity decoder requires $4c \cdot P \cdot M$, where c and P denotes the number of codewords and number of permutation groups in (5). As can be seen, the complexity saving is up to three order of magnitudes.

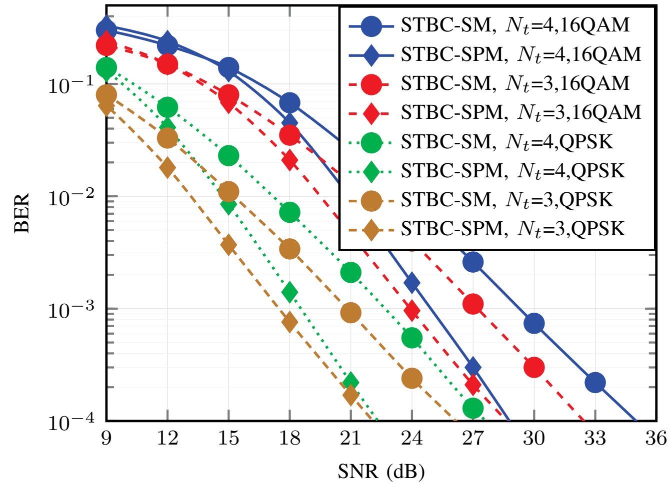


Fig. 1. BER comparisons of MIMO systems using STBC-SM and STBC-SPM with various N_t and modulation alphabets.

TABLE I
STBC-SM and STBC-SPM complexity comparison

N_t	modulation	exhaustive search	proposed method	complexity saving
3	QPSK	1024	96	90.6%
	16QAM	262144	384	99.9%
4	QPSK	4096	384	90.6%
	16QAM	1048576	1536	99.9%

VI. Conclusions

In this work, we elaborate on the STBC-SPM MIMO scheme and its low complexity ML decoder. We can see that STBC-SPM achieves better error rate performance than STBC-SM. Additionally, by using the complexity structure, the optimal STBC-SPM decoder only requires 0.1% complexity compared with the exhaustive search. The work report here can be generalized to STBC-SPM with arbrtiary number of transmit antennas, and the design of suboptimal decoder with further complexity reduction, which are our future research directions.

References

[1] R. Y. Mesleh, H. Haas, S. Sinaović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[2] I.-W. Lai, J.-W. Shih, C.-W. Lee, H.-H. Tu, J.-C. Chi, J.-S. Wu, and Y.-H. Huang, "Spatial permutation modulation for multiple-input multiple-output (mimo) systems," *IEEE Access*, vol. 7, pp. 68 206–68 218, 2019.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] E. Başar, Ümit Aygölü, E. Panayırcı, and H. V. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823–832, Mar. 2011.

Comparison of Inter-Cell and Co-Channel Interference Power between A-ZCZ and GMO-ZCZ Sequence Sets by Computer Simulation

Hideyuki Torii
Department of Information Network and Communication
Kanagawa Institute of Technology
Atsugi, Japan
torii@nw.kanagawa-it.ac.jp

Takahiro Matsumoto
Graduate School of Sciences and Technology for Innovation
Yamaguchi University
Ube, Japan
matugen@yamaguchi-u.ac.jp

***Abstract*—In the present paper, we compare inter-cell and co-channel interference power between A-ZCZ sequence sets and that of GMO-ZCZ sequence sets by computer simulation. The simulation results shows that the A-ZCZ sequence set is advantageous when the cell radius is small and the GMO-ZCZ sequence set is advantageous when it is large.**

***Index Terms*—ZCZ sequence sets, AS-CDMA, inter-cell interference, co-channel interference**

I. Introduction

Code-division multiple-access (CDMA) has been widely applied in digital cellular systems. In CDMA systems, channel separation is provided by pseudo-random codes referred to as spreading sequences.

In recent years, approximately synchronized CDMA (AS-CDMA) systems have attracted a great deal of attention because co-channel interference within a cell does not exist in some types of AS-CDMA systems. In such AS-CDMA systems, zero-correlation zone (ZCZ) sequence sets are used as spreading sequences in order to realize this advantage [1], [2]. In addition, asymmetric ZCZ (A-ZCZ) sequence sets [3] and generalized mutually orthogonal ZCZ (GMO-ZCZ) sequence sets [4] have been proposed by extending the ZCZ sequence sets. An A-ZCZ sequence set is composed of several ZCZ sequence subsets, and has the property whereby the zero-cross-correlation zone (ZCCZ) length between different sequence subsets is larger than the ZCZ length of the same sequence subset. Similarly, a GMO-ZCZ sequence set is composed of several ZCZ sequence subsets. However, it has the property whereby the ZCZ length of the same sequence subset is larger than the ZCCZ length between different sequence subsets. By assigning the sequence subsets of the A-ZCZ sequence sets or the GMO-ZCZ sequence sets to different cells, there is a possibility that inter-cell interference power from neighboring cells is reduced. However, concrete comparison of inter-cell interference power between the A-ZCZ sequence sets and the GMO-ZCZ sequence sets has not been done until now. In the present paper, we compare inter-cell and co-channel interference power of the A-ZCZ sequence sets and that of the GMO-ZCZ sequence sets by computer simulation, and clarify the property of their interference power.

II. Preliminaries

In this section, we define A-ZCZ sequence sets, GMO-ZCZ sequence sets, and related terms.

Let Z be a set with N sequence sets containing M sequences of period P. Z can be represented as

$$\begin{aligned} &Z = \{Z_n \mid 0 \le n \le N-1\}, \\ &Z_n = \left\{Z_m^{(n)} \mid 0 \le m \le M-1\right\}, \\ &Z_m^{(n)} = \left(z_0^{(n,m)}, \cdots, z_p^{(n,m)}, \cdots, z_{P-1}^{(n,m)}\right), \end{aligned} \tag{1}$$

where Z_n, $Z_m^{(n)}$, and $z_p^{(n,m)}$ denote a sequence subset, a sequence, and a sequence element, respectively. Suppose that all of the sequences in Z satisfy the following conditions:

$$\begin{aligned} &\forall n, \forall m, 1 \le |\tau| \le L_s, \\ &\quad R_{Z_m^{(n)}, Z_m^{(n)}}(\tau) = 0, \end{aligned} \tag{2}$$

$$\begin{aligned} &\forall n, \forall m_0 \ne m_1, |\tau| \le L_s, \\ &\quad R_{Z_{m_0}^{(n)}, Z_{m_1}^{(n)}}(\tau) = 0, \end{aligned} \tag{3}$$

$$\begin{aligned} &\forall n_0 \ne n_1, \forall m_0, \forall m_1, |\tau| \le L_d, \\ &\quad R_{Z_{m_0}^{(n_0)}, Z_{m_1}^{(n_1)}}(\tau) = 0, \end{aligned} \tag{4}$$

where $R_{x,y}(\tau)$ is the periodic correlation function between x and y. In addition, L_s is referred to as a ZCZ length in each sequence subset, and L_d is referred to as a ZCCZ length between different sequence subsets. If $L_s < L_d$ is satisfied, Z is referred to as an A-ZCZ sequence set, and it is represented as $A(P, \{N \times M\}, \{L_s, L_d\})$ in order to exhibit its parameters. On the other hand, if $L_s > L_d$ is satisfied, Z is referred to as a GMO-ZCZ sequence set, and it is represented as $G(P, \{N \times M\}, \{L_s, L_d\})$.

III. Computer Simulation

In this simulation, we assume two cells, that is, one is a desired cell and the other is a neighboring cell. In addition, we assume that a base station receives signals of the same power

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

from each terminal in a cell because of power control. However, inter-cell interference signals from the neighboring cell attenuate depending on the distance. $A(128, \{2 \times 8\}, \{6, 14\})$ and $G(128, \{2 \times 8\}, \{14, 6\})$ are constructed using a quadriphase perfect sequence of period 16 and a bi-phase 8 dimensional orthogonal matrix from the methods mentioned in [3] and [4], respectively. The A-ZCZ sequence set and the GMO-ZCZ sequence set have 2 sequence subsets, and each sequence subset includes 8 sequences. In this simulation, one sequence subset is assigned to the desired cell and the other sequence subset is assigned to the a neighboring cell. Therefore, the maximum number of terminals in each cell is 8. Table I shows simulation parameters.

TABLE I
SIMULATION PARAMETERS

Number of cells	2
Type of ZCZ sequence sets	A-ZCZ, GMO-ZCZ
Sequence length	128 [chips]
Total number of sequences	16
Number of sequence subsets	2
Number of sequences in each sequence subset	8
ZCZ length of the same sequence subset	6 [chips] (A-ZCZ), 14 [chips] (GMO-ZCZ)
ZCCZ length between different sequence subsets	14 [chips] (A-ZCZ), 6 [chips] (GMO-ZCZ)
Modulation method	QPSK
Chip rate	3.0 [Mchips/sec]
Cell radius	600(m), 1400(m)
Maximum number of terminals in each cell	8
Distribution of terminals	Uniform distribution

Figure 1 shows a simulation result when the cell radius is 600(m). In this case, although co-channel interference in a cell does not occurs theoretically, inter-cell interference may occurs. The horizontal axis is the number of terminals in each cell and the vertical axis is the ratio of average interference power to desired signal power. The result shows that the average interference power of the GMO-ZCZ sequence set is larger than that of A-ZCZ sequence set regardless of the number of terminals in each cell.

Figure 2 shows a simulation result when the cell radius is 1400(m). In this case, although co-channel interference of the GMO-ZCZ sequence set in a cell does not occurs theoretically, that of the A-ZCZ sequence set and inter-cell interference may occurs. The result shows that the average interference power of the A-ZCZ sequence set is larger than that of GMO-ZCZ sequence set when the number of terminals is greater than or equal to 3.

From both the results, the A-ZCZ sequence set is advantageous when the cell radius is adequately small, the GMO-ZCZ sequence set is advantageous when it is large. Note that in the all cases, the average interference power is extremely small.

IV. CONCLUSIONS

In the present paper, we have compared inter-cell and co-channel interference power of A-ZCZ sequence sets and that of GMO-ZCZ sequence sets by computer simulation. The simulation results shows that the A-ZCZ sequence set is advantageous when the cell radius is small and the GMO-ZCZ sequence set is advantageous when it is large. This result is expected to be useful for designing AS-CDMA systems using A-ZCZ sequence sets or GMO-ZCZ sequence sets.

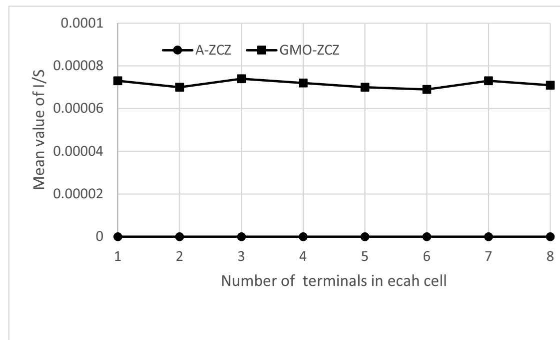


Fig. 1. Number of terminals vs. mean value of I/S (Cell radius is 600(m))

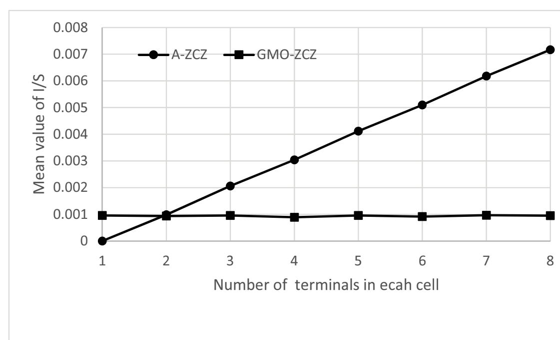


Fig. 2. Number of terminals vs. mean value of I/S (Cell radius is 1400(m))

REFERENCES

[1] N. Suehiro, "A signal design without co-channel interference for approximately synchronized CDMA systems," IEEE J. Select. Areas Commun., vol.12, no.5, pp.837–841, June 1994.

[2] P. Fan and L. Hao, "Generalized orthogonal sequences and their applications in synchronous CDMA systems," IEICE Trans. Fundamentals, vol.E83-A, no.11 pp.2054–2069, Nov. 2000.

[3] H. Torii, T. Matsumoto and M. Nakamura, "A new method for constructing asymmetric ZCZ sequence sets," IEICE Trans. Fundamentals, vol.E95-A, no.9, pp.1577-1588, Sept. 2012.

[4] H. Torii, M. Satoh, T. Matsumoto, and M. Nakamura, "Quasi-optimal and optimal generalized mutually orthogonal ZCZ sequence sets based on an interleaving technique," International Journal of Communication, issue 1, vol.7, pp.18-25, May 2013.

Single Carrier Block Transmission Schemes for Acoustic Communications and Their Field Evaluation Results

Ryuki Sano[1] , Hiroshi Kubo[2]
[1]Graduate School of Science and Engineering, Ritsumeikan University,
[2]Department of Electronic and Computer Engineering, Faculty of Science and Engineering, Ritsumeikan University,
[1,2]1-1-1 Noji-Higashi, Kusatsu-shi, Shiga 525-8577, Japan
[1]ri0050sr@ed.ritsumei.ac.jp , [2]kubohiro@fc.ritsumei.ac.jp

***Abstract*—This paper proposes transmission technologies for acoustic communications in order to mitigate doubly-selective environments. The single carrier block (SCB) transmission schemes using frequency domain equalization (FDE) are effective for frequency selectivity, but it suffers from performance degradation due to time selectivity. This paper proposes SCB transmission schemes for acoustic communications which introduce a frame format suitable for channel interpolation (CI), and employs cyclic delay diversity (CDD) in order to extend communication distance. In particular, the proposed schemes can mitigate tracking error caused by time-varying channels, owing to CI based on multiple estimated channel impulse responses (CIRs). Finally, it is confirmed that the proposed schemes of 8kbps have excellent performance in a delay spread of up to 10msec and a Doppler frequency of around 1Hz, on the basis of field evaluation results in terrestrial and shallow underwater environments.**

***Index Terms*—Acoustic Communications, Single Carrier Block Transmission, Frequency Domain Equalization, Doubly-Selective Channels, Channel Estimation**

I. Introduction

Acoustic communications suffer from severe doubly-selective environment regardless of terrestrial environment or underwater environments [1],[2]. It is difficult for acoustic communications to employ automatic repeat-request (ARQ), if acoustic communications assume point to multi-point (P-MP) communications and data communications with small latency. Thus, there are the three problems for the above mentioned acoustic communications: 1) to cope with several msec delay spread in acoustic communication environment; 2) to cope with slowly time-varying channel caused by human body movement in terrestrial acoustic communications (TRACs) or wave fluctuations in underwater acoustic communications (UWACs); 3) to maintain communication quality and communication distance regardless of speaker performance and communication environment.

This paper proposes single carrier block (SCB) transmission schemes for acoustic communications using frequency domain equalization (FDE) [3]. As a feature of the proposed schemes, for the item 1), it employs the SCB transmission schemes in the presence of intersymbol interference (ISI). For the item 2), it employs a frame structure suitable for channel interpolation (CI), using multiple estimated channel impulse responses (CIRs) to improve tracking performance. For the item 3), it employs cyclic delay diversity (CDD) to improve the receiver sensitivity [4]. Finally, it is confirmed that the proposed schemes have excellent performance in TRAC and shallow UWAC channels.

II. Proposed Frame Structure

Fig. 1 shows the proposed frame structure for acoustic communications. In the proposed frame structure, one training sequence (TS) is inserted every two data blocks. The data block transmits N_D-symbol information with an N_C-symbol cyclic prefix (CP). The training sequence consists of an N_M-symbol M-sequence and N_C-symbol CP, where the order of the primitive polynomial for the M-sequence is g.

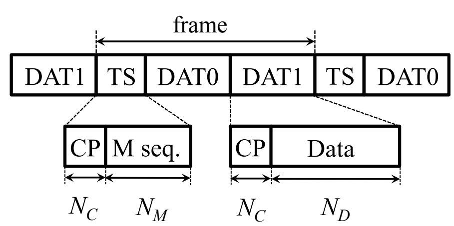


Fig. 1. Proposed frame structure.

III. Proposed Channel Interpolation Method

This section proposes CI in order to track time-varying channels. To improve the tracking performance, the proposed CI interpolates the two CIR estimated by the closest two TSs in the time domain, where CIR estimation in the time domain is described in [5]. The proposed CI estimates the CIR as follows: it derives discrete timing from TSs located prior to and posterior to data block. it derives the estimated CIR by means of CI according to the derived timing.

IV. TRAC Field Evaluation Results

This section shows the field evaluation results of the proposed SCB transmission schemes under the following two conditions: a) bit error rate (BER) and block error rate (BLER)

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

performances with and without CDD in a TRAC static environment; b) BER and BLER performances with and without CI in the terrestrial quasi-static environment. These experiments evaluated 8kbps communications. Table I shows experimental parameters of the proposed SCB transmission schemes.

Fig. 2 shows BER and BLER performances in actual field experiments as a function of distance in a 10m×10m room corresponding to the condition a). N of CDD (N) indicates the number of cyclic symbols in CDD. As shown in Fig. 2, CDD (4) which introduced 4 symbol delay, has obtained path diversity gain. The performance of CDD (0) at 7m seems to degrade owing to beat interference caused by zero delay CDD. Fig. 3 shows BER and BLER performances in an actual field experiment as a function of f_DT corresponding to the condition b). Fig. 3 shows that CI can improve tracking performance on time-varying channels.

TABLE I
COMMUNICATION SYSTEM CONFIGURATION

data modulation	QPSK	
frame structure	block size	N_D=512
	CP size	N_C=64
	SW length	N_M=127
filter	raised cosine filter	
	roll-off factor	0.2
demodulator	FDE weight	MMSE
error correction code	Reed-Solomon	RS(255, 225)
sampling rate	96kHz	
quantization bit	24bit	
carrier frequency	TRAC	22.6kHz
	UWAC	30kHz

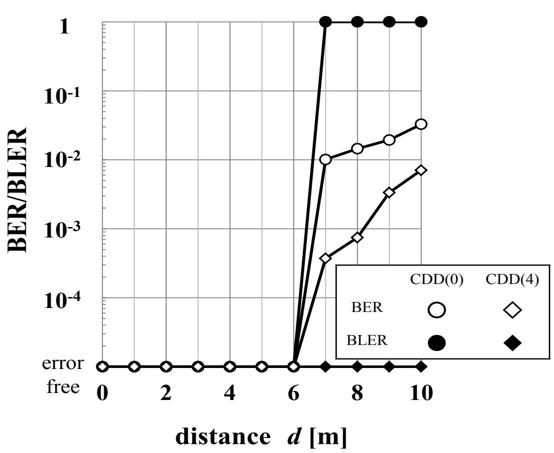


Fig. 2. BER and BLER performances in actual field experiment as a function of distance.

V. UWAC Field Evaluation Results

This section shows field evaluation results in the shallow-sea quasi-static environment, where the Tx transducer depth is 2m, the Rx transducer depth is 0.7m, and the communication distance is 370m. Table II shows the field evaluation results. In the shallow underwater channel, the proposed schemes improve BER and BLER performances for the sake of equalization using CI. This is because quasi-static shallow-sea channels change due to wave fluctuations.

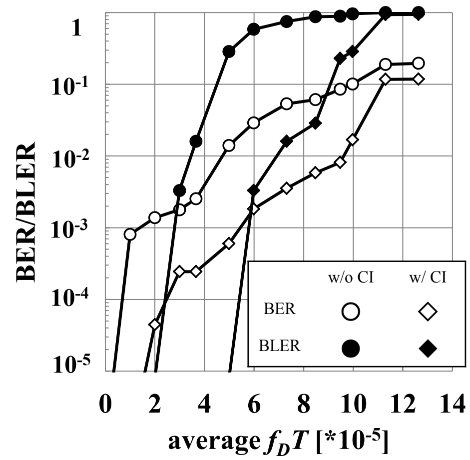


Fig. 3. BER and BLER performances in actual field experiment as a function of f_DT.

TABLE II
UNDERWATER ACOUSTIC COMMUNICATION RESULTS CONDUCTED AT A COMMUNICATION DISTANCE OF 370M

	BER	BLER
w/o CI	1.45×10^{-2}	1.00×10^{0}
w/ CI	1.49×10^{-3}	error free

VI. Conclusion

This paper has proposed the SCB transmission employing the CI and CDD for large time-dispersive channels of acoustic communications. Through field evaluation results using actual speakers, it has confirmed that CDD improves receiver sensitivity and CI improves tracking performance on time-varying channels. The field evaluation results have confirmed that the proposed SCB transmission schemes are suitable for actual quasi-static acoustic communication environments.

Acknowledgments This work was partly supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Numbers JP16K06374 and JP19K04382.

References

[1] A.C. Singer, J.K. Nelson and S.S. Kozat, "Signal processing for underwater acoustic communications," IEEE Commun. Mag., vol. 47, pp. 90–96, Jan. 2009.

[2] N. Shinmen, T. Ebihara and K. Mizutani, "Software-based modem for near field acoustic communication," The 1st IEEE Global Conference on Consumer Electronics, pp. 152–155, Oct. 2012.

[3] T. Walzman and M. Schwartz, "Automatic equalization using the discrete frequency domain," IEEE Trans. IT, vol. 19, pp. 59–68, Jan. 1973.

[4] T. Yamamoto, K. Takeda and F. Adachi, "Joint frequency-domain equalization and despreading for multi-code DS-CDMA using cyclic delay transmit diversity," IEICE Trans. Commun., vol. E92-B, no. 5, May 2009.

[5] H. Kubo and M. Miyake, "Single carrier modulation scheme employing a list Viterbi equalizer with a metric-criteria combining scheme," in Proc. of PIMRC'99, Sept. 1999.

A Differential Multiple Single Carrier Modulation Scheme for Underwater Acoustic Communications and Its Actual Evaluation Results

Rintaro Yoshii[1], Yoshiki Tsukamoto[1], Takuya Takahashi[1], and Hiroshi Kubo[2]
[1]Graduate School of Science and Engineering, Ritsumeikan University,
[2]Department of Electronic and Computer Engineering, Faculty of Science and Engineering, Ritsumeikan University,
[1,2]1-1-1 Noji-Higashi, Kusatsu-shi, Shiga 525-8577, Japan
[1]{ri0055hp, ri0051vf, ri0051se}@ed.ritsumei.ac.jp , [2]kubohiro@fc.ritsumei.ac.jp

***Abstract*—This paper discusses actual evaluation results of a differential multiple single carrier (MSC) modulation scheme for underwater acoustic communications (UWACs). UWACs suffer from severe doubly-selective channels and inter-carrier interference due to non-uniform frequency offset. In addition, assuming image data communications, UWACs require improved frequency efficiency. To resolve these issues, this paper proposes a wireless transmission scheme that employs differential MSC modulation scheme. Next, this paper investigates the parameters for 8kbps-UWACs and 64kbps-UWACs aiming at wireless control of autonomous underwater vehicles and image data communications, respectively. Finally, two actual evaluation results in an acoustic experimental pool and an actual shallow-sea environment confirm that the UWAC MODEMs have excellent performance on severe doubly-selective channels.**

***Index Terms*—Underwater Acoustic Communications, Channel Prediction, Doubly-Selective Channels, Per-Survivor Processing, Multi-Carrier Modulation Scheme, Multi-Level Modulation**

I. Introduction

Recently, wireless control of autonomous underwater vehicles (AUVs) and image data communications for ocean resource search are of increasing interest [1]. For realization by underwater acoustic communications (UWACs), the following issues must be resolved: 1) to improve performance on severe doubly-selective channels; 2) to cope with inter-carrier interference due to non-uniform frequency offset; 3) to improve frequency efficiency for image data communications. To resolve these issues, this paper proposes a communication scheme that employs multi-carrier (MC) modulation scheme and differential encoding/multiple differential detection based on per-survivor processing (PSP-MDD) with channel prediction [2] for frequency selectivity and time selectivity of the item 1), respectively. With respect to the items 2) and 3), this paper employs multiple single carrier (MSC) modulation scheme and multi-level modulation, respectively. MSC modulation scheme parallelly transmits information using narrowband single carriers. Next, this paper investigates transmission schemes that realize 8kbps and 64kbps for wireless control of AUVs and image data communications, respectively. Finally, this paper performs two experimental evaluations and confirms that the UWAC MODEMs using the proposed scheme are effective on severe doubly-selective UWAC channels.

II. Communication Scheme for UWACs

Fig. 1 shows a block diagram of the communication model employing the proposed scheme.

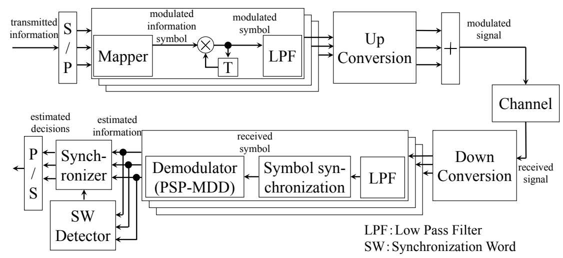


Fig. 1. Communication model employing the proposed scheme.

A. MSC Modulation Scheme

This paper employs MSC modulation scheme, which is one of MC modulation schemes. This is because the extremely large signal frequency bandwidth to center carrier frequency ratio in UWACs. Under the above-mentioned condition, it is problematic that the Doppler shifts differ significantly at the upper and lower ends of the occupied bandwidth. MSC modulation scheme can compensate the Doppler shift for each carrier, so MSC scheme is suitable for the above-mentioned condition.

B. PSP-MDD with Channel Prediction

PSP-MDD [2] discussed in this paper is equivalent to PSP-MLSE (maximum-likelihood sequence estimation). The previous work [2] discusses the channel estimation weights for the inverse-modulated value, which are used in calculating the estimated channel impulse response (CIR). The channel estimation weights for the improvement of the receiver sensitivity is averaging of the estimated CIR the most recent N symbols. N is the observation range of the received signals. Also, regarding the channel estimation weights for the improvement of the tracking capability on fast time-varying channels, the order of channel prediction L depends on N, and $L = N - 1$.

C. Algorithm to Reduce the Number of States

This paper employs decision-feedback sequence estimation (DFSE) [3]. The number of states for the Viterbi algorithm

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

is $M^{N-1-\Delta V}$, where the symbol length referring to the surviving path for DFSE is ΔV and the multi-level modulation index is M. PSP-DFSE with $\Delta V = 0$ corresponds to PSP-MLSE.

III. Experimental Evaluations

A. Parameters of UWAC MODEMs

Table I shows the parameters of the transmission schemes.

TABLE I
Parameters of the Transmission Schemes

Items	Wireless control of AUVs	Image data communications
Transmission rate	8kbps	64kbps
Modulation	QPSK	16PSK
Filter	20% root cosine roll-off	
Transmission rate per subcarrier	93.75sps	
The number of subcarriers	64	256
The center frequency	30kHz	28kHz
Reed-Solomon code	(240, 166, 8)	(255, 177, 8)
Abbreviation	DMSC2	DMSC4

B. Evaluation Results in Acoustic Experimental Pool

This section discusses an actual evaluation for a doubly-selective environment. In an acoustic experimental pool, two receive transducers are separated by 1.0m and fixed at a water depth of 1.0m. A transmit transducer is placed below the center of each receive transducer and moves periodically a distance of 0.25m to 1.25m from the receive transducer at average speed v. The moving speed is not constant, therefore, v is approximated using the average speed obtained from the moving distance and the corresponding time. In this experiment, a delay wave is generated mechanically by a reflector to simulate a doubly-selective environment.

Figs. 2 and 3 show block error rate (BLER) performances of PSP-MDD as a function of v, where the transmission schemes are DMSC2 and DMSC4, respectively. In addition, BLER performances of differential detection (DD) are also plotted for references. In Figs. 2 and 3, N_R is the number of receive antennas. In Fig. 3, PSP-DFSE exhibits better BLER performance than PSP-MLSE, therefore, the BLER performance of PSP-DFSE is plotted only when $L = 3$. ΔV in PSP-DFSE is 1. It seems that the results may be influenced by the estimated CIR obtained by channel prediction, and details are currently under investigation.

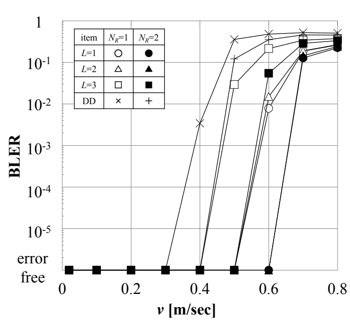


Fig. 2. BLER performance as a function of v (DMSC2).

Figs. 2 and 3 show the following results: PSP-MDD with 1st channel prediction has the best tracking performance in both

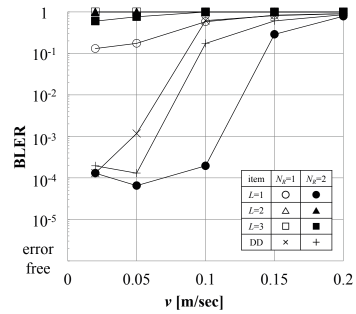


Fig. 3. BLER performance as a function of v (DMSC4).

DMSC2 and DMSC4; the tracking performance of PSP-MDD in the doubly-selective environment is improved compared with that of DD.

C. Evaluation Results in Actual Shallow-Sea Environment

This section discusses the BLER performances of UWAC MODEMs in an actual shallow-sea environment in the presence of frequency selectivity. When the transmission scheme is DMSC4, the evaluation results confirm that PSP-MDD ($L = 0$) has a BLER of less than 10^{-3}, assuming a 0.7m distance between the two receive transducers, a water depth of 25m, and a 10m distance between the transmit transducer and the receive transducers. When the transmission scheme is DMSC2, the evaluation results confirm that PSP-MDD ($L = 0$) is free of block error, assuming a 0.7m distance between the two receive transducers, a water depth of 2m, and 370m distance between the transmit transducer and the receive transducers.

IV. Conclusion

This paper has proposed a wireless transmission scheme based on differential MSC modulation scheme for severe doubly-selective channels like UWACs. This paper also experimentally has evaluated UWAC MODEMs with the proposed scheme in two environments. Evaluation results in an acoustic experimental pool have confirmed that PSP-MDD with 1st channel prediction has excellent performance in a doubly-selective environment. Next, evaluation results in a shallow-sea environment have confirmed that the proposed schemes using 16PSK and QPSK enable communication up to distances of around 10m and 370m, respectively.

Acknowledgments This work was partly supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Numbers 16K06374 and 19K04382.

References

[1] H. Ochi, Y. Watanabe and T. Simura, "Basic study of underwater acoustic communication using 32-quadrature amplitude modulation," Jpn. J. App. Phys., vol. 44, no. 6B, pp. 4689–4693, June 2005.

[2] H. Kubo, A. Okazaki, K. Tanada, B. Penther and K. Murakami, "A multiple-symbol differential detection based on channel prediction for fast time-varying fading," IEICE Trans. Commun., vol. E88-B, no. 8, pp. 3393–3400, Aug. 2005.

[3] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," IEEE Trans. Commun., vol. 37, no. 5, pp. 428–436, May 1989.

Image Quality Improvement of Underwater Images in Ideal HSI Color Space

Ibuki Yoshida
Department of Computer Science
Tokyo City University
Setagaya-ku, Tokyo 158-8557, JAPAN
g1881850@tcu.ac.jp

Akira Taguchi
Department of Computer Science
Tokyo City University
Setagaya-ku, Tokyo 158-8557, JAPAN
ataguchi@tcu.ac.jp

***Abstract*— In the underwater circumstance, the light attenuates, and its attenuation rate depends on the wavelength. Long wavelength light is easily absorbed. Also, the light attenuation rate of blue and green component changes according to the water quality, thus, the white balance of images is lost. The attenuation of the light reduces colorfulness and contrast of images, furthermore, causes blurring. When photographing, the illumination becomes uneven for the subject, thus, brightness unevenness occurs in the acquired captured images. In this paper, an underwater image signal is converted to the ideal HSI color space and apply processing suitable for various deterioration factors to saturation and intensity independently. High quality improvement of underwater image was realized.**

Keywords—white balance, ideal HSI color space, intensity, saturation

I. Introduction

Restoring degraded underwater images are a very important task. Underwater images generally are degraded by a lot of deterioration factors. Underwater is 800 times denser than in the air [1]. In the underwater, long wavelength light which is red component is easily absorbed and hardly scattered, while, short wavelength light which are blue and green components are hardly absorbed and easily scattered. Since the attenuation rate of green and blue light changes according to the water quality, the white balance of the underwater images is lost. The light attenuation causes low contrast and colorfulness and blurring. In addition, due to the illumination condition at the time of shooting, underwater images are a cause of uneven optical brightness.

For the color image processing, hue preserving is required. Therefore, in many techniques for processing intensity and saturation in color images are performed in transformed color space such as HSI, YIQ, HSV, etc. Transforming from RGB color space to another color space and image processing in these spaces usually generates the gamut problem. One of most popular HSI color spaces is defined in [2]. However, the HSI gamut is larger than the RGB gamut, thus, it is inappropriate to perform image enhancement in this HSI color space. In [3], we derive the ideal HSI color space whose gamut is same as RGB color space as an extended form of the HSI color space of [2]. By using this HSI color space, it became possible to carry out the image processing without the gamut problem. We use the HSI color space of [3] in this paper.

The deterioration factors of the underwater image are “loss of white balance”, “decreased contrast”, “blurring of image”, “uneven lighting”, and “reduction of colorfulness”. In this paper, the white balance is adjusted by gray world method (GW method) [4] in the RGB color space. After that, we convert the white balance adjusted image into the ideal HSI color space [3]. For the intensity component, uneven lighting effect is removed by Retinex processing [5], and contrast improvement and sharpening of the image are performed by the differential gray-level histogram equalization (DHE) [6]. On the other hand, the colorfulness is improved by applying the histogram equalization (HE) the saturation component (S). In the proposed method, we realized the image quality improvement of underwater images.

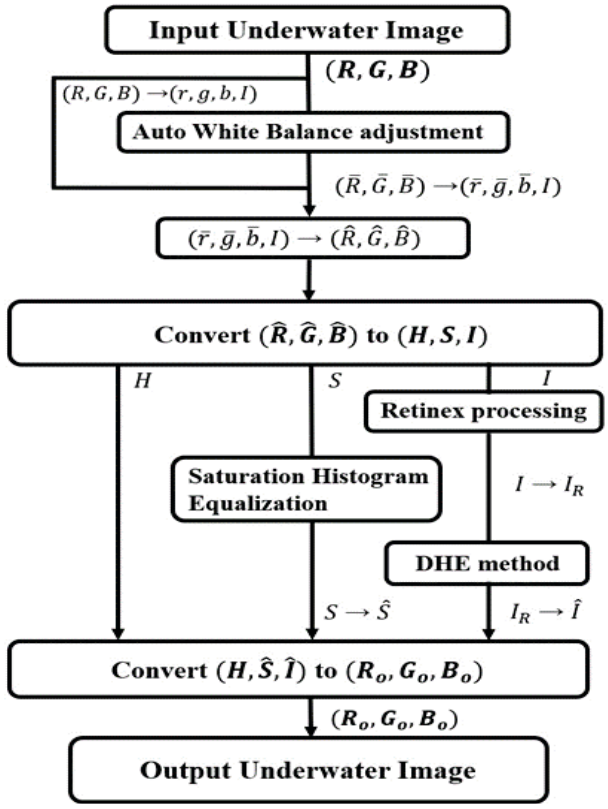


Fig. 1. Proposed method

II. The Proposed Method

To improve each degradation factor of the underwater image, we apply each processing to the most suitable color component for each deterioration factor. The block diagram of the proposed method is shown in Fig. 1.

A. Processing in RGB color space

By attenuating the red component, the underwater image is dominated by a single color such as blue or green. Therefore, the white balance of the underwater image is lost. It is very important to adjust the white balance to improve the quality of the image. In the proposed method, the white balance of the image is adjusted by using the GW method [4] for the RGB components of the color image signal. In order to preserve the intensity of the original image, the white balance adjusted RGB components are converted to chromaticity r, g, and b, and synthesized with the intensity component of the original image.

B. Processing in Ideal HSI color space

When photographing in the underwater circumstance, brightness unevenness occurs in the photographed underwater image. In addition, the underwater image causes blurring, low contrast, and low colorfulness. In order to improve these deterioration factors, the white balance adjusted underwater image convert to the ideal HSI color space [3]. Then, we process independently the intensity (I) and saturation (S).

First, we will explain the processing for intensity. We apply Retinex processing [5] to remove lighting unevenness. The Retinex processing calculates the local brightness (illumination light) average of intensity. By dividing the processed point of image to be processed by local average of the processed point eliminates uneven brightness condition. Since the DHE [6] equalizes the differential gray-level histogram including edge information, it has the image sharpening capability in addition to the contrast enhancement ability. The DHE is applied to the intensity component after the Retinex processing.

The decrease in colorfulness of the color image is due to the saturation distribution being biased towards lower saturation values. Thus, the improvement of the colorfulness is achieved by the equalization of the saturation distribution. Therefore, we apply the HE to saturation.

III. Experimental Results

We show the image of each processing result and clarify the effectiveness of the proposed method.

Fig. 2(a) is original image, and Fig. 2(b) is the result of the white balance adjustment method (i.e. GW method). The original image is an image in which the green component is strong and the white balance is broken. From Fig. 2(b), white balance adjustment was made by the GW method. Fig. 2(c) is the result of Retinex processing in which the brightness unevenness is removed from intensity component. Fig. 2(d) shows the result of Retinex processing and DHE method for intensity component. Fig. 2(e) is a processing result obtained by combining the result of the intensity component of Fig. 2(d) and the result of applying the HE method to the saturation component.

As a result of the proposed method, the blurring, contrast and colorfulness are improved, and the brightness unevenness of the image is removed. High quality improvement of underwater image is realized by the proposed method.

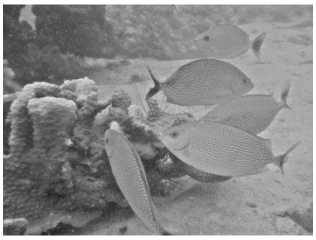
IV. Conclusion

The image quality of the underwater image was improved by applying appropriate processing to each deterioration factor in the ideal HSI color space.

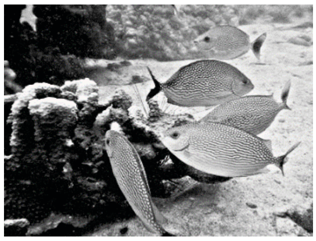
The white balance of the image was adjusted by using the GW method for the RGB component of the image. After that,

(a) Original

(b) GW method



(c) Retinex for Intensity



(d) (c) + DHE for Intensity

(e) Proposed method

Fig. 2. The output of proposed method

the RGB components was converted to the ideal HSI color space, and the intensity (I) and the saturation (S) were processed independently. Contrast enhancement and sharpening were performed on the intensity component by using the DHE method. Moreover, Retinex processing for intensity removed the brightness unevenness of the image. For the saturation component, the saturation histogram equalization processing improves the image colorfulness. We showed the result of images and clarified the effect of the image quality improvement of the proposed method.

References

[1] J.F. Anthoni, “Water and light in underwater photography,” http://www.seafriends.org.nz/phgraph/water.htm, June 2014.

[2] R. Gonzalez and R. Woods, *Digital Image Processing, 3rd ed*. Englewood Cliffs. NJ, USA: Prentice-Hall, 2007.

[3] M. Kamiyama and A. Taguchi, “HSI color space with same gamut of RGB color space,” *IEICE Trans. Fundamentals*, vol.E100-A, no.1, pp.341-344, January 2017.

[4] G. Buchsbaum “A spatial processor model for object colour perception,” *J. Franklin Inst*., 310, (1), pp. 1–26, 1980.

[5] D. J. Jobson, Z. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE Trans. Image processing*, vol.6, pp. 965 – 976, 1997.

[6] K. Murahira, A.Taguchi, “A novel contrast enhancement method using differential gray-levels histogram”, *Proc. 2011 International Symposium on Intelligent Signal Processing and Communications Systems* (*ISPACS*) , pp. 1–5, 2011.

Generalized Differential Gray-level Histogram Equalization

Hideaki Tanaka † and Akira Taguchi †

† Department of Computer Science, Tokyo City University, Tokyo Japan
ataguchi@tcu.ac.jp

***Abstract*—Histogram equalization (HE) is a simple and effective method for contrast enhancement as it can automatically define the gray-level transformation function based on the distribution of gray-level included in the image. HE fails to produce satisfactory results for broad range of low-contrast images because the HE does not use a spatial feature included in the input image. The differential gray-level histogram which is contained edge information of the input image, were defined. Furthermore, the differential gray-level histogram equalization (DHE) has been proposed. The DHE shows better enhancement results compared to the HE results for many kinds of images. In this paper, we propose a generalized DHE (GDHE) method. In GDHE, histograms are created using powers of gradients. If the power is set as 0, GHE is equivalent to HE, and if the power is set as 1, GHE is equivalent to DHE. That is, GDHE includes HE and DHE. GHE can preserve the mean brightness of the input image perfectly by setting the power appropriately and shows good contrast enhancement results at the same time.**

Keywords- contrast enhancement, histogram, differential gray-level histogram

I. INTRODUCTION

Contrast enhancement has an important role in the image processing applications [1]. The objective of this method is to make an image clearly recognized for a specific application. One of the most popular contrast enhancement methods is the histogram equalization (HE). The dynamical range of an image is increased by the HE. However, the HE often fail in producing satisfactory results for broad range of low-contrast images, such as images characterized by the fact that the amplitudes their histogram components are very high at one or several locations on the grayscale, while they are very small, but not zero, in the rest of the grayscale. The high amplitude of the histogram components corresponding to the image background also often prevents the use of the HE method, which could cause a washed-out effect on the appearance of the output image and/or amplify the background noise.

In order to overcome these problems, a lot of advanced HE-based contrast enhancement methods have been developed [2]-[5], such as bi-histogram equalization, dynamic histogram equalization, and so on. These methods were proposed in order to preserve the mean brightness of the input image.

The output image of the conventional HE-based methods cannot adapt to a spatial feature included in the input image, since the histogram shows the only information of the distribution of gray-level of the input image. On the other hand, a differential gray-level histogram (DH) was defined [6]. DH is created using gradient values which is typical special feature of the image. The DH equalization (DHE) shows better results compared with the HE, especially for broad range of low-contrast images.

In this paper, we propose a generalized DHE (GDHE). In GDHE, histograms are created using the power of gradient value. If the power is set as 0, GHE is equivalent to HE, and if the power is set as 1, GHE is equivalent to DHE. That is, GDHE is a generalized gradient histogram equalization including HE and DHE. Furthermore, the great advantage of GDHE is that it is possible to preserve the mean brightness of the input image by appropriately giving the power to each image. In other words, it is a method that preserves the mean brightness of the input image and is capable of good contrast enhancement. The effectiveness of GDHE is clarified from both subjective and objective evaluations.

II. GENERALIZED DIFFERENTIAL GRAY-LEVELS HISTOGRAM EQUALIZATIO

Consider an input image $f(i,j)$, which is a the total number of N pixels with gray-level in the range [0, L-1]. The gradient value of the gray level of the input image is given by follows:

$$d(i,j) = \mathrm{int}\{\sqrt{d_H(i,j)^2 + d_V(i,j)^2}\} \tag{1}$$

where

$$\begin{aligned} d_H(i,j) = &\{f(i+1,j+1) + 2\cdot f(i+1,j) + f(i+1,j-1)\} \\ &- \{f(i-1,j+1) + 2\cdot f(i-1,j) + f(i-1,j-1)\} \end{aligned}$$

$$\begin{aligned} d_V(i,j) = &\{f(i+1,j+1) + 2\cdot f(i,j+1) + f(i-1,j+1)\} \\ &- \{f(i+1,j-1) + 2\cdot f(i,j-1) + f(i-1,j-1)\} \end{aligned}$$

$\mathrm{int}\{\}$ in Eq.(1) represents the integer transform processing.

The histogram created by the power of α of the gradient value of gray-level is given by the following equation.

$$h^{\alpha}(r) = \sum_{(i,j)\in D_r} \{d(i,j)\}^{\alpha} \tag{2}$$

where D_r is a region composed of pixels whose value is r. If α is set as 0, $\{d(i,j)\}^{\alpha}$=1. Thus, $h^0(r)$ is the normal histogram. On the other hand, when α = 1, $h^1(r)$ is the DH. That is, the gradient histogram which is defined by Eq.(2) contains the normal histogram and DH.

GDHE will map an input gray-level *r* into an output gray level *s* using the following transformation function *T*(*r*).

$$s = T(r) = (L-1)\cdot c(r) \tag{3}$$

where

$$c(r) = \sum_{k=0}^{r} h^{\alpha}(k) \Big/ \sum_{k=0}^{L-1} h^{\alpha}(k)$$

GDHE can preserve the average value of input image strictly by properly determining α.

III. Experimental Results

Four 8-bit images of size 256x256, which are shown in Fig.1, are used for simulation. We would like to define the detail region and the background region. Let the variance of the 5x5 window centered on the pixel of interest be the local variance of that pixel. The detail region consists of pixels whose local variance is greater than 100. On the other hand, the background region consists of pixels whose local variance smaller than 100. We calculate the variance of the detail region and the background region as the detail variance (DV) and the background variance (BV), respectively. We use the ratio of DV of original image to enhancement image (DVR) for evaluation. In the same way, we define the ratio of background variance (BVR), and the ratio of mean brightness (MBR). MBR=1 means that the average value of the input image is preserved. The BVR and the DVR indicate the degree of enhancement. The larger the DVR, the better for the purpose of image enhancement.

Evaluation results are shown in Table 1. It can be seen that the mean brightness of the enhancement image obtained by GDHE can be made the same as the mean brightness of the original image by setting α appropriately. Although HE and DHE outperform GDHE in terms of the DVR, the mean brightness of these images are significantly different from the original image. It is a necessary condition that the tone before and after processing is not changed in image enhancement processing. From the point of view, GDHE is a very good method. Furthermore, in GDHE, DVR for all images greater than 2.3, thus, GDHE shows high enhancement ability.

Figure 2 shows the enhancement results of “Airplane”. In HE’s result, the tone of the image has changed. We can not read the character of the tail of the plane. In addition, the background is overemphasized. The enhancement results of DHE and GDHE are both excellent. In particular, GDHE sharpens the edges of the image and achieves good contrast improvement.

IV. Conclusions

In this paper, we proposed GDHE which includes HE and DHE. GDHE can perfectly preserve the mean brightness of the original image and can realize a good contrast improvement independent of the kind of image.

The future work is to improve this method so that it can be applied to video signals

(a) Airplane (b) Boat (c) Girl (d) Barbara

Fig.1 Test images

(a) HE (b) DHE (c) GDHE

Fig.2 Enhancement results (Airplane)

Table 1 Evaluation results of HE, DHE, and GDHE

		HE	DHE	GDHE(α)
Airplane	MBR	0.718	0.955	**1.000(α =1.5)**
	BVR	16.67	4.556	**3.161**
	DVR	2.376	2.331	**2.398**
Boat	MBR	0.900	1.148	**1.000(α =0.36)**
	BVR	22.07	**6.131**	14.11
	DVR	**4.483**	3.595	3.93
Girl	MBR	1.786	1.443	**1.000(α =6.9)**
	BVR	4.671	3.285	**2.476**
	DVR	**2.880**	2.746	2.420
Barbara	MBR	1.111	1.084	**1.000(α =4.6)**
	BVR	**2.249**	4.164	5.062
	DVR	2.349	2.892	**2.908**

References

[1] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, 2nd ed., New Jersey: Printice Hall, 2002.

[2] Y. T. Kim, “Contrast enhancement using brightness preserving bi- histogram equalization,” *IEEE Trans. Consummer Electron*., vol.43, no.1, pp.1-8, Feb.1997.

[3] Y. Wan, Q. Chen, and B.-M. Zhang, “Image enhacement based on equal area dualistic sub-image hustogram equalization method,” *IEEE Trans. Consummer Electron*., vol.45, no.1, pp.68-75, Feb.1999.

[4] S.-D. Chen, and A. R. Ramli, “Minumum mean brightness error bi-histogram equalization in contrast enhancement”, *IEEE Trans. Consummer Electron*., vol.49, no.4, pp.1310-1319, Nov.2003.

[5] M. Abdullah-Al-Wadud, Md. Hasanul Kabir, M. Ali Akber Dewan, and Okam Chae, “A dynamic histigram equalization for image contrast enhancement,” *IEEE Trans. Consummer Electron*., vol.53, no.2, pp.593-600, May.2007.

[6] F. Saitoh, “Image contrast improvement based on differential gray-levels histogram,” *IEEJ Trans. EIS*, vol.126, no.2, pp.228-236, Feb. 2006 (in Japanese).

Color Image Enhancement by Using Hue-Saturation Gradient

Hiromu Endo † and Akira Taguchi †

† Department of Computer Science, Tokyo City University, Tokyo Japan
ataguchi@tcu.ac.jp

***Abstract*— It is necessary to preserve the hue component of color image in color image processing. Although the hue component can not be processed, it can be used for color image processing. However, the hue component has not been used for color image processing until now. It is already known that saturation components work effectively in color image enhancement. Therefore, in this paper, we define the hue-saturation gradient which is a gradient combining the saturation gradient and the hue gradient. The differential hue-saturation histogram equalization with variable enhancement degree (DHSHEwVED) is proposed. The proposed method can emphasize both intensity and saturation. The proposed method has two parameters, which can be used to adjust the contrast and colorfulness independently. We shows the effectiveness of proposed method from image processing results.**

Keywords- color image enhancement, hue-saturation gradient, differential histogram equalization

I. Introduction

For the color image processing, hue preserving is required. The reason is that if hue is changed then the color is changed, humans feel the image is degraded. Therefore, many processing methods are given as processing in transformed color space such as HSI, YIQ, HSV, etc. Transforming from RGB color space to another color space and image processing in these spaces usually generates the gamut problem. In Ref.[1], we derive an HSI color space whose gamut is exactly same as RGB color space. The HSI color space is an extension of the HSI color space of Ref.[2] which is the most common HSI model. By using this HSI color space [1], it became possible to perform image enhancement without the gamut problem. In this paper, we propose a color image enhancement method for processing in HSI color space of Ref.[1].

It has been made clear that saturation information is effective for color image processing [3], however, the hue information has never been used for the color image processing. In the case of a color image, the change portion of hue is also regarded as the edge. Therefore, it is considered that the gradient of the hue component is effective for the intensity processing. Furthermore, since the gradient of hue is often sufficiently included even in the dark area of the color image where the amplitude of intensity is small, it is expected that the gradient of hue will effectively work for the contrast enhancement of the dark area of the color image. In this paper, the hue-saturation gradient is newly defined. By using the hue-saturation gradient, the differential hue-saturation histogram for intensity (DHSHI) and the differential hue-saturation histogram for saturation (DHSHS) are defined. Then, the DHSHI equalization with variable enhancement degree (DHSHIEwVED) and DHSHS equalization with variable enhancement degree (DHSHSEwVED) are proposed in order to emphasize the intensity and saturation, respectively. Finally, the differential hue-saturation histogram equalization with variable enhancement degree (DHSHEwVED) is proposed by combining DHSHIEwVED and DHSHSEwVED.

Both two enhancement methods have one free parameter. Thus, intensity enhancement and saturation enhancement are performed simultaneously, and contrast and colorfulness can be controlled independently by using two parameters of the DHSHEwVED.

II. Hue-Saturation Gradient

Regions of low contrast/colorfulness in color images are often shadow areas or saturated areas. In these regions, the gradient component of intensity is poor. On the other hand, the gradient of hue/saturation is often large in these areas. Therefore, we will define a hue-saturation gradient.

The RGB gradient is defined by

$$\Delta_{RGB}(i,j) = \sqrt{\{\Delta_R(i,j)\}^2 + \{\Delta_G(i,j)\}^2 + \{\Delta_B(i,j)\}^2} \quad (1)$$

where

$$\Delta_k(i,j) = \sqrt{\{\Delta_k^H(i,j)\}^2 + \{\Delta_k^V(i,j)\}^2}$$

$$\Delta_k^H(i,j) = k(i-1.j-1) + 2k(i.j-1) + k(i+1,j-1) - k(i-1,j+1) - 2k(i,j+1) - k(i+1,j+1)$$

$$\Delta_k^V(i,j) = k(i-1.j-1) + 2k(i-1.j) + k(i-1,j+1) - k(i+1,j-1) - 2k(i,+1j) - k(i+1,j+1)$$

(k=R, G, B)

$R(i,j)$, $G(i,j)$ and $B(i,j)$ represent the R, G and B components of the color image signal, respectively.

Let the intensity component of the color image be $I(i, j)$ [1], the intensity gradient be defined as follows.

$$\Delta_I(i,j) = \sqrt{\{\Delta_I^H(i,j)\}^2 + \{\Delta_I^V(i,j)\}^2} \quad (2)$$

where

$$\Delta_I^H(i,j) = I(i-1.j-1) + 2I(i.j-1) + I(i+1,j-1) \\ - I(i-1,j+1) - 2I(i,j+1) - I(i+1,j+1)$$

$$\Delta_I^V(i,j) = I(i-1.j-1) + 2I(i-1.j) + I(i-1,j+1) \\ - I(i+1,j-1) - 2I(i,+1j) - I(i+1,j+1)$$

To normalize both the RGB gradient and the intensity gradient, we apply following formula.

$$\overline{\Delta}_{RGB}(i,j) = \frac{\Delta_{RGB}(i,j) - \min\{\Delta_{RGB}(i,j)\}}{\max\{\Delta_{RGB}(i,j)\} - \min\{\Delta_{RGB}(i,j)\}} \quad (3)$$

$$\overline{\Delta}_I(i,j) = \frac{\Delta_I(i,j) - \min\{\Delta_I(i,j)\}}{\max\{\Delta_I(i,j)\} - \min\{\Delta_I(i,j)\}} \quad (4)$$

And we get the hue-saturation gradient $D_{HS}(i,j)$ by calculating the absolute differential value between $\overline{\Delta}_{RGB}(i,j)$ and $\overline{\Delta}_I(i,j)$.

$$D_{HS}(i,j) = \left|\overline{\Delta}_{RGB}(i,j) - \overline{\Delta}_I(i,j)\right| \quad (5)$$

III. DHSHEwVED

We propose a enhancement method that can change the degree of enhancement of both intensity and saturation using the hue-saturation gradient. Improve the contrast by processing the intensity and improve the colorfulness by processing the saturation. The improvement of contrast and colorfulness is done independently. Therefore, it is easy to obtain an color image of contrast and colorfulness desired by the user. A block diagram of the proposed method is shown in Fig. 1.

The transformation function of DHSHIEwVED is given as

$$s = (1-\alpha)\cdot r + \alpha \cdot f_{HS}^I(r) \quad (6)$$

where $f_{HS}^I(r)$

$$f_{HS}^I(m) = H_{HS}^I(m) / H_{HS}^I(M-1) \quad (7)$$

where M indicates the number of amplitude levels of intensity and

$$H_{HS}^I(m) = \sum_{r=o}^{m} h_{HS}^I(r) \quad (8)$$

$$h_{HS}^I(r) = \sum_{(i,j)\in D_r^{I(i,j)}} D_{HS}(i,j) \quad (9)$$

$D_r^{I(i,j)}$ is a set of pixels with the $I(i,j) = r$.

In DHSHSEwVED, an input saturation value q is mapped into an output saturation value t as follow.

$$t = (1-\beta)\cdot q + \beta \cdot f_{HS}^S(q) \quad (10)$$

where

$$f_{HS}^S(q) = H_{HS}^S(q) / H_{HS}^S(L-1) \quad (11)$$

$$H_{HS}^S(q) = \sum_{i=o}^{q} h_{HS}^S(i) \quad (12)$$

$$h_{HS}^S(q) = \sum_{(i,j)\in D_q^{S(i,j)}} D_{HS}(i,j) \quad (13)$$

L indicates the number of amplitude levels of saturation. $S(i,j)$ shows the saturation value of point (i,j) [1]. $D_q^{S(i,j)}$ is a set of pixels with $S(i,j) = q$.

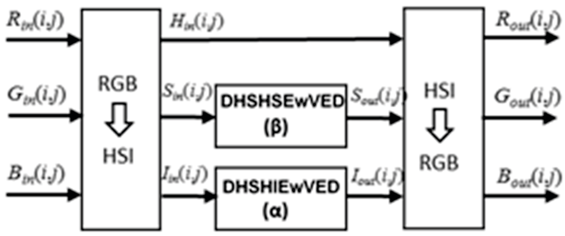


Fig.1 Proposed enhancement method for color images

(α, β)

(0, 0)	(0.5, 0)	(1, 0)
(0, 0.5)	(0.5, 0.5)	(1, 0.5)
(0, 1)	(0.5, 1)	(1,1)

Fig.2 Enhancement results for various (α, β)

IV. Experimental Results & Conclusions

We show the enhancement results in Fig.2. The enhancement image with α=β= 0 is the original image. The proposed method gives various results. It can be seen that this method has high emphasis ability.

Three images in the row direction, the colorfulness is almost constant and the contrast is changes. On the other hand, three images in the column direction, the contrast is almost constant and only the colorfulness changes. It can be seen that the contrast can be varied with the parameter α and the colorfulness can be varied with the parameter β. It is possible to adjust contrast and colorfulness independently. Since the contrast and colorfulness can be adjusted independently, the user can easily obtain the desired image.

References

[1] M. Kamiyama and A. Taguchi, "HSI color space with same gamut of RGB color space," *IEICE Trans. Fundamentals*, vol.E100-A, no.1, pp.341-344, January 2017.

[2] R. Gonzalez and R. Woods, *Digital Image Processing, 3rd ed.* Englewood Cliffs. NJ, USA: Prentice-Hall, 2007.

[3] R.N. Strickland, C.S. Kim and W.F. McDonnell, "Digital color image enhancement based on the saturation component, "*Optical Engineering*, vol.26, no.7, pp.609-616, July 1987.

Road Boundary Detection for Straight Lane Lines Using Automatic Inverse Perspective Mapping

Fabien Rakotondrajao
Computer Engineering Department
ESIEA Graduate Engineering school
Laval, France
rakotondrajao@et.esiea.fr

Kharittha Jangsamsi
Computer Engineering Department
Faculty of Engineering, KMUTT
Bangkok, Thailand
kharittha.jan@mail.kmutt.ac.th

***Abstract*—This paper addresses the difficulty of detecting and/or verifying road boundary markings from road images. The 2-D road images usually contain the perspective effect/distortion because of the acquisition process resulting in an inaccuracy of the lane detection. We thus propose in this paper an approach to automatically select four points for performing the Inverse Perspective Mapping (IPM) in order to diminish such effect/distortion. Our proposed method is efficient and very practical to implement with the existing lane detection method, as shown in the experimental results.**

Keywords—road boundary markings, perspective effect/distortion, lane detection, Inverse Perspective Mapping

I. INTRODUCTION

Advanced driver assistance systems (ADASs) are successively developed to assist drivers keeping their vehicle to stay in the lane and avoiding accidents. ADASs consist of several systems performed together, i.e. lane following systems (LFSs), lane keeping assistance systems (LKASs), lane departure warning systems (LDWSs), lateral control systems (LCSs), intelligent cruise control systems (ICCSs) and collision warning systems (CWSs). One of the critical components of ADASs, e.g. using in LKASs and LDWSs [1-2], is lane detection (LD). It gives meaningful information about the surroundings, including the road surface, road shape and objects. In a digital image, two stripes on a road surface that bound the road lane called left and right lane markings are used for LD. However, detecting lane is a difficult problem since various factors can affect its accuracy and robustness, for examples degradation of lane markings, presence of shadows projected by buildings, other vehicles or trees, lighting conditions, weather, traffic conditions, road geometries, etc.

A large volume of literatures attempts to reduce the effect of such factors on detecting the lane markings. Most of them are commonly proposed using vision-based methods owning to the low cost of camera devices and extensive background knowledge of image processing [3]. In the vision-based methods for LD, the 2-D road image captured from a camera which is mounted behind the front windshield to retrieve road information is examined by using image processing techniques. Vision-based methods can be categorized into two groups: model-based methods and feature-based methods [4]. To describe the lane structure, the model-based methods generally apply mathematical models with parameters [5] such as a linear model, a parabolic model, or various kinds of spline models. While the feature-based methods analyze the road images and detect the characteristics of lane markings [6] including color, edge, texture and geometry. Although the model-based methods are more robust against noise than the feature-based methods, they require some prior-known geometric parameters and heavy computation [7]. Hence, the feature-based methods are considered in this paper because they have better computational efficiency. Many existing feature-based methods have been proposed. In general, each feature of lane markings is preprocessed using a specific method. For examples, in 2002, Y. Otsuka *et al.* [7] utilized only the edge image obtained from extracting areas with differences of brightness. Each edge pixel contains its position and angle. The angles of all edge pixels were used to create the histogram for finding the left and right lane markings. In 2010, an ant colony optimization was developed by P. M. Daigavane *et al.* [8] for linking disjointed edges in the edge images. Then, the Hough transform was applied to extract left and right lane markings. In 2017, D. Clark [9] developed self-driving car program to identify road boundaries in images and then applied to video by breaking the video down into frames. The grayscale image was used instead of color image for feature detection, i.e. edge, to reduce the amount of data and complexity and Hough Transform was also applied to find where pixels form lines.

In this paper, we therefore propose a new approach to automatically select four points for performing the Inverse Perspective Mapping (IPM) in fitting step of lane detection. In general, the points used for the Inverse Perspective Mapping is fixed or selected manually. Sometimes they do not work for all road images. The results can hence affect the accuracy of lane detection.

II. PROPOSED METHOD

Generally, most of feature-based methods used for lane detection consist of three steps: image preprocessing, feature detection and fitting.

In image preprocessing step, an input image is prepared for the subsequent steps. We use a combination of color to generate a binary image, i.e. our method combines histogram equalization followed by thresholding and thresholding on V color channel of HSV color space for detecting the white lines and yellow lines, respectively. Both results from two sub-steps are combined using a logical operator, *OR* to obtain the binary image highlighted the actual line lanes.

In feature detection step, this step can be carried out in two sub-steps: feature selection and extraction. In this paper, edges are the selected features since we focus on the structured roads that the lane markings usually have clear edges and relatively high intensities. The Sobel edge detector is performed to extract the edge information from the binary image obtained from the previous step to acquire an estimation of the gradients of the lines. The closing morphological operation is then applied to fill the gaps in the edge images.

In fitting step, since the perspective effect/distortion is always occurred in the input image from the acquisition process, the Inverse Perspective Mapping (IPM) is applied to reduce this effect/distortion. The IPM will transform a perspective view to a view from the sky called bird's-eye view. In order to perform the IPM, four points of the result image from the previous step and four points of transformed space image are remapped. In this paper, we thus propose two following sub-steps to select both source and destination points automatically.

A. Finding Two Bottom Points

We perform an exhaustive search on the result image from the previous step. This search is implemented to find white pixels initiating from the bottom and moving to the upper side of the image. The search area will be starting from (150, h - 60) to (350, h - 20) for bottom left point and from (900, h - 60) to (1250, h - 20) for bottom right point, where h is the height of the input image. Our method will locate the positions of the continuous white pixels. The final positions of bottom left, and right points are then averaged from these positions.

B. Finding Two Top Points

This sub-step is not relevant to the previous sub-step. We perform an exhaustive search on the result image from the feature detection step. However, two-bottom left, and right points from the previous sub-step are used as the starting points of this sub-step. For finding top left point, two small windows are generated according to the bottom left points which is the reference position. Note that the size of windows is 15×20 pixels. The number of white pixels is then averaged for both windows. If the average value of the left window is more than the average value of the right window, the reference position will be changed to the center of the left window. This stage is repeated certain times.

Next stage, we create a sliding window starting from the latest reference position toward the upper side of the image to decide which pixel belong to the left lane marking. The neighbor pixels of the considered position are verified whether pixels are white or black. If some neighbor pixels on the left window is black, the reference position will be moved to right 5 pixels. Otherwise, the reference position will be moved up 3 pixels. This stage is repeated certain times. The moving of the reference position is the same as moving of snake. For finding the top right point, we do the same with finding the top left point.

After performing above two sub-steps, the IPM is applied to rectify the result image from the feature detection step. The IPM is defined by

$$[X_d \quad Y_d \quad Z_d] = [X_s \quad Y_s \quad 1] * \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & 1 \end{bmatrix} \tag{1}$$

where *a*-*h* are eight unknown parameters, which can be determined if four-point pairs are known. After determining, we get *Tform* structure containing the information for mapping as the output. Finally, pixels of the lane markings are detected and fitted their positions with a polynomial.

III. Experimental Setting and Results

The proposed method used Python language for development and implementation in the Atom. Our method was executed on compilation environment via a personal computer with an Intel i7-3537U 2.00 GHz CPU, and 8GB of RAM. The operating system of the experimental computer was Windows 10. In the experiments, twelve 2-D color road images were used as the test images. All test images used in this experiment are open source due to the lack of data available in the existing literature. However, all test images are the real road images captured using a camera mounted behind the front windshield including both straight and curve lane lines. The size of test images is 1280 × 720. Moreover, there are buildings, other vehicles, tree and shadows.

The experimental results showed that the detection accuracy of our proposed method was 91.67 % more than D. Clark's method [9], i.e. only 62.35 %. Fig. 1 illustrates the example of our result.

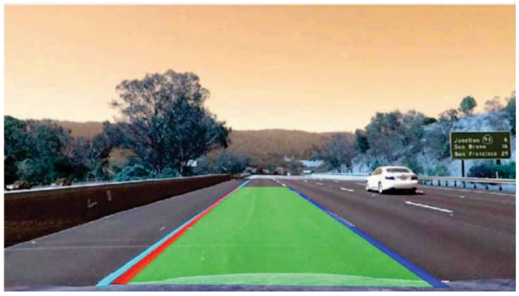


Fig. 1. The example of result image from our proposed method.

IV. Conclusion

In this paper, we have presented an approach to automatically select points for performing the Inverse Perspective Mapping in fitting step of lane detection to improve the accuracy of lane detection. The experimental results showed some improvements. However, the Hough transform based on the straight lane lines do not work well for the curve lane lines. Using a higher degree curve will be useful on the curve lane line.

References

[1] C. M. Martinez, X. Hu, D. Cao, E. Velenis; B. Gao; and M. Wellers, "Energy management in plug-in hybrid electric vehicles: recent progress and a connected vehicles perspective," IEEE Trans. Veh. Technol., vol. 66, no. 6, pp. 4534–4549, 2016.

[2] Y. Huang, A. Khajepour, T. Zhu, and H. Ding, "A supervisory energy-saving controller for a novel anti-idling system of service vehicles", IEEE/ASME Trans. Mechatronics, vol. 22, no. 2, pp. 1037–1046, 2017.

[3] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: A survey," Mach. Vis. Appl., vol. 25, no. 3, pp. 727–745, 2014.

[4] Y. Wang, N. Dahnoun, and A. Achim, "A novel system for robust lane detection and tracking," Signal Process., vol. 92, no. 2, pp. 319–334, 2012.

[5] M. Aly, "Real time detection of lane markers in urban streets," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 7–12.

[6] T. Y. Sun, S. J. Tsai, and V. Chan, "HSI color model based lanemarking detection," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Toronto, ON, Canada, Sep. 2006, pp. 1168–1172.

[7] Y. Otsuka, S. Muramatsu, H. Takenaga, Y. Kobayashi, and T. Monj, "Multitype lane markers recognition using local edge direction," in *Proc. IEEE Intell. Vehicles Symp.*, vol. 2. Versailles, France, Jun. 2002, pp. 604–609.

[8] C. Rose, J. Britt, J. Allen and D. Bevly, "An integrated vehicle navigation system utilizing lane-detection and lateral position estimation systems in difficult environments for GPS," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 6, pp. 2615–2629, Dec. 2014.

[9] D. Clark, "Advanced lane-detection for self-driving Cars," The Article. Becoming Human: Artificial Intelligent Magazine, 2017.

Semantic Traffic Light Understanding for Visually Impaired Pedestrian

Chanagan Pongseesai
Department of Electronic and Telecommunication Engineering
King Mongkut's University of Technology Thonburi
Bangkok, Thailand
Annnew.970@gmail.com

Kosin Chamnongthai
Department of Electronic and Telecommunication Engineering
King Mongkut's University of Technology Thonburi
Bangkok, Thailand
kosin.cha@kmutt.ac.th

***Abstract*—A visually impaired pedestrian normally requires information of walking/stop signs with time to decide whether it is available to cross a road on crosswalk. Since traffic lights are installed in verity formats such as, man figure, color and numerical characters, it is complicate to simply understand the meaning. This paper proposes a method of semantic traffic light understand for a visually impaired pedestrian. In the method, a picture of traffic light is initially assumed to extract and type of traffic light are first classified into man figure, color and numerical characters by some classification tools such as signature. In man figure case, color and man gesture are classified and both information are decided base on Fuzzy logic. In the case of numerical characters, they are first enhanced and then classified by a classifier. The experiments have been performed with 200 samples of pedestrian traffic lights and results reveal acceptable accuracy.**

***Keywords*— semantic Understanding; traffic light classification; visually impaired pedestrian**

I. INTRODUCTION

According to a report of WHO (World Health Organization) in 2018 [1], number of vision impaired people is increased to 1.3 billion approximately around the world which are 188.5 million distance-vision people, 217 million mild-vision people, 36 million blind people, and 826 million near-vision people. To live in the same environment with normal people, although they are lack with normal vision function, they have to be able to behave in the same manner, and survive by the same conditions, especially when they cross a road at a crosswalk. The statistic data reveal many people die by accidents at the crosswalks [2]. To help vision impaired people to survive in the environment with normal people, a pedestrian-traffic-light recognition system is urgently developed for assisting vision impaired people.

To develop the mentioned system, some researchers have tried to develop the system. For instance, Sergio Mascetti et al [3] proposed a robust setup for image capture that makes it possible to acquire clearly visible traffic light images regardless of daylight variability due to time and weather. The objective is for recognition of traffic light, and it worked well. However, traffic lights for pedestrian normally are not only color sign consisting of red and green, but also numerical characters expressing time of each color and human sign. These kinds of signs semantically give important information for pedestrian which vision impaired people really need to understand for considering to cross the road. The authors therefore try to first recognize meaning of each type of sign, and utilize Fuzzy logic to conclude all signs into information for crosswalk crossing decision for vision-impaired pedestrian.

II. PROPOSED METHOD

As an example of typical signs on a traffic light for pedestrian to cross a crosswalk as shown in Fig. 1, numerical characters of "50" and a sign of standing man with red color, as shown in Fig. 1 (a) reveals a message for pedestrian to stop and wait for 50 seconds on a footpath side, while numerical characters of "15" and a sign of walking man with green color, as shown in Fig. 1 (b) mean pedestrian has 15 second time to cross a crosswalk.

a) Red traffic light b) Green traffic light

Fig. 1 Traffic light

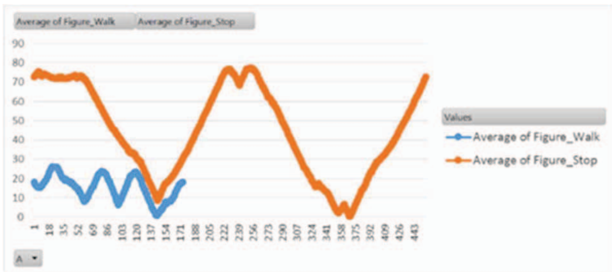


Fig. 2 Analysis of traffic light images by Signature

By analyze the images of traffic light by some feature extraction tools, features such as Signature and so on can be extracted, and it seems to possibly classify the images of traffic light into their types such as numerical number, human gestures, and colors. Due to possible conflict among signs, those extracted features are needed to decide as walkable or stop with time. These features should be fed to fuzzy logic for decision on ambiguous factors. The flowchart of processes can be expressed as shown in Fig. 3. An image of traffic light is input and classified into their types at type classification process, and each type of feature is then detected their boundaries at the process of boundary detection. All boundaries are decided as walkable or stop at the process of decision using Fuzzy logic, and the system finally outputs results.

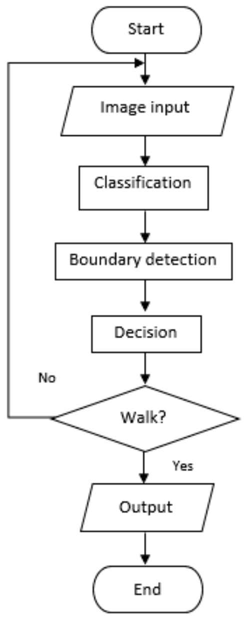


Fig. 3 Flowchart of software system

In application of Fuzzy logic, for instance, a couple of factors of numerical character, color and human gesture sign, which are learned levels of danger and safety, may be converted into a triangle model of the Fuzzy logic.

Fuzzy rules for decision is required to determine in advance as shown as follows.

1. If (Number is Danger) and (Color is Red) and (Sign is Stop) then output is Stop
2. If (Number is Danger) and (Color is Red) and (Sign is walk) then output is Stop
3. If (Number is Danger) and (Color is Green) and (Sign is Stop) then output is Stop
4. If (Number is Danger) and (Color is Green) and (Sign is walk) then output is Stop
5. If (Number is Safe) and (Color is Red) and (Sign is Stop) then output is Stop
6. If (Number is Safe) and (Color is Red) and (Sign is walk) then output is Stop
7. If (Number is Safe) and (Color is Green) and (Sign is Stop) then output is Stop
8. If (Number is Safe) and (Color is Green) and (Sign is walk) then output is Cross

When factors are applied to the mentioned Fuzzy rules, the following parameters are calculated, and center of gravity (COG) is then determined for final decision.

As an example, factors of numerical character, color, and human gesture sign are assumed. These factors are applied into the determined Fuzzy rules, and equations, and results are obtained by equations as follows.

These values are indicated by the colored area. The center of gravity (COG) of the colored areas is calculated using the following equations

$$\mathrm{COG}= \frac{\sum_{x=m}^{n} \mu(A)x\ x}{\sum_{x=m}^{n} \mu(A)x} \tag{1}$$

$$\mathrm{COG}= \frac{(Z0 \text{ x } S3)+(Z1 \text{ x } S2)}{S3+S2}$$

$$COG = \frac{(0.3\ x\ 0.3) + (0.6\ x\ 0.2)}{0.3 + 0.2}$$

$$COG = 0.42$$

In this case, the COG of 0.42 is obtained and finally plotted in Fig. 10, which indicates the probability.

III. Experimental Results

To evaluate the performance of the proposed method, experiments in eight cases have been performed, and results are shown in Table I.

Table I Recognition and decision results

Cases	**Numerical character (sec)**	**Color**	**Human gesture sign**	**Decision**
1	5	Red	Stop	Stop
2	5	Red	Walk	Stop
3	5	Green	Stop	Stop
4	5	Green	walk	Stop
5	15	Red	Stop	Stop
6	15	Red	Walk	Stop
7	15	Green	Stop	stop
8	15	Green	walk	cross

IV CONCLUSION

To implement an automatic system of semantic traffic light understanding for visually impaired pedestrian, numerical characters, color, and a sign of human gesture have to be recognized and decided. This paper proposed a method to recognize those visual data, and apply fuzzy logic to make decision. The effectiveness of the proposed method have been confirmed, and real time process remains as future research work.

References

[1] World Health Organization. (2018, October 11). Blindness and vision impairment. Retrieved May 5, 2019, from https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

[2] Traffic Signals Standard and manual. Accessed: May 1, 2019. [Online]. http://www.otp.go.th/index.php/post/view?id=2098

[3] Mascetti, Sergio, et al. "Robust traffic lights detection on mobile devices for pedestrians with visual impairment." Computer Vision and Image Understanding 148 (2016): 123-135.

Design and Implementation of Data Collection and Driving Behaviour Analysis Based on SAE J1939

B.V.P. Prasad
Department of Electronic Engineering
Southern Taiwan University of Science and Technology
Tainan City, Taiwan
bvpprasad@stust.edu.tw

Jing-Jou Tang
Department of Electronic Engineering
Southern Taiwan University of Science and Technology
Tainan City, Taiwan
tjj@stust.edu.tw

Sheng-Jhu Luo
Department of Electronic Engineering
Southern Taiwan University of Science and Technology
Tainan City, Taiwan
4a037025@stust.edu.tw

***Abstract*— Safe driving is essential for drivers, passengers, pedestrians and fleet executives in today's complicated traffic setting. In this article, we develop a hardware module for heavy-duty (HD) vehicles integrated with SAE J1939 communication protocol. Based on this scheme, it is possible to access, and store associated data, such as velocity, RPM, and coolant temperature to monitor driver behaviour and vehicle safety.**

The J1939 Bridge, used in this paper, is jointly developed by STUST and SwiSys, Taiwan. Our proposed system is simulated with the Vector ECU (Electronic Control Unit) design simulator-CANoe to validate the in-vehicle CAN (Controller Area Network) bus system behaviour. The Renesas RL78 MCU is used to realize the SAE J1939 protocol high-level software stack. The raw information accessed will then be viewed and transmitted to an internal user interface LCD display device. For further integration with external communication devices, a Bluetooth interface is also given.

Keywords—SAE J1939, CAN Bus, Vector CANoe, OBD II, Emulator, Driving Behaviour Analysis, In-vehicle data.

I. Introduction

Robert Bosch GmbH first created the CAN standard (Control Area Networks) in 1983. It's a serial communication type. ISO 11898, which defines the physical and data link layers of the CAN protocol, specified the current standard. The physical layer defines a two-wire differential bus, i.e. CAN H and CAN L, terminated with a 120-ohm resistor at both ends and linked at each node to a transceiver. Logic 0 and Logic 1 data are defined respectively as dominant bit and recessive bit [1]. Heavy vehicles play a significant role in the global economy, while in many ways physically distinct from consumer vehicles, heavy vehicles are internally comparable in that they are consist of a distributed ECU system that communicates through a CAN based network.

SAE J1939 is the suggested method used to interact with heavy duty (HD) auto parts in communication and diagnostics. SAE J1939 is a high-level communications protocol, defines five of the seven layers of the OSI model, with CAN2.0B being used for the physical and data-link layers (using only the 29-bit "extended" identifier). The physical layer J1939/11 describes the electrical interface to the bus. The data link layer J1939/21 describes the rules for constructing a message, accessing the bus, and detecting transmission errors. The application layer J1939/71and J1939/73 defines the specific data contained within each message sent across the network. The information speed is defined as 250 bit/s under J1939/11. SAE J1939/21 describes a unique set of parameter group (PG) called Request (RQST, PGN = 0x0EA00) PG that can be used to request any other group of parameters transmitted. The specification does not include the session and presentation layers.

A. Related Work

The purpose of our work is to analyze the relevant factors on heavy duty vehicles based on J1939. Wang Quanqi [2] introduced a system for vehicle data acquisition and fault diagnosis. The main components of this system are a CAN Transceiver, ARM Coretex M3 CPU, GPRS, LCD display and SD card. It aims only at light-duty (LD) vehicles on the CAN bus On-Board-Diagnosis (OBD II) instead of the SAE J1939.

Joe Grengs et al [3] characterization of driving behaviour is also helpful for vehicle insurance companies to measure accident risk and provide personalized rates state-of-art technology implements models, mostly based on GPS location, distance travel and coarse grain speed profile. Jorge Zaldivar et al [4] also proposed an accident detection system on the road. This system combines OBD interface, mobile phone, internet, emergency ambulance notification system, using vehicle timely data diagnosis to determine if the vehicle is abnormal, and early warning to prevent accidents. Unfortunately, a driving behaviour analysis for risk driving was not performed in the proposed system.

Pal-Stefan Murvay et al [5] focus on threats to CAN bus security in commercial HD vehicles and their findings indicate no important effect on the reliability of generic CAN bus communication. However, they also show that CAN-FD is an exceptional layer for carrying the extra authentication data. Unfortunately, the emerging CAN-FD standard is still not widely used in existing commercial vehicles.

II. Implementation and Verification

A. System Structure

One of the applications of the "J1939 Bridge" developed in this work on HD vehicles is shown in the Fig. 1. A heavy vehicle's diagnostic data will be transmitted via CAN bus to the "J1939 Bridge". This device receives data like speed, engine speed, coolant temperature...etc. The external mobile App can also record frame data whenever the vehicle being monitored sends a message from the Bluetooth interface. The data is stored either on the mobile phone or in the cloud for the driving behaviour analysis.

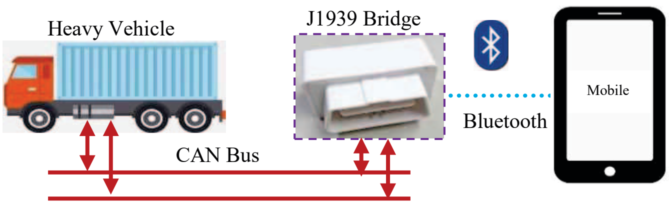


Fig. 1. Heavy Vehicle using J1939 Bridge.

As illustrated in the Fig. 2, there are four components of this "J1939 Bridge", CAN Transceiver (TJA1050), Main Controller (RL78 MCU), Bluetooth Transmission Module (HC-05) and LCD. The external oscilloscope is used between Emulator and CAN Transceiver to monitor CAN H or CAN L waveforms. Normally, raw CAN Bus data will be acquired by the CAN transceiver and sent to the RL78 CAN controller to interpret and display information on the 20x4 LCD display with J1939 data format.

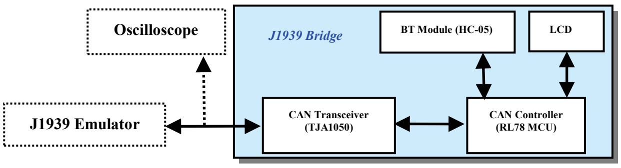


Fig. 2. Components in J1939 Bridge (solid lines).

B. Development Environment

Vector CANoe and Renesas CS+ development environments are used to check the system architecture and software stack to realize the proposed J1939 Bridge.

Vector CANoe

The Vector CANoe is an industry standard CAN analysis and simulation software tool developed by the Vector Informatik GmbH. It uses a hardware interface device which allows the user to interface via USB with various CAN protocols. We use CANoe for data gathering, and our more sophisticated attacks were implemented and executed using CANoe system [6].

Renesas CS+

Based on the Renesas RL78/F14 MCU, the "J1939 Bridge" is developed in this work. An RPB (Renesas Promotion Board) comprised of an integrated CAN Bus, LIN Bus, UART Port, HMI and I/O Port is used for development and testing. The RL78F14 is made up with a 16-bit single chip microcontroller. Data read from its CAN TX / Rx interface is viewed and sent to the Bluetooth module using SAU0-UART0.

III. EXPERIMENTAL RESULT

A. Functional verification (FV) – 1 & 2

Regarding the Functional Verification (FV), two J1939 emulators OBD SIM10 and J1939 ECU Simulator are used to assess the functionality and compatibility of our developed system. The J1939 emulator is a kind of hardware that allows the diagnostic interaction between HD vehicle and external scan tools. The basic signals analyzed include vehicle speed, RPM, temperature and coolant temperature. Fig. 3 shows the first FV-1 experiment using an oscilloscope and a CAN bus hub. This CAN Bus hub not only acts as a signal hub for CAN nodes, but also as a physical power supply for all components. The oscilloscope shows the monitored waveform of CAN_H signal and the decoded CAN message frame. From the decoded message frame, we can find that the CAN ID 18FE6CEEh we interpreted from our J1939 Bridge.

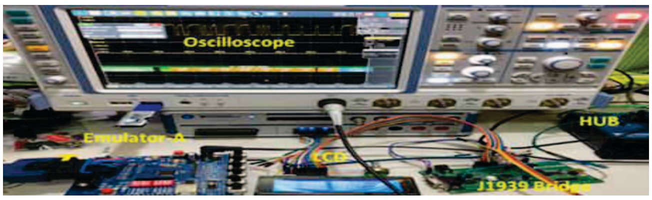


Fig. 3. FV-1 – By Emulator-A (OBD_SIM10).

As shown in Fig. 4, the J1939 ECU Simulator is used for FV-2. It is obvious from the interactive verification and testing of CAN message frames. It is obvious that our "J1939 Bridge" can effectively interpret vehicle information such as RPM, VSS, ECT, SPN, FMI, OC, EOT, TP, BPV and MAP. In addition, it is possible to display all the above information periodically on the LCD screen in every 5 seconds.

B. System architecture

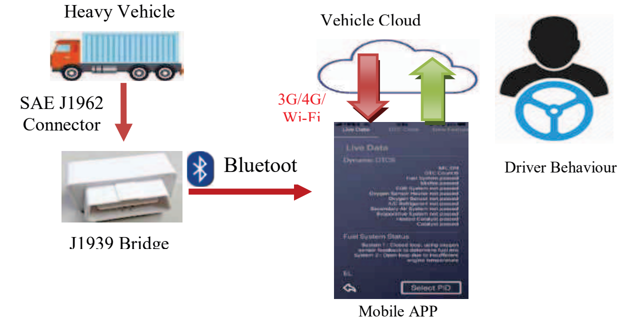


Fig. 5. System architecture.

The system architecture shown in Fig. 5 describes that heavy duty vehicle data is accessed from the J1962 interface based on the CAN bus J1939 protocol. This means that most data parameters are standardized and can be converted to human-readable form using a J1939 DBC file (PGN & SPN data). The mobile app is capable of connecting to J1939 Bridge via Bluetooth. Data stored on the mobile phone can be uploaded to the cloud with MySQL data base via 3G/4G or Wi-Fi interface. Further analysis on the fuel consumption or driving behaviour for UBI (Usage Based Insurance) can be conducted.

IV. CONCLUSION

In this paper the authors have proposed a "J1939 Bridge" based on the SAE J1939 communication standard for HD vehicles. From the experimental results, we have shown that it is possible to effectively access internal sensor information from HD vehicle linked to the J1939 Bridge. The data such as engine speed, rpm, coolant temperature, torque value, MAP, MAF...etc., can be further uploaded to the vehicle cloud via a mobile App. An artificial intelligent (AI) based algorithm can then be conducted for driver behaviour analysis.

REFERENCES

[1] J. R. Simma, "Understanding SAE J1939," Simma software, Inc, Terre Haute, Indiana, USA

[2] W. Quanqi, W. Jian and W. Yanyan, "Design of vehicle bus data acquisition and fault diagnosis system," in Proc. IEEE International Conf. CECNet, XianNing, China, pp. 245-248, April 2011.

[3] J. Grengs, X. Wang and L. Kostyniuk, "Using GPS Data to Understand Driving Behavior," Journal of Urban Technology, vol. 15, no. 2, pp. 33–53, 2008.

[4] J. Zaldivar, CT. Calafate, J. C. Cano and P. Manzoni, "Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones," IEEE 36th Conference on, pp. 813-819, 2011.

[5] P. S. Murvay and B. Groza, "Security shortcomings and counter measures for the SAE J1939 commercial vehicle bus protocol,"in IEEE Transactions On Vehicular Technology, vol. 67, no. 5, pp. 4325-4339, 2018.

[6] Vector CANoe, "user manual," Vector Informatik GmbH, 2010.

Multi-Agent Deep Reinforcement Learning for Cooperative Driving in Crowded Traffic Scenarios

Jongwon Park
Hanyang University
Seoul, Republic of Korea
pjw2091@hanyang.ac.kr

Kyushik Min
Hanyang University
Seoul, Republic of Korea
mks0813@hanyang.ac.kr

Kunsoo Huh
Hanyang University
Seoul, Republic of Korea
khuh2@hanyang.ac.kr

***Abstract*—For autonomous vehicles, lane changes on crowded roads are difficult to be performed without interactions and cooperation between vehicles. This paper proposes a novel method to learn interaction and cooperate between the multiple vehicles to solve the complex traffic problem through Multi-Agent Reinforcement Learning (MARL). The proposed network is designed based on the interaction network to learn optimal control strategies considering interaction between vehicles. By applying the proposed algorithm, the network can control and train the agents regardless of the number of agents. It is a practical advantage because the number of the vehicles is constantly changed in the real environment. The proposed method is evaluated in the connected car environment where all vehicles can exchange information with each other.**

***Index Terms*—multi-agent reinforcement learning, interaction network, cooperative driving, connected car**

I. Introduction

Driving on the congested roads is one of the most difficult problem for autonomous vehicles. Autonomous vehicles have mostly defensive path planning algorithms, so it is difficult to try to change lanes in congested situations. However, in an urban environment such as intersections, branch and merge roads, the ability to reach the desired lane in a complex traffic is essential. Fig. 1 (a) shows an example of a crowded traffic environment requiring the lane change. Human drivers, in the crowded roads, can change lanes by showing intention to other vehicles and understanding their interactions. In light of this fact, autonomous vehicle also requires the ability to understand the interactions with other vehicles.

The purpose of this study is to design and train a MARL network that solves the lane change problem in crowded roads without interfering the overall traffic flow. Several methods are applied to solve this problem. Interaction network [1] is designed to learn the physical interactions of the multiple objects and used to predict the motion of the objects. In this study, a new method, which applies interaction network to the reinforcement learning, is proposed. The reinforcement learning method is designed based on Proximal Policy Optimization (PPO). Curriculum learning [3] is applied from a few agents to many agents in the environment. It solves the difficulty of exploration in multi-agent environment. We assumed that all vehicles are connected via vehicle-to-vehicle (V2V) communication. The main contributions of this paper can be summarized as follows:

This work was supported by the Industrial Strategic Technology Development Program(10079730, Development and Evaluation of Automated Driving Systems for Motorway and City Road) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea)

Fig. 1. The image of the crowded roads and the simulation environment

- A new network structure is designed for multi-agent cooperative driving using PPO and interaction network. The proposed structure learns the best actions of multiple vehicles considering interactions.
- Since the proposed network can generate the actions regardless of the number of agents, training and inference can be performed even when the number of vehicles is changing.
- The proposed method have the centralized structure that can simultaneously control all vehicles with a single network.

II. Proposed Method

A. Environment

The proposed algorithm is trained and evaluated using the environment, which is made with Unity ML-agents [4]. The sample image of the simulation environment is shown in Fig. 1 (b). The environment provides the following information for each agent: [position, velocity, current lane, target lane]. Each agent has 5 discrete actions: [maintaining current state, acceleration, deceleration, lane change to left, lane change to right]. The reward function ($r(s)$) is designed to allow the agent to reach the target lane as fast as possible.

$$r(s) = \frac{v}{v_{max}} + r_{lc} + r_c + r_g \tag{1}$$

where v is the current velocity of the vehicle, v_{max} is the maximum velocity (40km/h) of the vehicle. r_{lc} is the lane change penalty for preventing unnecessary lane changes. r_c is

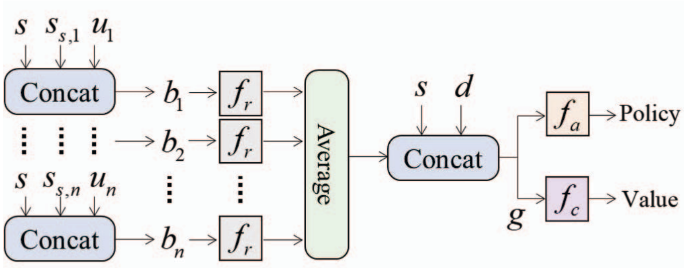


Fig. 2. Overall network structure of the proposed method

the collision penalty. r_g is the goal reward received when the agent reaches the target lane.

B. Network design

The proposed network creates an embedded feature through the interaction network. It creates relation vectors using information received from other vehicles (sender). The relation vector b can be expressed as follows.

$$b_i = [s_r; s_{s,i}; u_i] \tag{2}$$

where i is the index of other vehicles. s_r and s_s are the states of the receiver and sender, respectively. u is the relation attribute. In order to learn the intention of other vehicles, the relation attribute contains the target lane of the sender.

A Multi-Layer Perceptron (MLP) is used to learn the relation and to calculate the effect of each relation. This layer f_r can have multiple hidden layers.

$$e_i = f_r(b_i) \tag{3}$$

where e is output of f_r, representing the effect of each relation. All effect vectors of the relations are obtained on the same network f_r.

The original interaction network uses the sum operation to obtain the summarized effect for each agent. However, in the proposed structure, the summarized effect is calculated using average operation. It is designed to learn abstract interactions regardless of the number of agents. The summarized effect of each agent can be expressed as follows.

$$c = \frac{1}{N-1} \sum_{k=1}^{N-1} e_k \tag{4}$$

The summarized effect calculated in (4) is used as one of the input features of the PPO network. Other input features are the state of agent and external input. External input is additional information for making decision. The proposed method uses the target lane of the own vehicle as external input. The features are concatenated as follows.

$$g = [s; c; d] \tag{5}$$

where g indicates the input of the PPO network. s is the state of the own vehicle. d is the external input.

The PPO network consists of actor and critic network. The actor network generates the policy and critic network evaluates the value of the current state. The actor and the critic network use the same input g. The proposed network structure is shown in Fig. 2. All vehicles have the network with same parameters. The network is trained as the same method in [2].

TABLE I
PERFORMANCE COMPARISONS

Algorithm	Evaluation Factor	Number of Agents 3	6	9	12	15
CRL	# of Success	302	15	8	0	0
	Velocity [km/h]	37.9	36.5	37.3	38.7	38.6
CommNet	# of Success	258	83	46	47	40
	Velocity [km/h]	28.7	27.3	26.0	25.4	25.8
Proposed Method	# of Success	245	191	151	145	133
	Velocity [km/h]	37.1	36.3	34.7	33.2	33.2

III. RESULT

The proposed network is trained using curriculum learning from simple environment (3 vehicles) to complex environment (15 vehicles). To show the effectiveness of the proposed method, we compared it with the two other methods. The first method is centralized deep reinforcement learning (CRL) method with parameter sharing. The second is CommNet [5], which can also be trained even when the number of agents changes. The PPO algorithm is applied to all methods, and the structures of actor and critic are same.

Table 1 shows the test results for 2000 steps. There are the number of success that the vehicle reached the target lane and the average velocity. The results show that the proposed method outperforms other methods. Especially when there are many vehicles, the performance difference is remarkable. This suggests that the proposed method can learn interactions between vehicles more effectively.

The overall results of the proposed method can be found in the video at the following link.

`https://youtu.be/CacRZmjDIr4`

IV. CONCLUSIONS

In this paper, the novel MARL method is proposed to learn the cooperative driving strategy in crowded traffic scenarios. The result shows that the proposed method has better ability to understand the traffic situation and to find the optimal policy using the interactions between vehicles. It can be applied to other problems that many agents need to collaborate with.

REFERENCES

[1] P. Battaglia, R. Pascanu, M. Lai, D. J. Rezende, et al., Interaction networks for learning about objects, relations and physics, in Advances in neural information processing systems, 2016, pp. 45024510.

[2] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, Proximal policy optimization algorithms, arXiv preprint arXiv:1707.06347, 2017.

[3] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, Curriculum-learning, inProceedings of the 26th annual international conferenceon machine learning. ACM, 2009, pp. 4148.

[4] A. Juliani, V.-P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange, Unity: A general platform for intelligent agents, arXiv preprint arXiv:1809.02627, 2018.

[5] S. Sukhbaatar, R. Fergus, et al., Learning multiagent communication with backpropagation, in Advances in Neural Information Processing Systems, 2016, pp. 22442252.

New Asymmetric Data Transmission Method for In-Vehicle Network

Juho Lee and Sungkwon Park*
Department of Electronic and Computer Engineering
Hanyang University
Seoul, Korea
* Correspondence: sp2996@hanyang.ac.kr

***Abstract*— Recently, as the demand for technology related to autonomous vehicles increases, various standards have been established in order to meet these requirements. Among them, IEEE 802.3ch is establishing a PHY layer standard to support In-Vehicle Network's gigabit Ethernet. Among the standardization issues of IEEE 802.3ch, there is an asymmetric data transfer issue. This is an issue to discuss how to cope when the differences in data transmission rates are different between up and down directions. In this paper, we propose a link speed switching method using auto-negotiation as a way to provide solutions to these issues. This has the advantage of preventing data loss when switching the link speed. There is also a cost advantage because there is no need for new development.**

Keywords—In-vehicle network, Autonomous vehicle, auto-negotiation, asymmetric data transmission, Automotive Ethernet

I. INTRODUCTION

Recently, standards related to autonomous vehicles have been established in various fields. There are a variety of standards bodies, but the American standard body, the IEEE, has established standards for autonomous driving. Typical standard fields are V2X (Vehicle-to-X) communication and IVN (In-Vehicle Network) communication. Among them, IVN communication is divided into the IEEE 802.1 field, which mainly deals with communication of layer 2 or upper layers, and the IEEE 802.3 field, which mainly deals with the layer 1 (PHY layer) [1] [2]. In the case of IEEE 802.1, the AVB and TSN standards have been established for the past five years, and the communication requirements and their design have been fully discussed and are still in progress [3]. However, the IEEE 802.3 field is relatively slow to allocate the bandwidth required by the upper layer to the physical layer and to provide a standard for transmitting data. In the future, it is necessary to transmit Giga-level information in real time in order to transmit sensor information and camera information for autonomous driving [4]. For this purpose, IEEE 802.3ch, one of the working groups of IEEE 802.3, has begun to establish standards for configuring In-vehicle Ethernet links ranging from 2.5 to 10Gbps [5]. IEEE 802.3ch began its discussion at the beginning of 2018 and is now draft version 1.2 (March 2019).

IEEE 802.3ch is still in its early stages and various discussions are underway. One of the topics covered is asymmetric data transmission. Asymmetric transmission is a topic that discusses how to cope with the speed gap that can occur due to the difference in transmission speed between the upstream and downstream. In some automotive links, there is a case in which the transmission in one direction is configured to be a high-speed transmission and the transmission in the other direction is configured to be a low-speed transmission. As shown in Fig. 1, the ECU or IVN switch that transmits the control signal receives relatively small data, whereas the feedback to the ECU or IVN switch receives the data requiring high bandwidth such as image or sensor information for autonomous driving .

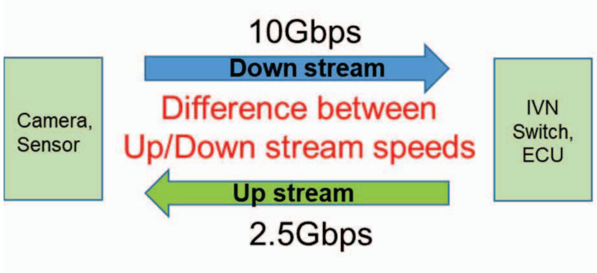


Fig. 1. Example of asymmetric data transfer in In-vehicle network

For asymmetric data transfer, it is necessary to consider switching to the commonly used symmetric data transmission. When a particular link switches to high speed or low speed, the upper layer must provide a corresponding response. If data to be transmitted suddenly in a link set at a low speed requires high-speed transmission, there may be some data loss while switching the physical link at a high speed. This is because there is a limitation on the buffer capacity of layer 2 to prevent data loss, and above all, the physical layer generally does not have a buffer. This problem can occur in the transition from high speed to low speed as well as the above-mentioned low speed to high speed link speed switching. The current IEEE 802.3ch standard does not provide a clear solution to this problem. There is a limit to find a solution only in the physical layer.

Therefore, in this paper, we propose a link speed conversion method for asymmetric data transfer between layer 1 and layer 2 by using auto-negotiation method. Thus, it is possible to determine whether the data to be transmitted in the upper layer is low or high speed and to provide the speed change according to the allocated buffer. Therefore, loss of data to be transmitted can be prevented, and it is possible to cope with situations requiring switching in real time.

This paper consists of two parts. The first part provides background information on auto-negotiation related to this paper's solution. The second part explains the speed conversion method of the asymmetric data transmission proposed in this paper.

II. BACKGROUND

Generally, devices such as ECUs, switches and gateways used in IVN are designed to support Ethernet 100Mbps. However, it is designed to cope with various link speeds 10Mbps or less and 1Gbps or more. The speed negotiation between these devices is auto-negotiation. The goal is to match the maximum communication speed that can be supported between communicating devices to provide the highest efficiency. In addition, full-duplex or half-duplex operation modes are matched to the situation. This function is defined mainly in IEEE 802.3cg [6].

Each device exchanges supportable link speeds information with electrical signals before initialization and the availability of full or half-duplex communications. At this time, the common maximum speed that can be supported is found out and the speed is set to the same. After a series of training steps, the matched device communicates at the maximum speed that it promises. When discussing how to use it for asymmetric data transmission, auto-negotiation is designed to provide maximum supportable speed for each device. By contrast, it is possible to support the lowest supportable speed so that the up and down speed differences can be accommodated. More details are covered in the next chapter.

III. New Asymmetric Data Transmission Method

Although auto-negotiation is intended to provide maximum speeds that can be supported between devices, new features have been added to suggest link rate switching methods in IEEE 802.3ch asymmetric data transmission. The new feature is to provide a common low speed. Figure 2 shows the newly added auto-negotiation mode in the existing function. In the section requiring high-speed transmission, the link speed is maximized through the auto-negotiation function. In the section requiring low-speed transmission, the link speed is configured at the lowest or the situation-specific speed. The modes for this are defined as AN(0) and AN(1), respectively.

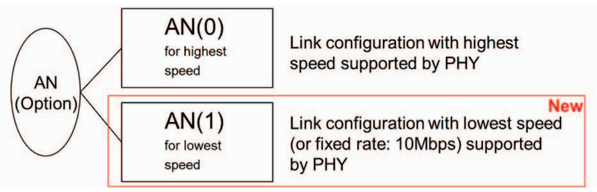


Fig. 2. New modes of auto-negotiation for asymmetric data transmission

In symmetric communication with the same uplink / downlink speed, since the maximum link speed is sought, the uplink communication rate is set to **AN(0)** as shown in Fig. 3 (1). However, if the uplink / downlink speed difference occurs as in the above-described use case, the link speed is adjusted through electrical message exchange between the devices. One of the features of auto-negotiation is that it can utilize layer 2 information. Therefore, we can access the I / O queue and utilize the data attributes of the corresponding queue. To achieve this, each device must classify data that requires high-speed and low-speed transmission into different buffers. If the low-speed data is configured to be transmitted upstream, it is set as the uplink through the **AN(0)** mode as shown in FIG. 3 (2), and the downlink is set through the **AN(1)** mode when the high-speed data is transmitted downstream. If it is necessary to set all the uplink and downlink at low speed, **AN(1)** mode should be applied to all directions. If low-speed data and high-speed data are competing on the upstream transmission side as shown in Fig. 3 (2), we assume that the upstream link is set to the highest speed on the assumption that the priority of the high speed data is high (This is because it is difficult to transmit high speed data over the lowest speed link).

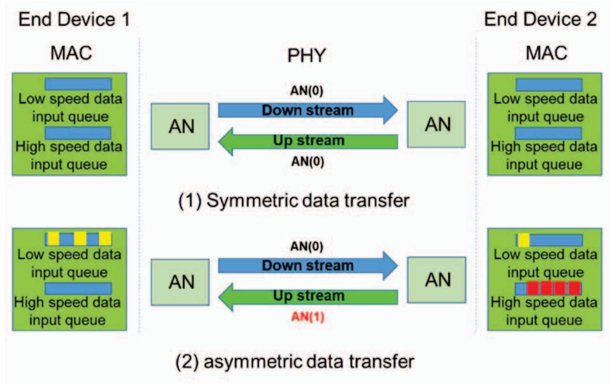


Fig. 3. Difference between symmetric and asymmetric data transmission

The advantage of asymmetric data transmission using auto-negotiation is that it is easy to apply without introducing a new solution because it is an existing technology, and it can be used immediately with a slight modification. Most importantly, it is compatible with the MAC layer, so that data can be prevented from being transmitted through the upper layer buffer in advance and data can be prevented from being lost through the upper layer buffer.

IV. Conclusion

Current IEEE 802.3ch standard activity tries to resolve this through only the OAM message of the physical layer and resource utilization of the layer without consideration of the upper layer. In this case, data loss is inevitable in speed conversion for asymmetric data transmission.

In this paper, we propose a link speed switching method based on auto-negotiation. Therefore, there is no need to present a new solution, and there are advantages that can be utilized through a slight modification. The most important advantage is that it is compatible with the MAC layer so that the upper layer can prepare for data transmission in advance and can prevent data loss by utilizing the upper layer buffer.

Acknowledgment

This work was supported by the Standard Technology Development and Spread Program of KATS/KEIT. [20001643 , Development of National Standards on Autonomous Vehicles and Use Case Guideline Provisioning]

References

[1] IEEE 802.1 Working Group. Available online: https://1.ieee802.org/ (accessed on 10 July 2019).

[2] IEEE 802.3 Ethernet Working Group. Available online: http://www.ieee802.org/3/ (accessed on 10 July 2019).

[3] Time-Sensitive Networking Task Group. Available online: http://www.ieee802.org/1/pages/tsn.html (accessed on 10 July 2019).

[4] TE Connectivity Germany GmbH, "THE ROAD TO AUTONOMOUS DRIVING: Transforming Vision into Reality," 2018.

[5] Draft Standard for Ethernet Amendment: Physical Layer Specifications and Management Parameters for Greater Than 1 Gb/s Automotive Ethernet—IEEE P802.3ch, Draft 1.2. Available online: http://www.ieee802.org/3/ch/index.html (accessed on 15 May 2019).

[6] Draft Standard for Ethernet Amendment: Physical Layer Specifications and Management Parameters for 10 Mb/s Operation and Associated Power Delivery over a Single Balanced Pair of Conductors—IEEE P802.3cg, Draft 2.4. Available online: http://www.ieee802.org/3/cg/index.html (accessed on 15 May 2019).

Heterogeneous Industrial IoT Integration for Manufacturing Production

Chao-Hsien Lee, Zheng-Lin Wu, Yun-Ting Chiu, and Yi-Shu Chen
Department of Electronic Engineering
National Taipei University of Technology
Taipei, Taiwan (R.O.C.)
chlee@mail.ntut.edu.tw (Correspondence)

***Abstract*—Due to the explosive growth and requirements of the industrial Internet of Things (IIoT), more than one standard and protocol have been defined and applied for the real environment. How to integrate these IIoT techniques for the demand for customization become one urgent and necessary research issue. This paper describes and presents one heterogeneous IIoT integration for manufacturing production.**

Keywords—Industrial Internet of Things (IIoT), Interoperability, Standardization, Manufacturing, Automation

I. Introduction

From Internet to Internet of Things (IoT), many industries are triggered to face the challenge of the next-wave industrial revolution. Industry 4.0, smart manufacturing or smart factory is the trend driving the digital transformation in the traditional manufacturing automation landscape. The convergence of information technology (IT), e.g., industrial IoT (IIoT) and big data analytics, and operational technology (OT), e.g., programmable logic controller (PLC) and computer numerical control (CNC), creates the next-generation cyber-physical factory. IIoT plays the key role in facilitating data collection from various sensors embedded on the production line and coordinating flexible automation of advanced manufacturing machines, i.e., robotic arms and automatic guided vehicles (AGVs). Taiwan is famous for producing ICT hardware and gradually joins the global supply chain. In order to increase high add-on value to related products, including smart machinery tools, ICT software uplift provides the pulling-up force. The industrial revolution of Industry 4.0 concentrates heavily on how to form better IIoT ecosystem. According to our observation, these IIoT standards roll-out of these frameworks and varying forms of industrial applications are still in their infancy. Many local manufacturing factories in Taiwan explore the IIoT techniques and see something of renewal with researchers and engineers now starting to develop Industry 4.0. However, high heterogeneity of IIoT standards, frameworks, and protocols obstructs the process of applying IIoT to the automation practice. In this paper, we describe and presents our work-in-progress project that cooperates one world-top-20 motor manufacturing company. In order to customize and fit the requirements, we integrate more than one IIoT techniques and keep it flexible.

This paper is organized as follows. Section II reviews the existing IIoT standards. Section III shows our proposed heterogeneous IIoT integration in this paper. Finally, Section IV discusses and concludes this paper.

II. Related Works

In order to provide global services in the Internet, application-layer network protocols are defined to specify how service data exchange. Thus, vertical IoT solutions are initially proposed and designed for some specific application domains. ETSI, that is a European standardization organization in the telecommunication industry, started to study and define one globally applicable and access-independent common M2M service layer [1]. That is, one horizontal M2M service layer providing a common set of APIs standardizes the common M2M operations within various hardware and software components, and then simplifies the development of IoT applications. ETSI formed the international OneM2M initiative and focused on European adaptation of OneM2M specifications called SmartM2M. In order to guarantee interoperability, OneM2M proposed an interworking framework called interworking proxy application entity (IPE) to bridge and collaborate with other IoT standards or protocols [2]. According to the second release, OneM2M featured Allseen AllJoyn, OCF IoTivity, and OMA LWM2M. Kim et al. tested several OneM2M implementations [3], including KETI Mobius, InterDigital oneMPOWER and FOKUS openMTC. Furthermore, these OneM2M implementations support three application-layer protocols, i.e., HTTP, CoAP and MQTT. From the interworking test experiments, this work shares their implementation experience and explains the key issues that need to be considered in the future. Different from the ICT industry, manufacturing vendors may not have a comprehensive knowledge of IoT. Thus, these manufacturing vendors pursued the help of some ICT enterprises or formed some kind of consortiums. For example, OPC supported by Microsoft had defined the communication standard in the industrial automation space and provided the information flow among machinery tools from different vendors. Although the initial OPC standard was limited to the Windows operating system, OPC introduced service-oriented, vendor-independent, and open-platform unified architecture (UA) for the manufacturing industry. Similar to the IoT/M2M interworking development, some studies started to integrate OPC UA with other IoT/M2M standards. For example, Cavalieri et al. aimed to integrate OPC UA and OCF IoTivity, which is one of the biggest industrial connectivity standard organizations [4, 5]. That is, an OCF device can publish its information to an OPC UA server that allows making this information available to any OPC UA compliant devices, and vice versa. In this paper, Cavalieri et al. specified the defect of OCF IoTivity to bridge the mapping with OPC UA and contribute one general architecture of OCF bridge device.

III. Heterogeneous IIoT Integration

From the aforementioned survey in Section II, we are motivated to propose one heterogeneous IIoT integration because no IIoT standard can fit all possible conditions. Furthermore, regarding the manufacturing automation, the production machines purchased in different time period may support different standards or equip different interfaces. If the proposed IIoT integration is scalable, the manufacturing company has the flexibility to gradually extend their

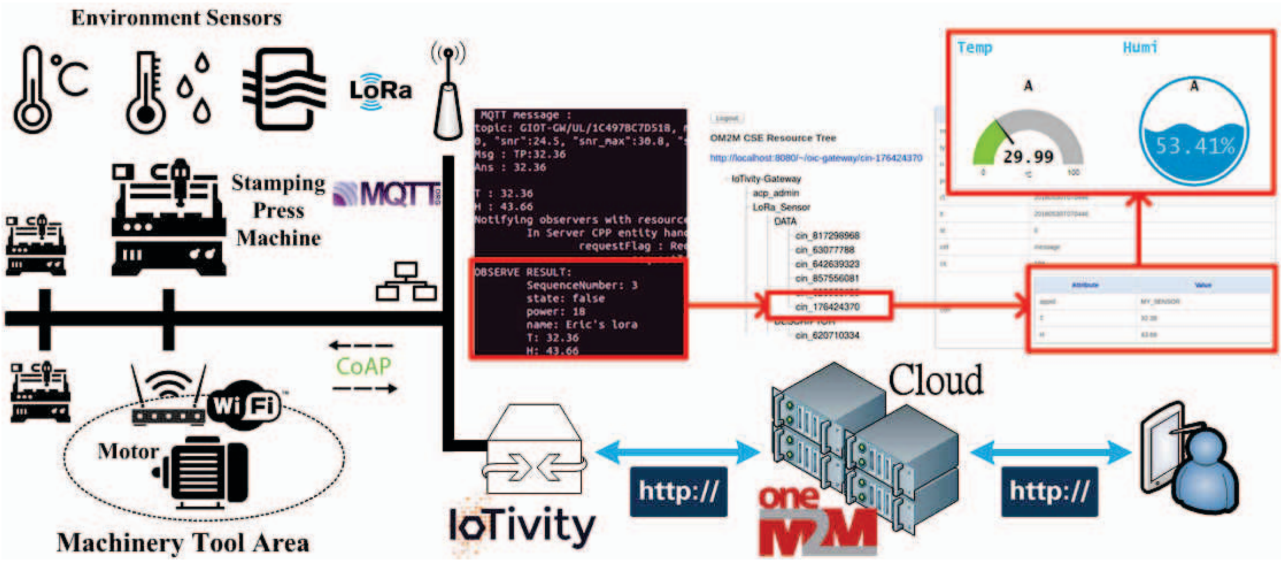


Fig. 1. Our proposed heterogeneous IIoT integration, including OM2M, IoTivity, HTTP, CoAP, MQTT, LoRa, Ethernet, and Wi-Fi.

production lines. Fig. 1 depicts the proposed heterogeneous IIoT integration. First, regarding the evolving progress of interoperability, OneM2M is selected to play the core role of our proposal. In this paper, we utilize one open-source project called OM2M which is based on the OneM2M and ETSI SmartM2M specifications. The IN-CSE component of OM2M is configured to our cloud server and responsible for collecting all sensor data. Then, there are two different realistic scenarios discussed in this paper. In the first scenario, we consider the continuous online data streaming requirement. We install our own sensors to motors and stamping press machines. Regarding the motor evaluation, the real-time values of voltage, current, power, power factory, shaft temperature, coil temperature, and vibration frequency are collected. On the other hand, the real-time values of strain and temperature are collected for the maintenance of stamping press machines. Since the sampling rate of these sensor values is high, the corresponding data volume should be transmitted using Ethernet or Wi-Fi. According to the OneM2M specification, OneM2M does not provide its own device management. Therefore, we adopt the OCF IoTivity for device management. IoTivity supports dynamic device discovery. In other words, it is scalable for one-by-one motor evaluation or machine maintenance schedule. In order to keep flexible, we utilize the HTTP RESTful API to interwork the OM2M and OCF IoTivity instead of the specific IPE. In the second scenario, we consider the wide-area environmental monitoring requirement. We build the LoRa LPWAN network in the factory building and interconnect the sensors of air quality and machine power. Hence, when machines in the factory building operate the production process, the air condition would be detected for safety concern. The Gemteks LoRa router provides the bridge function from LoRa to MQTT. However, the default transmission protocol of OCF IoTivity is CoAP. Although IoTivity declares it is compatible with MQTT. There is no existing MQTT module inside IoTivity. In this paper, we refer the IoTivity plug-in example and then implement our own MQTT module. Thus, the factory environmental data can be converged into OCF IoTivity like the data of motors and stamping press machines. Finally, we utilize the NodeRED that is one graphic flow programming tool. Since all sensor data of the two aforementioned scenarios are collected and converged into OM2M, we utilize the NodeRED to access the OM2M resource tree and then visualize all data in the web dashboard.

IV. Discussion and Conclusion

In this paper, we presented our work-in-progress heterogeneous IIoT integration for manufacturing production. All equipment and sensors distributed in 3 wired/wireless access networks, i.e., Ethernet, Wi-Fi, and LoRa, are interconnected using OCF IoTivity and then all data are transmitted using 3 application-layer protocols, i.e., CoAP, MQTT and HTTP, to the private cloud. Finally, we followed OneM2M to keep all resources be standardized and easily visualized in the customized format. Based on our implementation experience in this paper, the efficiency of integration and automation is deeply affected by the interaction of triggers and data flows. As the scale-up of data and device volume, the static settings and configurations become impractical. Thus, our future work is trying to further study the unified and comprehensive semantic ontology by integrating the OPC UA framework.

Acknowledgment

This research is supported by the Ministry of Science and Technology of the Republic of China (Taiwan) under the grant numbers MOST 107-2218-E-027-018 and 108-2637-E-027-005.

References

[1] J. Swetina, G. Lu, P. Jacobs, F. Ennesser and J. Song, "Toward a standardized common M2M service layer platform: Introduction to oneM2M," in IEEE Wireless Communications, vol. 21, no. 3, pp. 20-26, June 2014.

[2] Hyuncheol Park, Hoichang Kim, Hotaek Joo, JaeSeung Song, “Recent advancements in the Internet-of-Things related standards: A oneM2M perspective,” in ICT Express, vol. 2, no. 3, pp. 126-129, Sep. 2016.

[3] J. Kim et al., "Standard-based IoT platforms interworking: implementation, experiences, and lessons learned," in IEEE Communications Magazine, vol. 54, no. 7, pp. 48-54, July 2016.

[4] S. Cavalieri, M. G. Salafia and M. S. Scroppo, "Realising Interoperability Between OPC UA and OCF," in IEEE Access, vol. 6, pp. 69342-69357, Nov. 2018.

[5] S. Cavalieri, M. G. Salafia and M. S. Scroppo, "Mapping OPC UA AddressSpace to OCF resource model," in Proceedings of IEEE Industrial Cyber-Physical Systems (ICPS), St. Petersburg , Russia, May 15-18, 2018, pp. 135-140.

Integrated Simulator for Evaluating Cooperative Eco-driving System

Geonil Lee, Jae-il Jung*
Department of Electronics and Computer Engineering
Hanyang University
Seoul, Republic of Korea
{lgi1992, jijung}@hanyang.ac.kr

***Abstract*— In recent years, connected and automated vehicles (CAVs) have attracted considerable attention because they can improve driver convenience and safety using vehicle to everything (V2X) communication. CAV system must satisfy safety and performance requirements, and they are required to go through rigorous evaluation processes. In this study, we design an integrated simulator to evaluate a CAV system, cooperative eco-driving system, which models vehicular ad hoc network (VANET) topology, driver models, vehicle models, and vehicle control algorithms. This is done by integrating three simulators: network simulator, traffic simulator, and driving simulator. Cooperative eco-driving system requires precise and accurate map data including real-time traffic conditions and surrounding environmental information. Therefore, we construct a local map system based on V2X communication to provide a host vehicle (HV) with surrounding traffic and environment information. A local map system generates a local map by utilizing surrounding traffic information of HV acquired from each simulator and shares it with the simulators.**

Keywords*— *CAV, VANET, V2X, Eco-driving, Local Map, Integrated Simulator

I. Introduction

Connected and automated vehicles (CAVs) provide convenience and safety to drivers based on the connectivity between cars and surrounding IT devices. It is possible to provide a variety of safety and infotainment services by utilizing vehicles as an Internet of Things (IoT) platform, and there is ongoing development of technology and standard for CAV system. Many engineers have developed various services for CAV system, and these services must satisfy the safety and performance requirements by undergoing rigorous evaluation processes. In this study, we design a simulation test bed that integrates network, traffic, and driving simulators to evaluate the performance of a CAV system, namely, the cooperative eco-driving system, taking into account both V2X communication and vehicle control system. The integrated simulator that is used to evaluate CAV system can model vehicular ad hoc network (VANET) topology, various large-scale scenarios, vehicle models and vehicle control algorithms.

There are many studies related to VANET simulation tool [1]-[3]. However, because many existing tools use two simulators for traffic and networks, they cannot reflect sophisticated vehicle dynamics models and driver models. To solve this problem, there is a need to integrate a driving simulator with the existing system, which enables us to simulate vehicle dynamics models and driver models. Zhao et al. [4] explained the need to integrate traffic, network, and driving simulators and presented the concept of their development for an integrated traffic-driving-networking simulator. However, they did not consider the V2X communication-based local map of the HV, which is an essential element for the CAV system scenario. In this paper, we propose an our integrated simulator that evaluates a cooperative eco-driving system. The proposed simulator

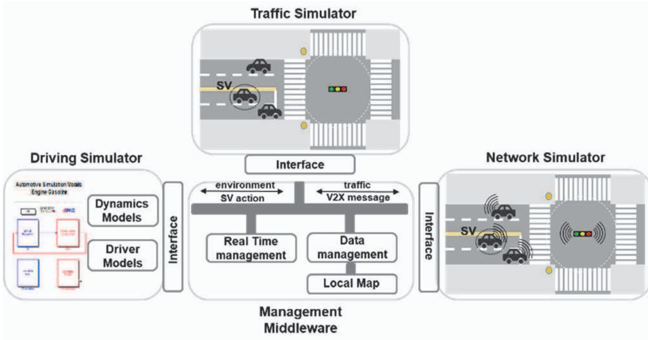


Fig. 1. Architecture of integrated simulator for evaluating connected vehicle based eco-driving systems

focuses on a local map system that creates the local map of the HV by utilizing the surrounding traffic information of HV acquired from V2X communication messages and the interaction among simulators. A local map created by taking the dynamic traffic, V2X, and environmental information generated during the simulation process is shared among the simulators; hence, it can be used to increase the performance of the eco-driving system.

II. Integrated simulator to evaluate cooperative eco-driving system

Fig. 1 depicts the architecture of our integrated simulator for the evaluation of cooperative eco-driving system. To develop and evaluate cooperative eco-driving system considering both V2X communication technology and vehicle control technology, we developed an integrated simulator that consists of QualNet, which serves as the network simulator, VISSIM, which serves as the traffic simulator, Autonomie, which serves as the driving simulator, and management middleware, which manages the parameters as well as time synchronization and simulation data.

First, the behavior of each simulator for the cooperative eco-driving system is described, after with the time-sync management system for sequential operation and time synchronization between simulators and data management system for local map generation is described.

A. *QualNet*

In our work, we used a QualNet 4.5 network simulator [5]. QualNet is used to model and simulate vehicular networks, including various protocols at different layers of the protocol stack, and the wireless communication channel. However, QualNet does not contain an IEEE 802.11p/wireless access in vehicular environment (WAVE) protocol which is standard for vehicular communication system, so we implemented IEEE 802.11p protocol and WAVE/1609 protocol stack to simulate the VANET environment. In addition, we modified suitable propagation models for V2X communication, such as the Nakagami fading model, two-ray path loss model, and corner path loss model. With these protocol stacks and channel models, V2X communication can be simulated precisely.

B. VISSIM

The traffic simulator used in our work is VISSIM [6]. VISSIM is a discrete-time, microscopic road traffic simulation package designed to handle large road traffic scenarios. VISSIM provides a GUI for users to easily configure traffic scenarios. For our simulation, we implemented a simulation scenario for eco-driving system based on Google map data using the VISSIM GUI. We also implemented the VISSIM controller-server in VISSIM and the VISSIM controller-client in the management middleware which interacts with each other using UDP/IP.

C. Autonimie

We used the driving simulator Autonomie, which is a driving simulation tool developed by Argonne National Laboratory, to provide the vehicle modeling function in our integrated simulator [7]. Autonomie provides the function of customizing the structure of the vehicle, and it allows specifications of each component of the vehicle to be set up, so that various vehicle models can be constructed for the simulation. We implemented the vehicle model used in the simulation using Autonomie, and instead of the control algorithm provided by Autonomie, we used an in-house developed HCU control algorithm.

D. Time-sync Management System

The simulation proceeds by repeating the interactions between the five main entities: the time-sync management system, VISSIM, data management system, QualNet, and Autonomie, and the simulation is initiated when the time-sync management system sends a SIMSTEP message to VISSIM which indicates the start of each step. The start time of each step is recorded in the time-sync management system. After 100ms from the start time of the previous step recorded, the time-sync management system sends the SIMSTEP message of the next step so that the simulation can be repeated every 100 ms.

E. Data Management System

Fig. 2 shows the idea of the local map. The local map is a map data store that can manage both static and dynamic information about stationary objects (e.g., roadside infrastructure, road) or moving objects (e.g., pedestrian, vehicle) within a certain range of HV. We generate the local map by storing simulation information to obtain the static and dynamic information required to operate cooperative eco-driving service.

Fig. 3 depicts the architecture used to generate the local map in management middleware. Local map is generated

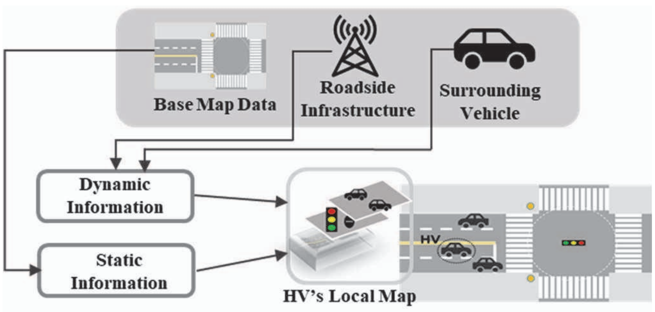


Fig. 2. Composition of the local map.

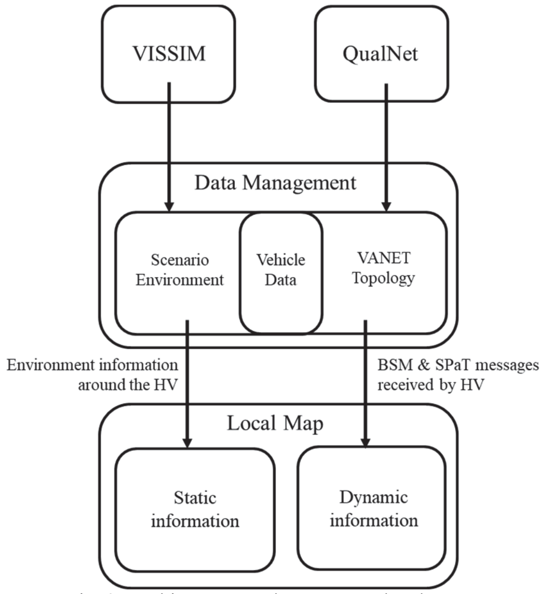


Fig. 3. Architecture used to generate local map.

using scenario environment information obtained from VISSIM, and VANET topology information obtained from QualNet stored in the data management system.

III. Conclusion

We have designed a simulation test bed to evaluate the cooperative eco-driving system. The designed simulation test bed models the VANET topology, driver model, vehicle model and vehicle control algorithm by integrating three simulators: a network simulator, traffic simulator, and driving simulator. The most important contribution of this paper is the local map system, which generates a local map by considering the dynamic map information based on V2X communication messages. Our integrated simulator can be used as a test bed to evaluate the hardware equipment used in actual vehicles by integrating then with hardware such as communication equipment and vehicle ECU, and we are currently conducting these additional studies.

Acknowledgment

This work was supported by a grant from the Korea Evaluation Institute of Industrial Technology (KEIT) funded by the Republic of Korea (No. 20000578).

References

[1] C. Sommer, R. German, and F. Derssler, “Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis,” IEEE Trans. on Mob. Comput., vol. 10(1), pp. 3-15, January 2011.

[2] B. Schunemann, “V2X simulation runtime infrastructure VSimRTI: An assessment tool to design smart traffic management systems,” Comput. Netw., vol. 55, pp. 3189–3198, October 2011.

[3] V. Kumar, R. Bauza, F. Filali, J. Gozalvez, L. Lin and M. Rondinone, “iTETRIS – A Large Scale Integrated Simulation Platform for V2X Communications: Application to Real-Time Traffic Management,” 9th Int. Conf. on ITS Telecommun., Lille, France, 2009.

[4] Y. Zhao, A. Wagh, Y. Hou, K. Hulme and C.Qiao, “Integrated traffic-driving-networking simulator for the design of connected vehicle applications: eco-signal case study,” J. of Intell. Transp. Syst., vol. 20, pp. 75–87, 2016.

[5] Scalable-Networks, “QualNet(Quality Networking),” http://web.scalable-networks.com

[6] VISSIM (Verkehr In Stadten – SIMulationsmodell), http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim

[7] Autonomie, http://www.autonomie.net

Pilot Allocation for Iteration Reduction of Channel Estimation Using MF-Based Interpolation

Norisato Suga, Ryohei Sasaki, and Toshihiro Furukawa
Faculty of Engineering, Tokyo University of Science, Tokyo, Japan
E-mail: norisato.suga@rs.tus.ac.jp

***Abstract*—This paper studies an iteration reduction of channel estimation method using matrix factorization (MF) based interpolation. It is important to set the initial value close to the optimal solution to reduce the iteration. To obtain an appropriate initial value, exploiting the existing methods with low computational complexity such as linear interpolation is reasonable. While MF-based interpolation can achieve high accuracy when randomly determined pilot symbol allocation, it is not suitable for the existing method. This paper proposes a pilot allocation where fixed position and the randomly determined position are mixed to perform both the linear interpolation and MF-based interpolation. Computational simulation result shows the iteration reduction by the proposed pilot symbol allocation.**

***Index Terms*—iteration reduction, channel estimation, interpolation, matrix factorization, OFDM**

I. INTRODUCTION

In current OFDM (orthogonal frequency division multiplexing) based wireless communication systems such as LTE (long term evolution) and WLAN (wireless local area network), channel estimation is essential. Since the performance of channel estimation affects the following equalization and error-correcting, the channel estimation should be performed as accurately as possible. Recently, the channel estimation using matrix factorization (MF) based interpolation [1] has been proposed, which shows high accuracy compared with the conventional interpolation method such as sinc interpolation and DFT based interpolation [2].

In MF-based channel estimation, the optimization problem is iteratively solved by alternative least square (ALS) [3]. This results in a large number of iterations and a long processing delay. It is important to set the initial value as close to the optimal solution as possible. To obtain an appropriate initial value, exploiting the existing methods with low computational complexity such as linear interpolation is reasonable. On the other hand, the estimation accuracy of the MF-based interpolation method is significantly degraded when the pilot symbols arranged in periodic. Moreover, if some subcarriers/OFDM symbols do not include any pilot symbol, the performance is also degraded. Therefore, in [1], pilot symbols are randomly allocated under the constraint that at least one pilot symbol is included for all the subcarriers and OFDM symbols. However, since the pilot symbols may not surround some data symbols, the estimation by the conventional method is difficult to be performed.

Therefore, this paper proposes a pilot allocation that enables to perform both the linear interpolation and MF-based interpolation. In the proposed allocation, fixed position and the randomly determined position are mixed. Then, the proposed allocation enables MF-based interpolation with the estimated value by the linear interpolation as the initial value and reduces the number of iterations. The contribution of this paper is that we propose the pilot symbol allocation for the MF-based interpolation method to reduce the number of iteration and confirm the iteration reduction by the computational simulation.

Notations: Underlined letters represent vectors. Boldface capital letters represent matrices. Calligraphic letter means sets. $(\cdot)^H$ represents Hermitian transpose.

II. PROPOSED METHOD

A channel frequency response (CFR) of k-th subcarrier in i-th OFDM symbol is denoted by $H_{i,k}$. When the number of subcarriers is N_{sub}, the i-th CFR $\underline{H}_i \in \mathcal{C}^{N_{\mathrm{sub}}\times 1}$ is given as $\underline{H}_i = \mathbf{G}\underline{h}_i$ where $\mathbf{G} \in \mathcal{C}^{N_{\mathrm{sub}}\times L}$ is DFT matrix, $\underline{h}_i \in \mathcal{C}^{L\times 1}$ is i-th CIR, and $\mathcal{C}^{m\times n}$ is the complex matrix set with size $m\times n$. The channel matrix is represented as $\mathbf{H} = [\underline{H}_1, \cdots, \underline{H}_{N_{\mathrm{sym}}}]$ where N_{sym} is the number of OFDM symbols. Each column of the channel matrix is a linear combination of Fourier basis, and the rank of the channel matrix is limited to L.

Here, we define a two-dimensional integer set $\mathcal{K}$ whose element (i', k') indicates pilot symbol location in the time-frequency domain. The CFR at pilot symbol location $H_{i',k'}$ can be estimated by, for example, linear square (LS) estimation or minimum mean square error estimation [5]. Since the rank of the channel matrix is L, the channel matrix can be factorized to $\mathbf{H} = \mathbf{P}^H\mathbf{Q}$ where $\mathbf{P} \in \mathcal{C}^{L\times N_{\mathrm{sub}}}$ and $\mathbf{Q} \in \mathcal{C}^{L\times N_{\mathrm{sym}}}$. According to [4], MF problem is formulated as follows:

$$\min_{\mathbf{P},\mathbf{Q}} \sum_{(i',k')\in\mathcal{K}} \left(\hat{H}_{i',k'} - \underline{p}_{i'}^H \underline{q}_{k'}\right)^2 + \lambda\left(||\underline{p}_{i'}||^2 + ||\underline{q}_{k'}||^2\right) \quad (1)$$

where $\underline{p}_{i'}$ and $\underline{q}_{k'}$ are the i' and k'-th column of $\mathbf{P}$ and $\mathbf{Q}$, and λ is the parameter which controls the extent of regularization. The optimization problem in Eq. 1 can be iteratively solved by ALS [3].

Since we update matrices $\mathbf{P}$ and $\mathbf{Q}$ alternately, a large number of updating is required to get an accurate channel estimate. However, in the communication systems such as LTE and WLAN, processing delay is an issue. It is expected to obtain an accurate estimated value with less updating by using roughly estimated value by the existing interpolation method as an initial value. While MF-based interpolation can achieve high accuracy when randomly determined pilot

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

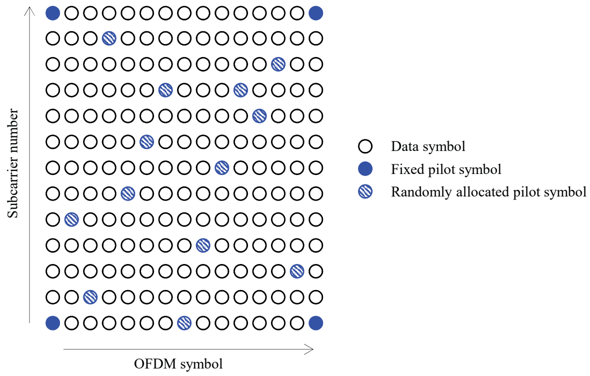


Fig. 1. Pilot symbol allocation in the proposed method.

TABLE I
SIMULATION PARAMETERS

Frequency band	1.7 GHz
Bandwidth	10 MHz
Number of Subcarriers	600
Number of OFDM symbols	600
OFDM symbol duration (include CP)	66.67 + 4.69 μsec
Delay profile	ETU [6]
Maximum Doppler frequency	300 Hz
Number of rays	20

symbol allocation, It is not suitable for the existing method. Therefore, we propose new pilot allocation where some of the pilot symbols are arranged to a fixed location, and the others are arranged to a random location. The example of the proposed pilot allocation is shown in Fig. 1. In this example, the four corner shows the fixed pilot allocation, and the other pilot symbol allocation is randomly determined under the constraint that each row and column includes at least one pilot symbol.

III. Computational Simulation

This section shows simulation result to confirm the reduction of number of iteration. We assume LTE downlink systems and the Rayleigh channel with additive white Gaussian noise. The simulation parameters are shown in Table I. The least square estimation is used as the estimation of CFR at each pilot symbol location. The estimated value is given as $\hat{H}_{i,k} = H_{i,k} + w_{i,k}$ where $w_{i,k}$ is estimation error which obeys complex Gaussian distribution with mean zero and variance σ^2. SNR (signal to noise ratio) is defined as $\frac{1}{\sigma^2}$. For the MF-based interpolation, we set λ is 10^{-5}. The estimation accuracy is measured by mean absolute error (MAE) which is calculated by $\frac{1}{N_{\mathrm{sym}} N_{\mathrm{sub}}} \sum_{i,k} |H_{i,k} - \hat{H}_{i,k}|$.

Figure 2 shows the MAE performance of the proposed method ('Prop.' in Fig. 2) and the conventional ones ('Conv.' in Fig. 2) in the ETU scenario [6] with maximum Doppler frequency 300 Hz. The conventional method uses the estimated value by LS estimation at the pilot location and zero-padded at the data symbol location. We can observe that the proposed method shows fast convergence performance compared with the MF-based method [1]. The proposed method enables to start the iteration algorithm from a position which closes to the

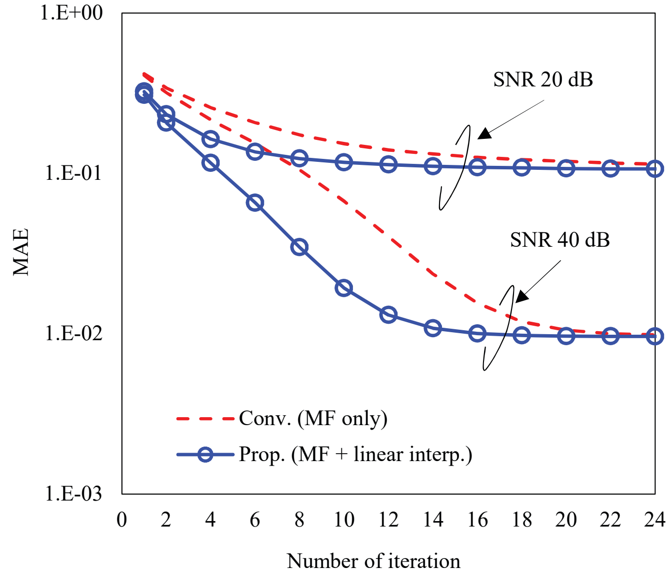


Fig. 2. MAE performance comparison between the conventional method and the proposed method.

optimal solution because the initial value is roughly estimated by linear interpolation. We can confirm the reduction of the iteration count by the proposed method when SNR is both 20 and 40. In particular, when the SNR is 20 dB, the proposed method converges to the optimal solution in about ten times, but the conventional method requires about twenty times.

IV. Conclusion

In this paper, we proposed the new pilot symbol allocation to apply the conventional interpolation method such as linear interpolation as the initial value of MF based interpolation method. We validated that the proposed method can reduce the number of iteration of MF based interpolation method in LTE downlink channel. A part of our future work will consider the effect of frequency offset for the MF-based interpolation method.

References

[1] N. Suga, R. Sasaki, and T. Furukawa, "Channel Estimation Using Matrix Factorization Based Interpolation for OFDM Systems," in *Proc. 2019 IEEE 90th Vehicular Technology Conference (VTC-Fall)*, Sept. 2019, to appear.

[2] M. Zourob and R. Rao, "2×1-D fast Fourier transform interpolation for LTE-A OFDM pilot-based channel estimation," in *Proc. Int. Conf. on Electrical and Computing Technologies and Applications (ICECTA)*, pp. 1–5, Nov. 2017.

[3] X. Liu, C. Aggarwal, Y. Li, X. Kong, X. Sun, and S. Sathe, "Kernelized Matrix Factorization for Collaborative Filtering," in *Proc. SIAM Int. Conf. on Data Mining*, pp. 378–386, May 2016.

[4] Y. Koren, R. Bell and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," in *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[5] M. Zourob and R. Rao, "Hybrid Lower-Complexity Wiener Filter for Pilot-Based Channel Estimation for C-RS in LTE-A DL System," in *Journal of Mobile Networks and Applications*, vol. 23, no. 4, pp. 921–939, Nov. 2017.

[6] 3GPP TS 36.101. "User Equipment (UE) Radio Transmission and Reception." 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA).

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

A 0.8V 14bit 62.5kSPS non-binary cyclic ADC using SOTB CMOS technology

Shuichiro Yamada, Toshiki Ohtsu, Minami Sasaki, Hao San, Tatsuji Matsuura, Masao Hotta

Integrated System Laboratory, Tokyo City University

***Abstract*— This paper presents a proof-of-concept of the low supply voltage circuit technique for high resolution cyclic analog-to-digital converter (ADC). By utilizing substrate-voltage-control technique for SOTB CMOS, high resolution cyclic ADC can be realized at supply voltage as low as Vdd= 0.8V. A prototype 14bit cyclic ADC is designed and fabricated in 65nm SOTB CMOS technology. Measured DNL=-0.80/1.11LSB, INL=-5.05/3.49LSB are achieved while a 7.81kHz sinusoidal input is sampled at 62.5ksps.**

Keywords*— *SOTB, Non-binary ADC, Dynamic analog circuit

I. INTRODUCTION

High resolution ADCs are widely used in mixed-signal SoC in the fields of both consumer and industrial applications. In the edge-node of sensor networks powered by energy harvesting technology, low-voltage ADC embedded wireless sensor device is required as the interface of the analog world to digital domain. In nano-meter CMOS technology, the linearity of ADC strongly depends on the non-ideality of analog elements' characteristic. Therefore, high gain wideband amplifiers and well-matched circuit elements such as transistors and capacitors are required to achieve a high accuracy ADC. In previous works, we have proposed a switched capacitor (SC) non-binary cyclic ADC tolerates the capacitors mismatch, amplifier and comparator offsets, and amplifier finite DC gain [1], [2]. However, the dynamic range and linearity of cyclic ADC are limited by the performance of amplifier while the supply voltage is less than 1V. In this study, we proposed a non-binary cyclic ADC using dynamic amplifier for low voltage operation. In order to demonstrate the feasibility and effectiveness of the proposed circuit, a 14bit non-binary cyclic ADC with inverter-based amplifier is designed and fabricated in 65nm SOTB CMOS technology. Measured DNL=-0.80/1.11LSB, INL=-5.05/3.49LSB are achieved at 62.5ksps while the differential sinusoidal input of 7.81k Hz is sampled at 62.5kSPS under supply voltage of 0.8V.

II. NON-BINARY CYCLIC ADC

Fig.1 shows the block diagram of the conversion stage of non-binary cyclic ADC. This conversion stage resolves a 1-bit per conversion, residual signal (Vres) is feed backed to the input node of conversion stage for the next AD conversion step. Re-sampled residue is processed step by step to realize a high resolution conversion. Since cyclic ADC uses the same conversion stage repeatedly, so that it can realize high resolution in small chip area. For high resolution conversion, although multi-step conversion is required, and hence the conversion speed is limited, however, cyclic ADC is still widely used for the low frequency applications such as edge senor of IoT network. The configuration of non-binary cyclic ADC is similar to a conventional binary cyclic ADC, however the residual signal is not amplified by 2 but β (1 < β < 2), and the output signal of the DAC is scaled by (β-1).

When the analog input range is within [−Vref, Vref], the output code of sub ADC in the *ith*-step is b_i (b_i= -1 or 1) and the corresponding output voltage of 1-bit DAC is -Vref or Vref (the full-scale-voltage of the analog input signal, V_{FS} = 2Vref), then the input/output transfer function of the cyclic AD conversion stage can be expressed as,

$$V_{res} = \beta V_{in} + b_i(\beta - 1)V_{ref} \quad (1)$$

In addition, the ratio of input voltage to full-scale voltage of ADC can be expressed as

$$\frac{V_{in}}{V_{FS}} = (\beta - 1)\sum_{n=1}^{N} \frac{1}{\beta^n} b_n + \frac{1}{\beta^N}\frac{V_{resN}}{V_{ref}} \quad (2)$$

Here, V_{FS} = 2Vref, which is the full-scale analog input voltage of the ADC. Eq.(2) is a non-binary expansion of the analog input signal, and β is the radix value of the non-binary ADC's output code. Since radix is 1 < β < 2, the redundancy of this conversion stage tolerates the offset caused by comparator and/or amplifier [1],[2].

III. THE PROPOSED LOW-VOLTAGE NON-BINARY CYCLIC ADC

Fig.2 shows the circuit configuration of the proposed non-binary cyclic ADC. The sampling capacitances C_S, C_f are designed as C_S/C_f= 0.93/1 to realize multiply-by-β (β＝1.93) amplification. Although a single-ended structure is shown for simplicity, the actual implementation is fully differential configuration. Bootstrapped switches are used as the analog sampling switches to reduce the nonlinear effect of ON-resistance [3], whereas all the others are CMOS switches. In order to realize high linearity of ADC at low voltage, a pseudo differential dynamic amplifier is proposed. The schematic of core amplifier is shown in Fig.3. It is an inverter-based dynamic amplifier [4], however, in order to enhance the output impedance of amplifier, high V_{th} transistor and the substrate voltage biased to V_{cm} of 3rd-stage is used. Therefore, high DC gain of amplifier is realized even at low supply voltage. A dynamic comparator circuit as shown in Fig.4 is also proposed for low voltage operation [5].

In order to confirm the feasibility and performance of the proposed low supply voltage circuit technique for high resolution ADC, a prototype non-binary cyclic ADC is designed and fabricated in 65nm SOTB CMOS technology. Fig.5 shows the chip microphotograph of the proposed ADC, its active area is 0.1mm^2. Fig.6 shows the measured output power spectrum of proposed ADC. SNDR=58.6dB is achieved while the differential sinusoidal input of 7.81 kHz with 1.2Vppd is sampled at 62.5kSPS while supply voltage is 0.8V. Fig.7 shows the measured linearity characteristics of the proposed ADC, DNL=-0.80/+1.11LSB and INL=-5.05/+3.49 LSB, respectively.

IV. CONCLUSION

A non-binary cyclic ADC composed of dynamic analog circuits is proposed. Substrate voltage control technique for SOTB CMOS is utilized to realize high resolution ADC at low

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

supply voltage. Measurement results show the effectiveness of the implementation method of the proposed dynamic analog circuit and the feasibility of low supply voltage non-binary cyclic ADC.

V. ACKNOWLEDGMENT

The authors would like to thank Renesas Electronics which is supporting this research. This work was supported by JSPS KAKENHI 19K04518. This research was supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Cadence Design Systems, Inc.

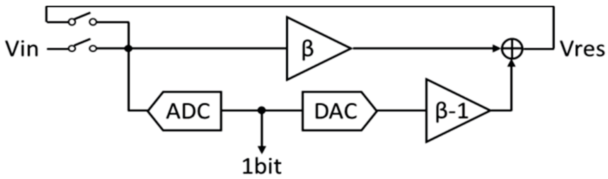


Fig. 1. Block diagram of conversion stage of non-binary cyclic ADC.

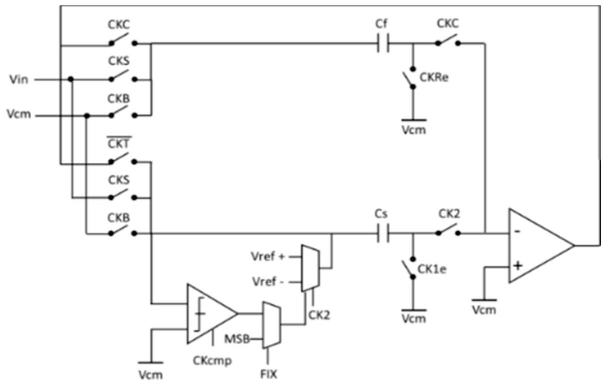


Fig. 2 Circuit configuration of non-binary cyclic ADC.

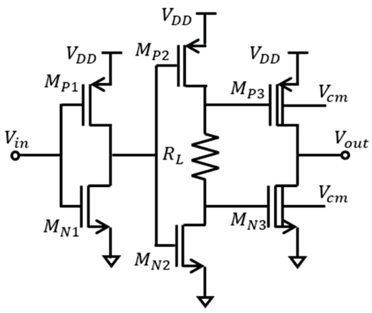


Fig. 3. Circuit structure of proposed dynamic amplifier.

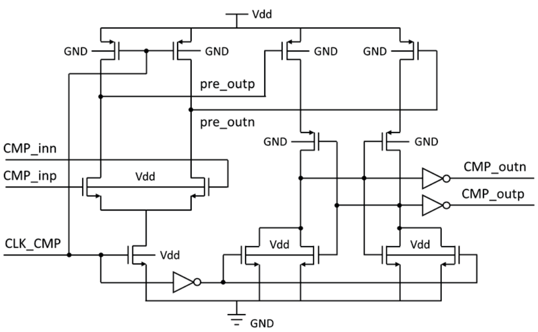


Fig. 4. Circuit structure of proposed dynamic comparator.

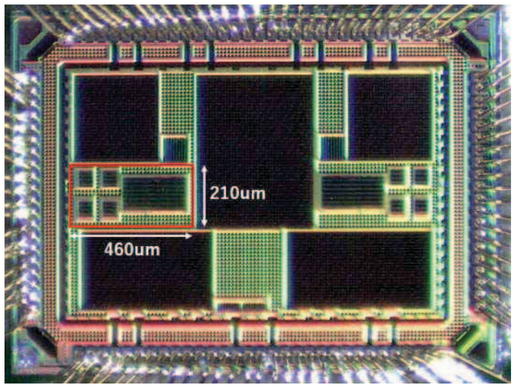


Fig. 5. Chip microphotograph

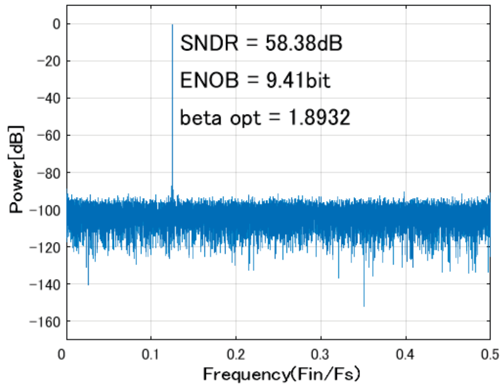


Fig. 6. Measured output power spectrum of proposed ADC.

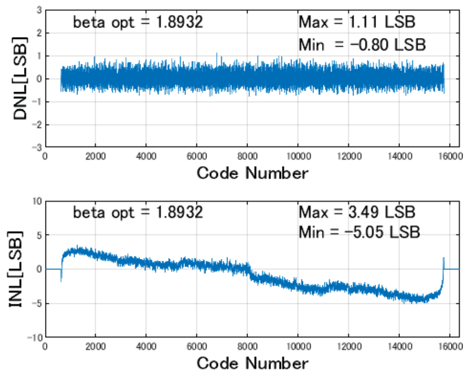


Fig. 7. Measured DNL and INL of proposed ADC.

REFERENCES

[1] Y. Watanabe, et al., "Experimental implementation of a 14 bit 80kSPS non-binary cyclic ADC," Analog Integrated Circuits and Signal Processing, Springer, Volume 97, Issue 2, pp 207–214, Nov, 2018.

[2] H. San, et al., "A 12-bit 1.25MS/s Area-efficient Radix-value Self-estimated Non-binary Cyclic ADC with Relaxed Requirements on Analog Components," IEICE Trans. Fundamentals, Vol.E100-A, No.2, pp.534-540, Feb. 2017.

[3] Ohtsu, et al., "Bootstrapped switch in SOTB CMOS," Papers of Technical Meeting on Electric circuit, ECT-18-072, IEE Japan, October 2018.

[4] Hershberg,et al., "Ring amplifiers for switched capacitor circuits," IEEE JSSC, Vol.47 No.12, pp2928-2942, Dec. 2012.

[5] Sasaki, et al., "Study on the design of low voltage dynamic comparator," Papers of Technical Meeting on Electronic circuits, ECT-18-073, IEE Japan October 2018.

An Evolutionary Computation Approach for Approximate Computing of PNN Hardware Circuits

1s Ching-Yi Chen
Department of Information and Telecommunications Engineering, Ming Chuan University, Taoyuan, Taiwan.
chingyi@mail.mcu.edu.tw

2nd Che-Wei Chang
Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan.
yc09264fgg@gmail.com

3rd Zih-Ching Chen
Department of Economics, National Taiwan University, Taipei, Taiwan.
virginia.chen2007@gmail.com

***Abstract*—Approximate computing is a way to help computer system greatly improve computing efficiency. Compared with traditional computer system's computing mode, the approximate computation of computers may be able to accomplish more tasks under the same resource consumption. In general, DSP hardware architecture requires using a large number of floating-point operations and multiplier, which will cost a large amount of hardware resources; using fixed-point arithmetic implemented in hardware enables the DSP algorithm to processing the constant multiplication simultaneously. However, on the other hand, this can affect the accuracy of the calculation results. Facing these hardware circuit design problems, this research attempt to realize the Probabilistic Neural Network (PNN) hardware architecture of approximate calculation with using genetic algorithm (GA). Considering hardware resource consumption and computing speed, we hope that by sacrificing the precision of operational results to reduce the hardware complexity of the PNN hardware circuit.**

Keywords—Approximate computing, floating-point, fixed-point, Probabilistic Neural Network, genetic algorithm

I. Introduction

The purpose of Approximate Computing is to improve the computing efficiency of the system by eliminating some minor computation. Approximate computing is a way that might help computer systems improve computing performance. For some systems, even if the imprecision of the result, such inaccuracy will not affect users' final judgment or perception. Take the algorithm of machine learning as example, the reduction of accuracy in the process can be tolerated as long as the final classification result is correct. In this sense, there is no dramatic difference between an imprecise output and an exact output. When implementing DSP algorithms in hardware, the floating-point multiplication and division often appear in the algorithm, are the main reasons that cause the calculation time-consuming. These modules are also significant at consuming hardware resources when it comes to hardware. To improve the performance of hardware and reduce the cost of hardware resources, we often convert part of a floating-point operation to a fixed-point operation. However, using different bits for a fixed-point operation will affect both the cost of hardware resources and the accuracy of the system. In general, we can reduce the cost of hardware resources by approximating computations while maintaining acceptable accuracy.

For an evolvable hardware design problem, the most basic standard of performance evaluation is the correctness of the circuit output and cost both hardware resources. Our efforts to achieve is to make an efficient compact circuit approximate with the original algorithm of calculation results while ensuring the precision conditions. In this study, we use Genetic Algorithm (GA) to realize approximate calculation of the hardware circuit of Probabilistic neural network (PNN)[1], the main purpose is to achieve the best balance between maintaining good classification ability and the least hardware resource consumption.

II. Materials and Methods

A. *Probabilistic neural network*

Probabilistic neural network (PNN) is proposed by Specht [1], which is a supervised learning neural network based on Bayesian decision rule. Its network architecture is divided into four layers, including input layer, hidden layer, summation layer and output layer, as shown in Fig.1.

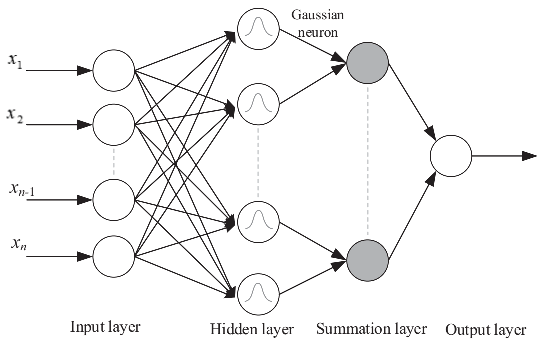


Fig. 1. The structure of PNN model

The input layer represents each feature of the input vector as a node. In hidden layer, the probability density function of the *i*th neuron is the following form

$$f_{Ai}(X) = \exp\left(-\frac{(X-Y_{ai})^t (X-Y_{ai})}{2\sigma^2}\right) \qquad (1)$$

where X is the input vector, Y_{ai} is the *i*th pattern from class C_A, σ is the smoothing factor of Gaussian function.

The third layer is the summation layer. In the summation layer, the probability density function $P_A(X)$ of class C_A is defined by

$$P_A(X) = \frac{1}{m}\sum_{i=1}^{m} \exp\left(-\frac{(X-Y_{ai})^t (X-Y_{ai})}{2\sigma^2}\right) \qquad (2)$$

where m is the number of training patterns from class C_A.

The output layer unit classifies the input vector in accordance with the Bayesian decision rule based on the output of all the summation layer neurons.

B. *Optimized design of PNN hardware using GA*

1) *Basics of GA algorithm*

GA algorithm [2] was proposed by Prof. J.Holland in the United States in 1975. The key concept of GA is to imitate the natural evolution law of natural selection in nature and to solve the optimization problem. There are three main operators in the GA algorithm: reproduction, crossover, and mutation. Firstly, encode all the parameters into chromosomes, and define fitness function. The evolution started from the population of completely random individuals, evaluated the adaptability of each chromosome to the environment in each iteration process, and then generate the new population through natural selection and mutation. Repeat until the final break conditions are met.

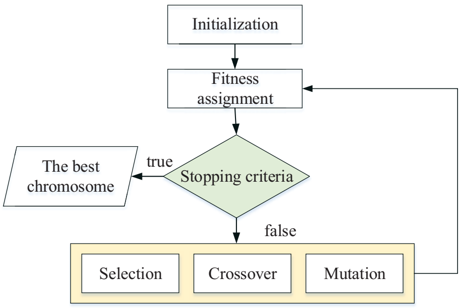


Fig. 2. The flowchart of GA

2) *The proposed method*

The hidden layer neurons of PNN are responsible for the computer rate density function, which performs the nonlinear transformation from the input space to the hidden layer. The weight vector of hidden layer neurons represents a training pattern, and the probability density function is a Gaussian function in multidimensional feature space, which is a nonlinear function. Such nonlinear functions are often implemented on hardware. In addition, the Gaussian function is decided by a smoothing coefficient of σ its distribution scope. The larger the σ, the wider the breadth, and the smaller σ, the narrower the breadth. When the input vector is near the center of the Gaussian function, the hidden layer node will produce a larger output. In practical engineering, the look-up table is often used to approximate these nonlinear functions. In this study, the number of bits encoded by the smoothing parameters and probability values of PNN is used as the gene encodes of each individual in GA, and the recognition rate of the PNN classifier is used as the fitness function, using GA to optimize the parameters to obtain the circuit structure with both the correct rate and the low memory resource consumption.

III. RESULTS

In order to reduce the amount of logic gate usage, we store the Gaussian probability values in the memory, so the resource usage is expressed in memory units (bits). In order to understand the impact of different indicators on the system, we conducted two experiments with IRIS dataset separately. Table I shows the effect of different smoothing parameters on the amount of memory used in the case of the same number of bits (7 bits) and the same accuracy of classification. It can be known that the smaller the smoothing parameters, the lower the amount of hardware resources usage is. Table II shows the influence of different bits on the usage and accuracy of memory after obtaining the best smoothing parameters, reducing the number of data bits while also reducing the number of logic gates of synthesis register.

TABLE I. EFFECTS OF DIFFERENT SMOOTHING PARAMETERS ON MEMORY USAGE

smoothing parameter	25.283	22.412	20.096	18.763	11.571
No. of probability values	7,077	5,561	4469	3,879	1,482
Memory usage (bits)	49,539	38,927	31,283	27,153	10,374

TABLE II. THE EFFECT OF DIFFERENT BITS ON MEMORY USAGE

	16 bits	8 bits	4 bits	3 bits	2 bits
The best smoothing parameter	396.509	18.275	4.860	2.528	2.238
Classification accuracy	0.962	0.962	0.962	0.971	0.962
No. of probability values	3,705,185	4,164	160	33	17
Memory usage (bits)	59,282,960	33,312	640	99	34

Under the condition of ensuring classification accuracy, after using GA to search the optimal solution set of the best parameters of the PNN hardware architecture, that is, the smoothing coefficient σ and the number of encoded bits of the probability value, and implementing the circuit on Altera MAX 10 device, the circuit module with simplified circuit and high performance can be obtained. The hardware resources used are as shown in Table III. Fig.3 and Fig.4 are the waveform simulation results of the partial operation. The number in the red line is the classification test result of inputting a feature vector. When the stimulation clock frequency is set to have 2 ns clock period with 1 ns high and 1 ns low, the circuit outputs the classification result at 373 ns. In conclusion, 186 clock cycles are required to complete a classification process.

TABLE III. THE RESOURCE CONSUMPTION OF PNN HARDWARE CIRUCIT DESIGN FOR IRIS CLASSIFICATION PROBLEM.

FPGA Device	Altera MAX 10 10M50DAF484C7G
EDA Tool	Altera Quartus
Total Logic Elements	97 / 49,760 (<1%)
Total Register	55
Total Pins	11 / 360 (3%)
Total memory bits	2,560 / 1,677,312 (<1%)
Performance (Slow 1200mV 85C Model)	144.68 MHz
Performance (Slow 1200mV 0C Model)	159.85 MHz

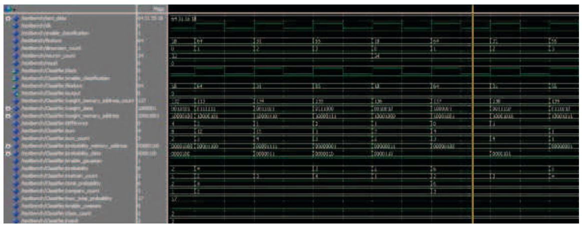


Fig. 3. Waveform simulation result

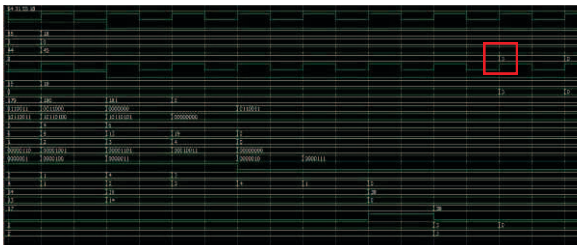


Fig. 4. Simulation of classification result

IV. DISCUSSION AND CONCLUSIONS

This study introduces a method for implementing the approximate calculation of PNN hardware circuits using GA. While ensuring that the correct rate is not affected, we use GA to search for the optimal parameters of the PNN and establish a look-up table method for nonlinear functions to simplify the complexity of the hardware architecture, reduce the use of logic gates, and increase the operation speed.

ACKNOWLEDGEMENTS

This research was supported by the Ministry of Science and Technology of the Republic of China (Taiwan) under Contract No. MOST 107-2221-E-130 -008.

REFERENCES

[1] Specht, D. F. "Probabilistic neural networks". *Neural Networks*. 3: 109–118, 1990.

[2] Eiben, A. E. and Smith, J.E., *Introduction to Evolutionary Computing*, Springer, First edition, 2003

Novel Subblock Partitioning for PTS Based PAPR Reduction of OFDM Signals

Miin-Jong Hao
Department of Computer and Communication Engineering
National Kaohsiung University of Science and Technology
Kaohsiung, Taiwan
miinhao@nkust.edu.tw

Hong-Han Yao
Department of Computer and Communication Engineering
National Kaohsiung University of Science and Technology
Kaohsiung, Taiwan
0151812@nkust.edu.tw

Shou-Sheu Lin
Department of Computer and Communication Engineering
National Kaohsiung University of Science and Technology
Kaohsiung, Taiwan
ssl@nkust.edu.tw

***Abstract*—Three novel subblock partitioning schemes for partial transmit sequence (PTS) based peak-to-average power ratio (PAPR) reduction of Orthogonal Frequency Division Multiplexing (OFDM) signals are proposed. The Geometric Series Division (GSD), the Bidirectional Geometric Series Division (BGSD), and the Quadratic Polynomial Partition (QPP) schemes all have the PAPR performance close to that from the random partition schemes but with much lower computational complexity. Numerical simulations are carried out to verify that the proposed methods can effectively solve the problem of high complexity and to reach the required PAPR reduction performance.**

Keywords*—*PAPR, Subblock Partition, PTS-OFDM, GSD, BGSD, QPP

I. INTRODUCTION

To reduce large PAPR of OFDM signals, various schemes for reducing the PAPR have been presented [1]. Among these existing schemes, the partial transmit sequence (PTS) scheme is very promising due to its good PAPR reduction performance without no signal distortion [2]. The random partition scheme has the best effect on PAPR reduction in PTS techniques but comes with a very high computational complexity. Also, it requires a great amount of computations to search for the optimal weighting phases [3]. In this paper, three low-complexity subblock partition schemes, i.e. the Geometric Series Division (GSD), Bidirectional Geometric Series Division (BGSD), and Quadratic Polynomial Partition (QPP) [4] are proposed for achieving the same performance in PAPR as the pseudo random schemes. Each scheme is presented by an illustration. Simulations are performed to verify their advantages.

II. PTS-BASED PAPR REDUCTION TECHNIQUES

A. Partial Transmit Sequence (PTS) OFDM Method

Let the data block $\bar{\boldsymbol{X}} = [X_0, X_1, \cdots, X_{N-1}]^T$ denote the collection of all data modulation symbols X_k , $k = 0,1, \cdots, N-1$. The OFDM signal x_n consisting of N subcarriers is given by

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k\, e^{j\frac{2\pi kn}{N}}, 0 \le n \le N-1. \qquad (1)$$

The PAPR value of OFDM signals can expressed by

$$PAPR \triangleq \frac{max|x_n|^2}{E[|x_n|^2]}, 0 \le n \le N-1 \qquad (2)$$

In the PTS-OFDM scheme, the input data sequence $\bar{\mathbf{X}} = [X_0, X_1, \cdots, X_{N-1}]^T$ with N subcarriers is partitioned into N disjoint subblocks $\bar{\mathbf{X}}^{(v)}$, where $\bar{\mathbf{X}}^{(v)} = \left[X_0^{(v)}, \cdots, X_k^{(v)}, \cdots X_{N-1}^{(v)}\right]^T$, $v = 1,2, \cdots, V$, $k = 0,1, \cdots, N-1$. Data assigned to the element $X_k^{(v)}$ according to different schemes. Each subblock $\bar{\mathbf{X}}^{(v)}$ is multiplied by a phase factor $p^{(v)}$ and then added together. The PTS operation can be expressed by

$$\bar{\mathbf{x}} = \sum_{v=1}^{V} IFFT\{p^{(v)}\bar{\mathbf{X}}^{(v)}\} = \sum_{v=1}^{V} p^{(v)}\bar{\mathbf{x}}^{(v)}, \qquad (3)$$

where $v = 1,2, \cdots, V$, $p^{(v)} = e^{j\phi^{(v)}}$, $\phi^{(v)} \in [0,2\pi)$ are the phase of the phase factor, $\bar{\mathbf{x}}^{(v)}$ $v = 1,2, \cdots, V$ are the partial transmit sequences. The objective of phase optimization is to select a proper phase factor so that an optimal $p^{(v)}$ with the smallest PAPR value is chosen. The algorithm is given by [2]

$$\underset{\{p^{(1)},p^{(2)},\cdots,p^{(V)}\}}{\operatorname{argmin}} \left(\max_{0\le n\le N-1}\left|\sum_{v=1}^{V} p^{(v)}\, x_n^{(v)}\right|\right) . \qquad (4)$$

The complexity of the searching procedure increases exponentially with the number of subblocks.

B. The Geometric Series Division (GSD) and Bidirectional Geometric Series Division (BGSD) Partition Schemes

The GSD method allocates data in disjoint subblocks according to the geometric sequence

$$\{ar^{m-1}\}_{m=1}^{M} , \qquad (5)$$

where $r \ne 0$ is the common ratio, m the positive integer index, and a a scale factor of the series. It is usually to choose the ratio value of $r = 2$, and the scale factor of $a = 1$ for this application. Each data symbol is assigned to the subcarrier index pointed by the index m for $m = 0,1, \cdots, M_{GSD}$, where $M_{GSD} = \log_2(N/V) - 1$. An illustrated example of the GSD partition scheme with $N = 32$ subcarriers and $V = 2$ subblocks is shown in Fig. 1.

The BGSD method combines two GSD sequences in bidirectional way of ascending and descending as follows:

$$\{ar^{m-1}\}_{m=1}^{M/2} , \{ar^{M-m}\}_{m=M/2}^{M} . \qquad (6)$$

Each data symbol is assigned by the index m for $m = 0,1, \cdots, M_{GSD}, M_{GSD}, \cdots, 1,0$. An illustrated example of the BGSD scheme with $N = 32$ and $V = 2$ is shown in Fig. 2.

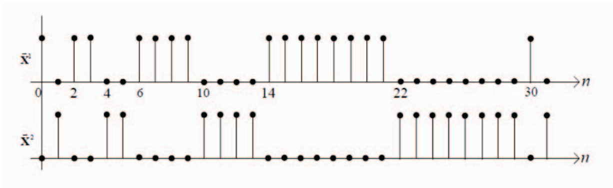


Fig. 1 An example of GSD partition scheme

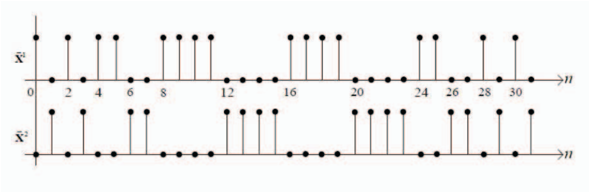


Fig. 2 An example of BGSD partition scheme

C. The Quadratic Polynomials Partition (QPP) Scheme

The proposed QPP method assigns data. in disjoint subblock according to the QPP interleaver based index. The QPP interleavers have been presented in [4]. The idea of the Quadratic Polynomial Partition interleaver is to generate a sequence by the generator.

$$I(n) = (q_1 n + q_2 n^2)\,\mathrm{mod}(N), \qquad (7)$$

where $I(n)$ is the index pointing to the new sequence locations, n is the locations to the original sequence elements, $n = 0,1,\cdots,N-1$,and N is the length of the interleaver. It follows that the element can be expressed:

$$X_k^{(v)} = \begin{cases} X_k, & k = I(n), \quad (v-1)N/V \le n < vN/V \\ 0, & \text{otherwise} \end{cases}. \qquad (8)$$

For different numbers N of subcarriers, the corresponding values of the two parameters q_1 and q_2 are various, as given in [4]. An illustrated example of the QPP scheme with $N = 40$, $V = 2$, $q_1 = 3$, and $q_2 = 10$ is shown in Fig..3.

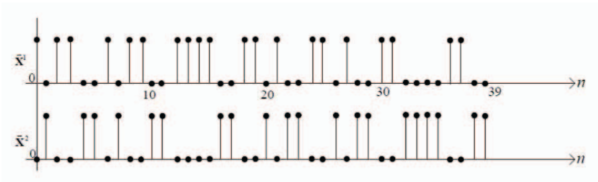


Fig. 3 An example of QPP based index partition scheme

III. Simulation Results

Simulations are carried out for PAPR reduction performance of PTS-OFDM with the three proposed partition methods. The simulation parameters are listed in Table I.

Table. I Simulation Parameters.

Parameter name	Value	
Number of symbols	10^5	
Number of subcarriers	128	512
Modulation types	QPSK	64QAM
Number of sub-blocks	4	8
Phase weighting	{1, -1}	

The PAPR reduction performance of the PTS OFDM for various partition methods is shown in Fig. 4 and Fig. 5, where the cases of QPSK with N=128, $V = 4$ and 64QAM with $V = 8$, N=512 at $CCDF = 10^{-3}$ are given respectively. It can be seen that the proposed three schemes have better PAPR reduction performance compared to the traditional methods and almost have the same performance as the pseudo-random methods. However, the computational complexities of the proposed methods are much lower than the pseudo-random partition schemes. In general, the performance can be improved by increasing the number of subblocks.

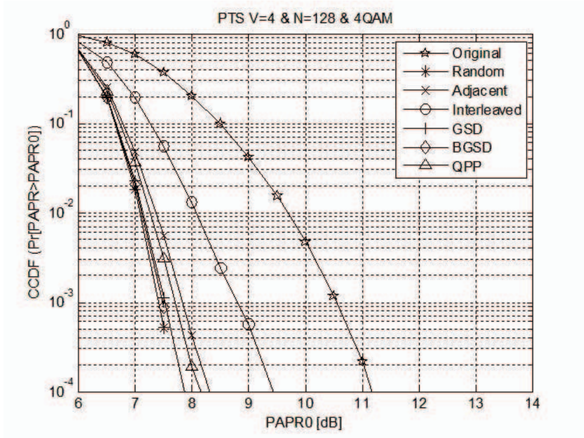


Fig. 4 PAPR reduction performance of PTS-OFDM with QPSK, $N = 128$, $V = 4$ for various partition schemes.

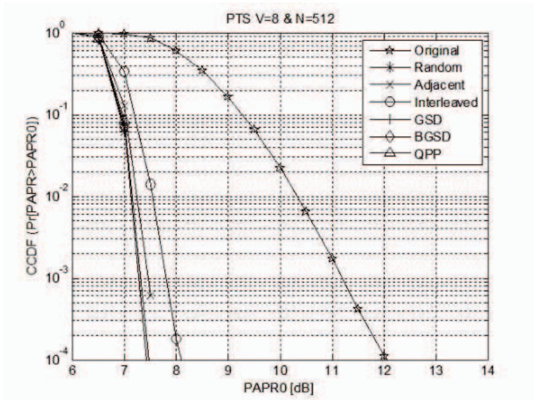


Fig. 5 PAPR reduction performance of the PTS-OFDM with 64QAM, $N = 512$, $V = 8$ for various partition schemes.

IV. Conclusions

Three effective partition schemes for PAPR reduction of PTS-OFDM are proposed. The GSD and BGSD methods apply a special geometric series for data assignment. The QPP partition scheme assigns data based on a fast pseudo random sequence generator. All the proposed schemes can reach the PAPR reduction performance around that of pseudo-random partition schemes at low complexities.

References

[1] T. Jiang, and Y. Wu, "An overview: peak-to-average power reduction techniques for OFDM signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp.257-268, Jun. 2008.

[2] C. Ye, Z. Li, T. Jiang, C. Ni, and Q. Qi, "PAPR reduction of OQAM-OFDM signals using segmental PTS scheme with low complexity," *IEEE Trans. Broadcast.,* vol. 60, no. 1, pp. 141-147, Mar. 2014.

[3] C.-X, Ni, J. Tao, and P. Wei, "Joint PAPR reduction and sidelobe suppression using signal cancelation in NC-OFDM-based cognitive radio systems," *IEEE Transaction on Vehicular Technology*, vol. 64, no. 3, March 2015.

[4] O. Y. Takeshita, "Permutation polynomial interleavers: an algebraic-geometric perspective," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp.2116–2132, Jun. 2007.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Famileaf: Flowerpot robot for dementia prevention

Nagisa Ishizumi and Manabu Gouko
Graduate School of Engineering
Master's program in Mechanical Engineering, Tohoku Gakuin University
Miyagi Prefecture, Japan
s1894102@g.tohoku-gakuin.ac.jp

***Abstract*—In Japan, as the rate of dementia increases, the interest in preventing dementia also increases. In this paper, we propose a flowerpot-type robot system called Famileaf. Growing a plant using Famileaf is expected to prevent dementia. The robot comprises a flowerpot, microcomputer, multiple sensors, and a battery. Famileaf can post a comment on a social networking site (SNS) to interact with an elderly person who is a grower. The users of the robot can grow plants in the same manner as they take care of a pet. Because the users can grow plants with more attachment, the use of Famileaf is expected to prevent dementia. Herein, the concept and the prototype of Famileaf have been described.**

***Index Terms*—Famileaf, Dementia prevention**

I. Introduction

A survey conducted by the Cabinet Office of Japan [1], estimates that the number of people suffering from dementia aged 65 or above will be approximately 7 million in 2025 (comprising approximately 20% of the population of this age group). Furthermore, the number of carers needed is increasing. Therefore, efforts to prevent dementia are urgently required for patients with dementia in their homes as well as assisted living residences.

It has been reported that, the activation of brain functions through activities such as reading, writing, conversing, physical exercise, and improving eating habits are effective for preventing dementia [2]. Hence, various activities for residents are performed in assisted living residences in an effort to prevent dementia.

Herein, we aim to develop a flowerpot-type robot system called Famileaf that combines a planter with plants, sensors, etc. Famileaf enables its users to grow plants with more attachment, which is expected to prevent dementia. This study introduces the concept and prototype of Famileaf.

II. Conventional approach to dementia prevention

In assisted living residences, at present, various efforts are undertaken for preventing dementia. This chapter describes the methods based on the measure of dementia prevention, particularly the interaction of an elderly person with an animal, robot, and plant.

Generally, interaction with animals is considered beneficial for dementia prevention. In assisted living residences, animal therapy, wherein animals such as dogs and cats are in contact with each other, is implemented to this end; it has been reported to a positive mental and physical impact on participants [3]. However, in animal therapy, there are problems related to the hygiene and time-consuming training of animals.

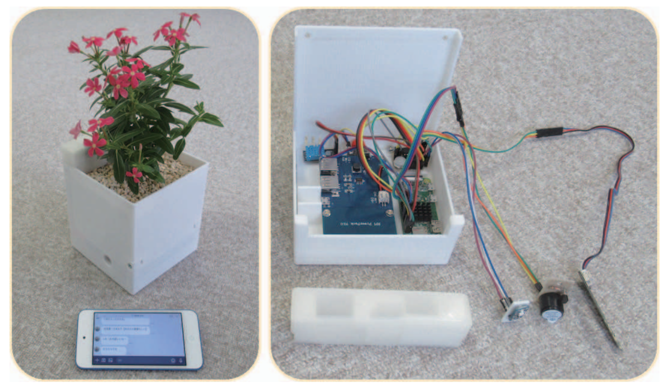


Fig. 1. The prototype of the flowerpot robot called 'Famileaf' (left) and its internal structure (right).

In recent years, robot therapy that uses an robot instead of a real animal has been proposed [4]. Although robot therapy can solve the problems associated with animal therapy [5], [6], the expense of a robot for the therapy is high, which is a major obstacle for its wide application.

Plant therapy has been implemented; it is expected to foster the user's sense of responsibility and attachment to his/her plant through cultivation. Further, plant therapy is beneficial for motor function maintenance [7]. However, unlike with animals and robots, it is difficult for the grower to immediately interact with plants during cultivation. A grower often becomes tired of growing a plant and therefore fails to maintain it.

Here, the following is an example of an immediate interaction is explained. For example, when a dog owner feeds a hungry dog, the owner can immediately confirm the dog's happy reaction about receiving the food. Hence, there is an immediate response to one's own behavior, and confirming this response is called immediate interaction in this research.

III. Flowerpot-type robot

We propose a flowerpot-type robot system called Famileaf to realize the immediate interaction between growers and plants. Famileaf is a planter equipped with a micro¬ computer and sensors that are used while growing plants (Fig. 1). Based

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

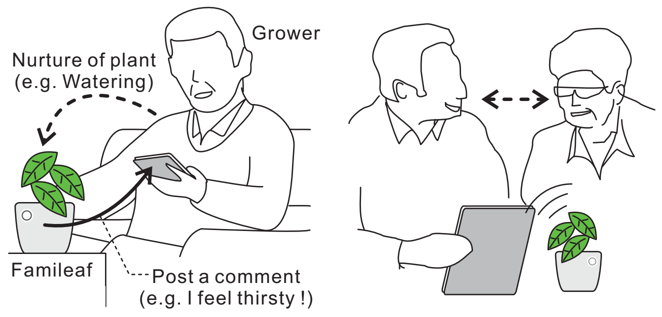


Fig. 2. Interaction of Famileaf and the grower (left). Encouragement of conversation through Famileaf (right).

on sensor output, the robots post various comments on a social networking service (SNS). Famileaf is intended for use in environments wherein multiple subjects of dementia prevention reside, such as assisted living residences.

A. System overview

Famileaf has functions such as nurturing and situation feedback that enables immediate interaction. The details of each function are described below.

The nurturing and situation feedback function is used to generate a greater sense of responsibility and attachment in users (Fig. 2 (left)). The function detects changes in the grower's care and its environment by means of sensors and accordingly posts comments on SNS, e.g., wanting water when the soil dries up and thanking when the plant receives water. Such a process enables immediate interaction with plants, which is similar to an interaction with animals and robots. Additionally, to prevent user boredom, it posts meaningless comments. Additionally, Famileaf sounds a buzzer when posting a comment; the aim is to make users aware that the plants need attention because producing sounds like a pet is also expected to create more attachment.

Famileaf is expected to solve the problems by using the advantages of the conventional dementia prevention methods described in the previous chapter. Additionally, because this robot does not require complicated components like actuators, it can be realized at a relatively low cost. Therefore, the hurdles of operation in assisted living residences are lowered and widespread application of Famileaf is expected.

B. Prototyping of Famileaf

Till date, the hardware for potted robots has been created and the nurturing and situation feedback function has been implemented. The hardware is designed using 3D CAD and created using a 3D printer. The weight of the potted robot is about 1 kg including soil and plants, and the size is W120 mm $\times$ D120 mm $\times$ H150 mm.

It comprises a microcomputer (Raspberry Pi ZERO W), multiple sensors, and a battery. A soil humidity sensor (SEN0114, DFRobot) and ultraviolet sensor (GUVA-S12SD 3528, Keyestudio) are installed to monitor plant conditions. Temperature and humidity sensors (DHT11, Aosong Guangzhou Electronics) monitor the environment where the flowerpot robot is placed. It costs only one hundred dollars to produce this prototype.

To implement the nurturing and situation feedback function, the microcomputer submits messages to the SNS based on sensors' output. There are two categories of messages: one is about the demands of the plant and the other is about consideration for the user. For example, when a plant needs water, the comment "I feel thirsty!" is posted on the SNS.

IV. Related Approaches

Over the past few years, studies have been conducted on engineering systems about growing plants [8], [9], [10], [11].

Among others, the flowerpot-type robot system has been proposed [11]. The system can create relaxed communication with a third person in a public space. This system allows the users to indirectly recognize third persons.

Our proposed robot also aims at fostering direct conversation between growers with the viewpoint of dementia prevention (Fig. 2 (right)). Generally, direct communication involves significant psychological hurdles. However, robot users are assumed to belong to the same facility (such as an assisted living residence), so the psychological hurdles in communication are considered relatively low.

V. Summary and Future Issues

This research proposes a flowerpot-type robot system called Famileaf, and its application is expected to prevent dementia. Herein, the concept and prototype of Famileaf have been explained.

In future, a field study will be conducted to verify the effectiveness of Famileaf.

References

[1] HP of Japanese cabinet office (http://www8.cao.go.jp/kourei/whitepaper/w-2016/html/gaiyou/s1_2_3.html), in Japanese.

[2] K. Nishino, " The Advance of Dementia Prevention and Dentistry, " Ronen Shika Igaku, vol.29, no.3, pp.278-281, 2015, in Japanese.

[3] Y. Futoyu, H. Kobayashi, H. Nagase, T. Ikenaga, " Usefulness of dog-assisted therapy in elderly patients with dementia, " Kawasaki medical welfare journal, vol.17, no.2, pp.353-361, 2008, in Japanese.

[4] K. Wada, T. Shibata, " Robot Therapy: A New Approach for Mental Healthcare of the Elderly ? A Mini-Review, " Gerontology, vol.57, no.4, pp.378-386, 2011.

[5] K. Wada, T. Shibata, " Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house, " IEEE Transactions on Robotics, vol.23, no.5, pp.972-980, 2007.

[6] M. Kanoh, T. Shimizu, " Developing a Robot Babyloid That Cannot Do Anything, " Journal of the Robotics Society of Japan, vol.29, no.3, pp.76-83, 2011, in Japanese.

[7] S. Koura, " Present Status of Horticultural Therapy and Future Possibilities in Japan, " Horticultural Research (Japan), vol.12, no.3, pp.221-227, 2013, in Japanese.

[8] M. Kinoshita, N. Ezaki, " Uekimochi; that communicates with Plants, " IPSJ SIG Technical Report, CDS-vol.5, no.11, pp.1-7, 2012, in Japanese.

[9] F. Sawaki, K. Yasu, M. Inami, " A Specialized Actuation Method for Expression of Emotion by Plants " IPSJ Interaction 2012, pp.617-622, 2012, in Japanese.

[10] T. Nishida, S. Owada, " MOEGI: Plant Fostering by the Assistance of Augmented Reality, " 14th Workshop on Interactive Systems and Software (WISS), pp.23-26, 2006, in Japanese.

[11] T. Yoshino, T. Yamanaka, " OSHABERI-HACHIBE: Messaging Bot System Aiming Encouraging a Sense of Affinity, " IPSJ SIG Technical Report, GN-vol.74, no.14, pp.1-6, 2010, in Japanese.

An Enhanced Method for Turbo Code Interleavers

Shou-Sheu Lin
Computer and Communication Engineering
National Kaohsiung Univ. of Science and Technology
Kaohsiung, Taiwan
ssl@nkust.edu.tw

Sun-Ting Lin
Electronic Engineering
National Kaohsiung Normal University
Kaohsiung, Taiwan
stl@nknu.edu.tw

Yu-Cheng Li
Computer and Communication Engineering
National Kaohsiung Univ. of Science and Technology
Kaohsiung, Taiwan
9851814@nkust.edu.tw

***Abstract*—Design of interleaver for turbo codes to achieve excellent BER performance is a very difficult task. The brute-force searching method is not only time-consuming but also impractical due to extreme large of search space. To reduce the design complexity while maintaining the performance, we propose a bubble-sort like method to generate the new interleaver candidate which is a swapped sequence from a baseline interleaver. The distance spectrum is used as the BER evaluation tool to select the best interleaver from a limited search space. Simulation results show that the proposed design method can improve the performance effectively for generic interleavers.**

Keywords—turbo code; interleaver; distance spectrum

I. INTRODUCTION

Turbo code [1] is a capacity-approaching error correcting code and its bit error rate (BER) performance is close to the theoretical Shannon bound. Once the polynomials of the recursive convolutional code are selected, the permutation sequence of the interleaver will be an important factor of the performance of a turbo code. The purpose of the interleaver is to reduce the correlation between information bits and prevent burst errors caused by decoding error propagation. Some generic interleavers with good performance are used in turbo codes such as uniform random interleaver (URI) [2] and S-random interleaver (SRI) [3]. Some improved methods for interleaver design are proposed [4][5]. The distance spectrum (DS) [2] analysis provides BER prediction with reasonable computational time instead of time-consuming Monte Carlo method. The difficulty of designing an interleaver is the extreme large of search space for all possible interleaver candidate, which is proportional to the factorial of information length. For example, if the information length is 100, the total number of possible interleaver is 100!. In this paper, an enhanced algorithm for generic interleaver design with reduced design complexity is proposed.

II. ENHANCED GENERIC INTERLEAVER ALGORITHMS

The concept of the proposed algorithm is to swap two positions of an interleaver to generate a new interleaver, which may perform better than the original one. The swapped position is similar to the procedure in the bubble-sort algorithm. Therefore, the search space is reduced. The detail flow chart is shown in Fig. 1. It can briefs as follows. First, given a baseline interleaver $\pi(x)$ such as URI. Set the maximum number of successful enhancement K to avoid unlimited search, i is the index used as the counter for successful enhancement. N is the length of interleaver. The j is the swap base index which is similar to base index in the bubble-sort algorithm. The initial value of both i and j is zero. There are $N-1$ batch of try. Each batch of try will generate $N-j-1$ of new interleavers which are the candidates for performance evaluation. The new interleaver is to swap two position similar to the bubble-sort algorithm. Note that we always do the swap no matter what value is in these two positions . Then use DS analysis [2] in (1) to evaluate the BER as follows

$$P_n = \frac{N_{d_{free}} \bar{W}_{d_{free}}}{N} Q\left(\sqrt{d_{free} \frac{2RE_b}{N_0}}\right) \qquad (1)$$

where P_n is the low bound of the BER for each interleaver which is denoted as $\pi(j) \leftrightarrow \pi(j+n)$.

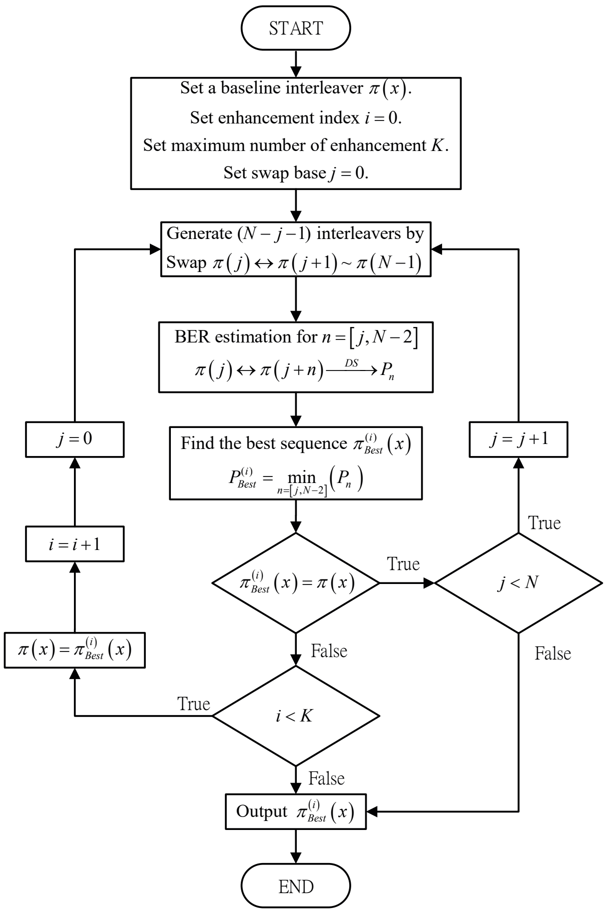


Fig. 1. Flow chart of the proposed enhanced algorithm.

Compare the BER of these interleavers as in (2) and the ith best interleaver $\pi_{Best}^{(i)}(x)$ is the interleaver with lowest BER.

$$P_{Best}^{(i)} = \min_{n=[j,N-2]} \left(P_n \right) \qquad (2)$$

If no interleaver is better than baseline or previous one, then try the next batch of interleavers by increasing swap base index j by one. If a better interleaver is found, then record the baseline interleaver as the new ith best interleaver and go to the next batch of try. Repeat the mentioned procedure until all the interleaver candidates in search space is finished. Obviously, the complexity of the enhanced algorithm is much lower than the brute-force method. The search space size of the proposed algorithm is $\frac{N(N-1)}{2}$, which is much smaller than $N!$.

III. Simulation Results

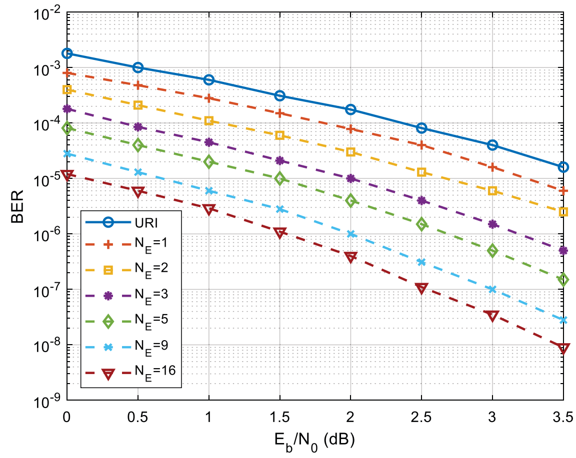
In order to demonstrate the performance of the proposed enhanced algorithm, simulations were conducted. The BER performance are evaluated by distance spectrum analysis. The simulation parameters are set as follows. The turbo code is the (13, 15) code with zero-state trellis termination. The data block length is 256 and the code rate is 1/3. AWGN channel is used and N_b / N_0 ranges from 0 to 3.5 dB. The enhanced algorithm is applied to the URI and SRI interleavers. The maximum number of successful enhancement K is set to 20.

Define N_E as the N_E-th successful enhancement for a given baseline interleaver. The BER performance for the URI as the baseline interleaver and its enhanced interleavers are shown in Fig. 2(a). For SRI interleaver, the results are shown in Fig. 2(b). The proposed enhanced algorithm can find enhanced interleaver successfully for both URI and SRI interleavers. URI finds 16 enhanced interleavers and SRI finds 19 interleavers, respectivly. For clarity, the figures do not show all the BER curves of enhanced interleavers. At the 16-th enhancement URI achieve two order of magnitude of BER improvement.

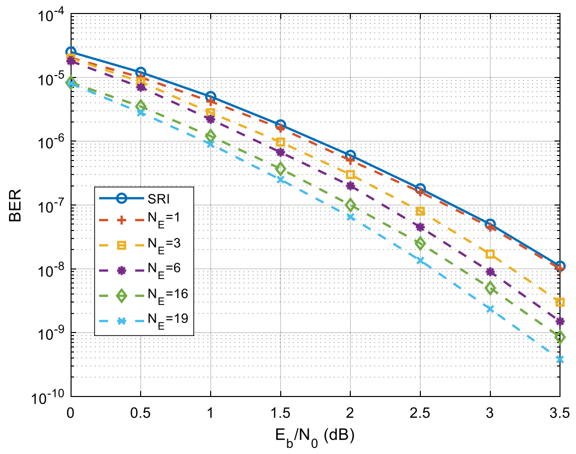
It is well known that the BER performance of the URI interleaver is poor than the SRI. However, the improvement is more obvious for the URI after enhancement because there are much improvement margin than SRI. We also calculate the free distance d_{free} of each enhanced interelavers in Table I to analysis the improvement. Obviously, the free distance d_{free} becomes larger and larger compared to each baseline interleavers. At 16-th successful enhancement, the d_{free} difference of URI is 7 and SRI is 4, respectively. That is the reason why the improvement of SRI is not as apparent as the URI.

TABLE I. Free Distance of The Enhanced Interleavers Based on URI and SRI

Baseline interleaver	N_E					
	0	3	6	9	16	19
d_{free} (URI)	9	12	13	16	16	-
d_{free} (SRI)	16	18	20	20	20	22



(a)



(b)

Fig. 2. BER performance for various generic interleavers. (a) URI, (b) SRI.

IV. Conclusions

A performance enhancement interleaver design method are proposed for the generic interleaver. Using distance spectrum as the BER performance prediction tool, these methods provide an systematic way to generate finite interleaver sequences candidates so that the design complexity is reduced dramatically. Performance improvements are obtained for all the considered interleavers.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Commun. Conf.*, pp. 1064-1070, Geneva, Switzerland, May 1993.

[2] L. C. Perez, J. Segher, and D. J. Costello, "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1698-1709, Nov. 1996.

[3] D. Divsalar and F. Pollara, "Turbo codes for PCS applications," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, vol. 1, pp. 54-59, Jun. 1995.

[4] G. M. Kraidy, "On Progressive Edge-Growth Interleavers for Turbo Codes," *IEEE Comm. Let.*, vol. 20, no. 2, pp.200–203, Feb. 2016.

[5] J. J. Boutros and G. Zemor, "On quasi-cyclic interleavers for parallel turbo codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1732–1739, Apr. 2006.

Performance Analysis of Power Outage Probability for Drone based IoT Connectivity Network

Sarun Duangsuwan, Anna Chusongsang[*,1] and Sathaporn Promwong[*,2]
[1]Information Engineering, Department of Engineering
King Mongkut's Institute of Technology Ladkrabang Prince of Chumphon Campus
[*]Telecommunication Engineering, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
1 Chalongkrung Rd., Ladkrabang, Bangkok, Thailand
sarun.du@kmitl.ac.th, [1]annaannoy@hotmail.com, [2]kpsathap@kmitl.ac.th

***Abstract*—In this paper, we present a performance analysis of power outage probability of using on drone communicated with an internet of things (IoT) connectivity network. We obtain that there are several IoT devices are connected with a drone as simultaneously. Therefore, the power outage probability (POP) is analyzed based on the verifying of drone height which ensures the best performance in term of connectivity network. Experimental results confirm that this analysis is very useful for drone usage as drone small cell or hot-spot to distribute the transmitting power to coverage all IoT devices.**

Keywords—IoT connectivity network, Drone, air-to-ground communication, Power outage probability, smart farming.

I. Introduction

Internet of Things (IoT) is spreading much faster than the speed at which the supporting technology is maturing. Today, there are tens of wireless technologies competing for IoT and a myriad of IoT device with disparate capabilities and constrains. Moreover, each of many verticals employing IoT networks dictates distinctive and differential network qualities [1].

Drone has been rapidly growing in nowadays. For the surveillance and agriculture, in particular, with their inherent attributes such as mobility, flexibility, and adaptive altitude, drone admit several key potential applications in wireless systems. The key drone challenges such as 3-dimensional deployment, performance analysis, channel modeling, and energy efficiency are explored to support many devices of IoT connectivity. There are several researches described to drone for wireless communications, as related work [2] – [4]. In [2], the control drone for wireless communication was presented where drone has provided the array antenna to connect ground users. In the considered model, the service time is minimized by minimizing the wireless transmission time as well as the control time that is needed for movement and stabilization of the drones. The results also shown that (therein figure 9), in comparison with the fixed-array case, the network spectral efficiency can be improved by 32% while leveraging the drone antenna array. In [3], an amateur drone surveillance system based on the cognitive IoT has been described. Also, in [4], optimal dimensioning and performance analysis of drone based wireless communications has been interestingly proposed. Drone small cell (DSC) is that the mobility base station. The advantage is to serve specific applications like hot-spot, disaster management, etc. In addition, most of research paper consider a service for air-to-ground communication between drone and IoT devices.

Air-to-ground communication is a research challenging for design and implementation of DSC. The authors in [5] introduced the use of DSC to provide a service for air-to-ground devices in the scenario of device centric architecture and circular coverage area. However, there is no consider the impact of multiuser (MU) uncertainly accessible. Moreover, path loss of air-to-ground for low attitude platform in urban environment was investigated in [6], but it is not used with drone. It is important to note that for IoT dense of inter-connectivity network. The power constrain of transmission link is able to supply all IoT devices.

In this paper, we present the performance analysis of power outage probability of air-to-ground communication. By using the deployment geometry, we consider a drone under the control of a ground positioning system verifying the height of the drone to achieve the performance of power outage probability. As illustrated in Fig. 1, we model a drone to coverage all IoT devices in the large-scale area of smart farming. We analyze the performance of power outage probability in different drone height levels.

The rest of paper as following. The performance analysis describes in Section II, and conclusion is given in Section III.

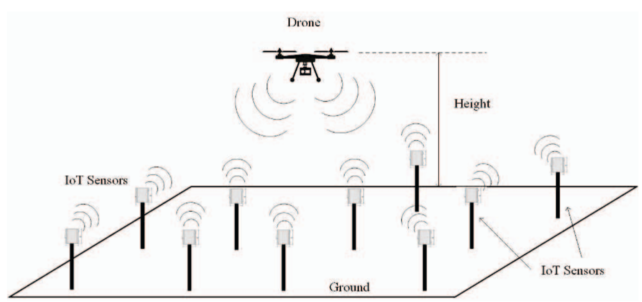


Fig. 1. Model of Drone based IoT connectivity network.

II. PERFORMANCE ANALYSIS

A. Drone to IoT Sensors Link Budget

Consider the link between the drone (hot-spot) and IoT sensor (receiver). The total link budget is expressed in terms of the received signal to noise ratio (SNR) at the receiver antenna that can be written as

$$\gamma_n(h,\phi) = \frac{|\varphi|^2 E_s G_t(\theta,\phi)}{L(h,\phi)N_o} \tag{1}$$

where φ is the multi-path fading operator, E_s is the transmitted symbol energy, N_o is the noise power spectral density of the zero mean complex additive white Gaussian noise (AWGN), $L(h,\phi)$ is the height and angle of arrival (AoA) path loss and $G_t(\theta,\phi)$ is the effective gain of transmitter antenna (drone). In addition, the SNR constrain depends on the number of n IoT sensor.

B. Power Outage Probability

The power outage probability $\hat{P}_{out}(h)$ is defined as the probability at which the SNR value falls below a certain specified of threshold $\hat{\gamma}_{th}$. We consider the Rayleigh fading channel. Therefore, the power outage probability over all the IoT coverage area can be written as

$$\hat{P}_{out}(h) = E\left[1-\exp\left(-\frac{\hat{\gamma}_{th}}{\gamma_n(h,\phi)}\right)\right] \tag{2}$$

where $E[\cdot]$ is an expectation operator [4].

C. Measurement Setup

The experimentally measured setup was conducted in the smart farming of deer farm, Ramkhamhaeng University. Note that the measurement location is in ruzi grass farm where the experiment and dimension can be shown in Fig. 2. The number of IoT devices were installed by 25 sensors. Additionally, WiFi 2.4 GHz was provided to the connectivity between drone and IoT devices. The height of drone was test from 5 m. to 100 m. in line of sight (LOS) communication.

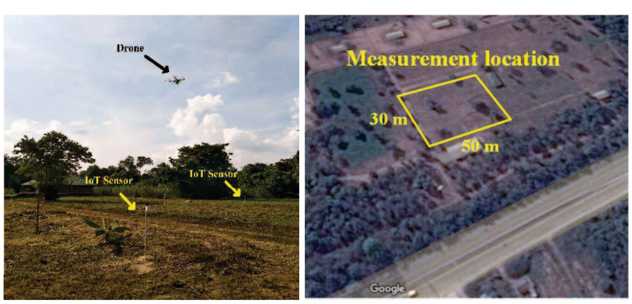


Fig. 2. Experiment and measurement locarion in smart farming.

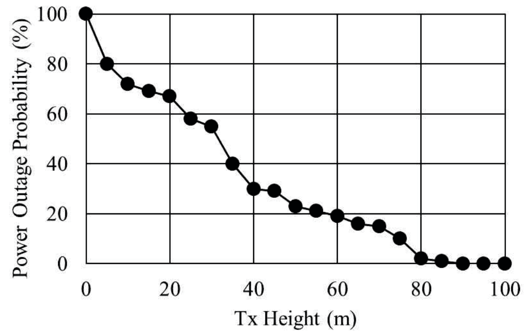


Fig. 3. Power outage propabability compared with Tx height.

D. Result and Discussion

We measure the power constraint of the air-to-ground radio link between Tx antenna (drone) and IoT devices. Fig. 3 shows the power outage probability percentage, in the experiment, Tx power was 10 dBm from the drone. We compared with Tx height from 5 m to 100 m of the experimental test. The percentage of power outage probability is lower than 50 % when drone height was at 30 m. However, we found that the accuracy of air-to-ground communication, in this case, IoT devices cannot link the data streaming to drone when the Tx height is higher than 80 m because the power outage probability is zero.

III. CONCLUSION

In this paper, we have presented the performance analysis of power outage probability based on experimental evaluation between drone small cell or hot-spot and IoT devices. The performance result of power outage probability of air-to-ground communication has been proposed. In this experiment, we performed in the smart farming scenario which has more IoT devices installed. In future work, the path loss analysis of smart farming environment will be discussed.

REFERENCES

[1] M. Ozturk, M. Jaber, and M. A. Imran, "Energy-aware smart connectivity for IoT networks: Enabling smart ports," *International Journal on Wireless Communication and Mobile Computing*, vol. 2018.

[2] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Comminication and control for wireless drone-based antenna array," *IEEE Transaction on Communication*, vol. 67, no. 1, Jan. 2019.

[3] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y. D. Yao, "An amateur drone surveillance system based on the cognitive internet of things," *IEEE Comminication Magazine*, vol. 26, no. 1, Jan. 2018.

[4] A. M. Hayajneh, S. A. R. Zaidi, D. C. Mclernon, and M. Ghogho, "Optimal dimensioning and performance analysis of drone-based wireless commuications," in *IEEE Globecom Workshops*, Washington, DC, USA, Dec. 2016.

[5] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah,, "Drone small cells in the clouds: Design, deployment and performance analysis," in *IEEE Globecom Workshops*, San Diego, CA, USA, Dec. 2015.

[6] A. AI-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platform in urban environment," in *IEEE Globecom Workshops*, Austin, TX, USA, Dec. 2014.

Channel Selection Metric by the Number of Users and SNR in WLAN

Taiga Okutate, Ikuo Oka, Shingo Ata
Graduate School of Engineering,
Osaka City University, 3-3-138 Sugimoto,
Sumiyoshi-ku, Osaka, 558-8585 Japan
Email{okutate@c., oka@, ata@}info.eng.osaka-cu.ac.jp

***Abstract*— In this paper, a channel selection metric is presented for WLAN. It is based on SNR and the number of users. A delay time is employed to define metric of channel selection. Nth-order polynomials are introduced to obtain a polynomial metric in a tractable form. The advantage of formulation is that more parameters other than SNR or the number of users can be added to metric without taking up complex delay analysis. Numerical results are presented to demonstrate metric as a function of SNR and the number of users. The effective range of metric is discussed with a difference between the polynomial and delay time metrics.**

***Index Terms*— SNR, CSMA/CA, Channel Selection, a polynomial metric.**

I. Introduction

A channel selection improves the communication efficiency, which often deteriorates by a change of channel conditions. The channel selection is employed to switch dynamically the channel to the optimal channel using channel state information (CSI). The key parameters representing CSI in WLAN include signal to noise power ratio (SNR), as packet error rate (PER) is dominated by SNR. Another important parameter is the number of users which is closely related with the packet collisions in CSMA/CA. The transmission opportunities decrease as the number of users increases.

For 802.11 a channel model is presented by Bianchi using Markov chain for performance analyses of CSMA/CA [1]. The influence of bit error rate (BER) on the performance of CSMA/CA is evaluated in [2]. Bianchi takes into account the saturation that each user always holds the packet to send. However, the saturation assumption is not in most real 802.11 networks. An analysis method in the non-saturation state in proposed by defining the probability of packet occurrence [3]. With regard to the metric for the channel selection, an analysis formula is derived for a delay time in the non-saturation state [4]. In [5], a delay time is derived considering packet errors with SNR. The analytical results show that a delay time is increased significantly when SNR is low.

II. Channel Selestion Metric

In this paper, metric is presented for a channel selection or an access point selection in WLAN. A delay time [5] is employed as metric. It is defined from the occurrence of a packet to the completion of transmission. Each station first estimates the delay time for each AP, and then selects the AP being expected to provide the minimum delay time. In order to have more tractable metric than the delay time metric, a polynomial metric is proposed. The polynomial metric has a form of Nth-order polynomials, and derived from the delay time metric by a least square method. The first order and third order polynomials are examined for the polynomial metric. The least square method yields the polynomial metric Δ_1 of the first order and Δ_3 of the third order.

$$\Delta_1 = \theta_0 + \theta_1 n + \theta_2 \gamma, \tag{1}$$

$$\theta = \begin{pmatrix} 747.50 \\ 1513.6 \\ -118.57 \end{pmatrix}$$

$$\begin{aligned} \Delta_3 = \theta_0 + \theta_1 n + \theta_2 \gamma + \theta_3 n^2 + \theta_4 n\gamma + \theta_5 \gamma^2 \\ +\theta_6 n^3 + \theta_7 n^2\gamma + \theta_8 n\gamma^2 + \theta_9 \gamma^3, \end{aligned} \tag{2}$$

$$\theta = \begin{pmatrix} 61407 \\ 3705.9 \\ -15091 \\ 28.659 \\ -375.13 \\ 1212.9 \\ -0.6589 \\ 0.1350 \\ 13.839 \\ -31.972 \end{pmatrix}$$

where γ and n are SNR and the number of users. θ is a coefficient matrix representing θ_0,θ_1,θ_2 or θ_0,θ_1,$\cdots$,θ_9. The polynomial metric Δ_1 and Δ_3 are shown in Figs.1 and 2. In Figs.3 and 4, the polynomial metric Δ_1 of the first order has some differences with Δ_3 of the order in the low SNR, and can not express rapid increase of delay time with SNR.

Differences of metric between the polynomial and delay time metrics are shown in Fig.3 for Δ_1 and Fig.4 for Δ_3, where the difference of 5% or less, that from 5% to 10%,

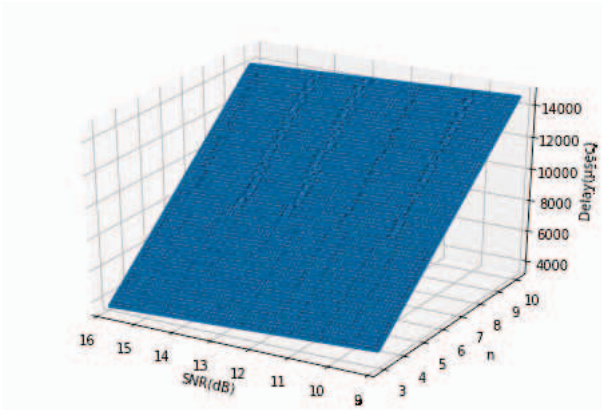


Fig. 1. Polynomial metric Δ_1

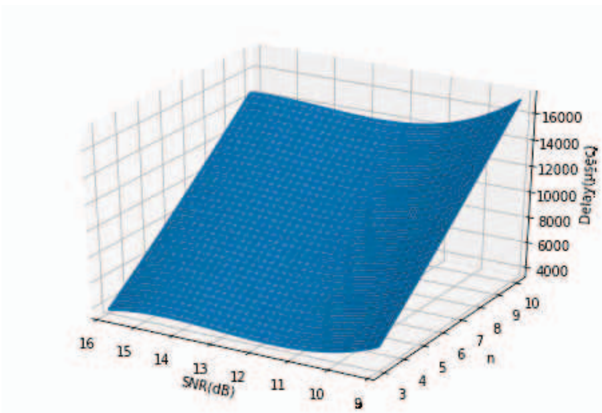


Fig. 2. Polynomial metric Δ_3

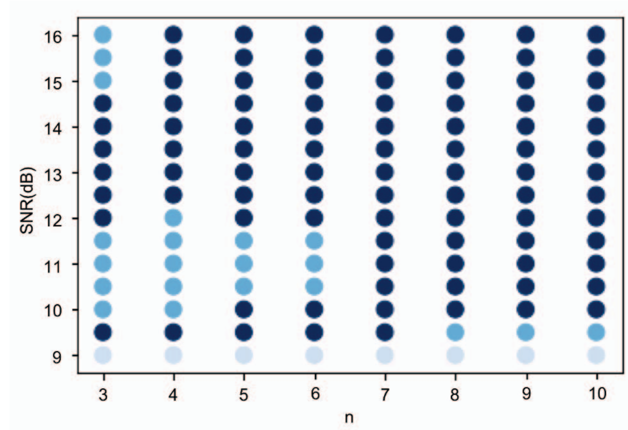


Fig. 3. Difference of metric between Δ_1 and the delay time metric

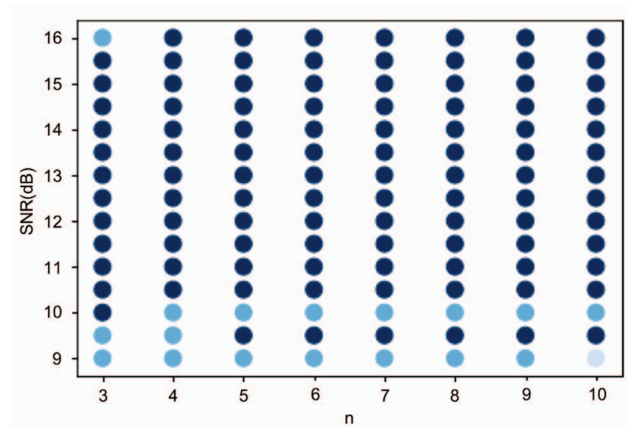


Fig. 4. Difference of metric between Δ_3 and the delay time metric

and that over 10% are denoted by rich blue, light blue, and white circles respectively. From Figs.3 and 4, it is noted that the polynomial metric Δ_1 of the first order has more light blue and white circles in the low SNR and small n (the number of users) regions than Δ_3 of the third order. In most parameters of SNR and n, the polynomial metric Δ_3 of the third order has the small difference with the delay time metric.

III. Conclusion

A channel selection is a significant issue as it improves the communication efficiency. In this paper, a channel selection metric based on SNR and the number of users has been presented for WLAN. It is based on SNR and the number of users. we define a polynomial metric. The channel selection metric may include more parameters other than SNR or the number of users without taking up complex delay analysis. The effective range of metric has been discussed with a difference between the polynomial and delay time metrics.

This work was supported in parts by the Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science, No. 16K06362

References

[1] Giuseppe Bianchi "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on selected areas in communications*, vol. 18, no. 3, pp. 3854–3858, 2004.

[2] Periklis Chatzimisios, and Anthony C. Boucouvalas, and Vasileios Vitsas, "Performance analysis of IEEE 802.11 DCF in presence of transmission errors," *Proc.2004 IEEE International Conference on Communications*, vol. 7, no. 8, pp. 535–547, 2000.

[3] Ken Duffy, David Malone, and Douglas J. Leith "Modeling the 802.11 distributed coordination function in non-saturated conditions," *IEEE communications letters*, vol. 9, no. 8, pp. 715–717, 2005.

[4] David Malone, Ken Duffy Ken, and Doug Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 159–172, 2007.

[5] Shinji Nishijima,Ikuo oka, Shingo Ata, "A Study on Channel Select by Traffic and SNR Estimation," *Technical Report IEICE,*, vol. 116, no. 394, pp. 31–34, 2017.(in Japanese)

Robot Knows Where Human Are Through Sensory Data Fusion

Shih-Huan Tseng
Computer and Communication Engineering
National Kaohsiung University of Science and Technology, Kaohsiung
Kaohsiung,Taiwan
shtseng@nkust.edu.tw

***Abstract*—In a situation where a robot initiates interaction with groups of people, questions such as ”where is the people” should be addressed. This paper proposed a real-time method that enables a robot to find the position and orientation of people. The method is mainly to fuse depth-related data to track the positions of people and compute orientations of those people by using depth-related data. The conducted experimental results demonstrate the properness of the proposed method in knowing where people are.**

***Index Terms*—human-robot interaction, sensory data fusion, human tracking**

I. Introduction

Human-robot interaction (HRI) is an important issue concerning the sociality between robots and humans. A significant aspect of this is the robot’s initiation of interactions upon encountering a group of people. If the robot were to be replaced by a human, the above-mentioned aspect is actually a simple matter for most humans, because they are able to achieve sociality almost unconsciously; however, we know it is definitely not a trivial matter for any robot. To realize this, the robot requires a variety of skills demonstrating three fundamental abilities [1]:

- *human-oriented perception*: skills in detection and tracking of humans as well as recognition of face, speech, and gestures.
- *user modeling*: ability to understand human behavior and to make appropriate decisions [2].
- *sensitivity to the user*: capability of measuring feedback from users and adapting to the responses received.

This paper addresses first issues: 1). human-oriented perception, which refers to the skills of a robot used to detect human’s anatomical features, such as the legs [3], torsos, and faces.

II. Tracking People through Data Fusion

Figure 1 presents the flowchart of this work. Essentially, the proposed method that allows a robot to find the position and orientation of people is through the following three steps: 1) to get the positions of people through the fusion of different depth-related data, 2) tracking positions using Kalman filter, 3) to compute the orientations of people based on human joints.

This research was supported by Ministry of Science and Technology of R.O.C. (MOST grants 108-2218-E-992-310)

A. Data fusion using covariance intersection

we employ multiple sensing devices in our system and adopt a sensor fusion approach referred to as covariance intersection (CI) [4]. This method is able to overcome the disadvantages of using single sensors and thereby enhance the accuracy of detection. We also employ Gaussian distribution for representation of the data attributes of each sensor, wherein the probability distributions are fused into a robot-centered map using the CI algorithm.

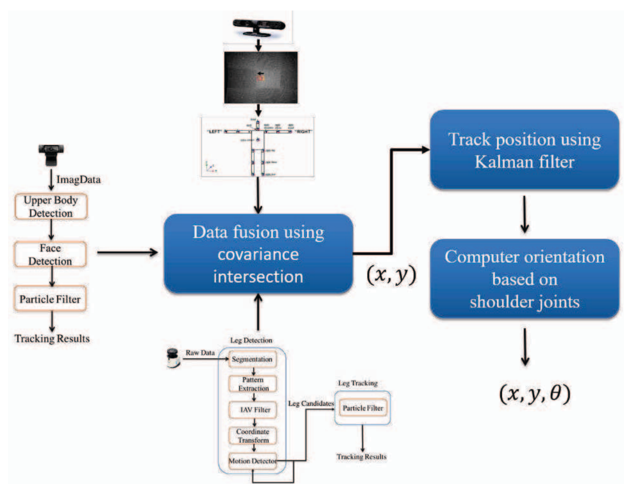


Fig. 1. The three layers of the system architecture

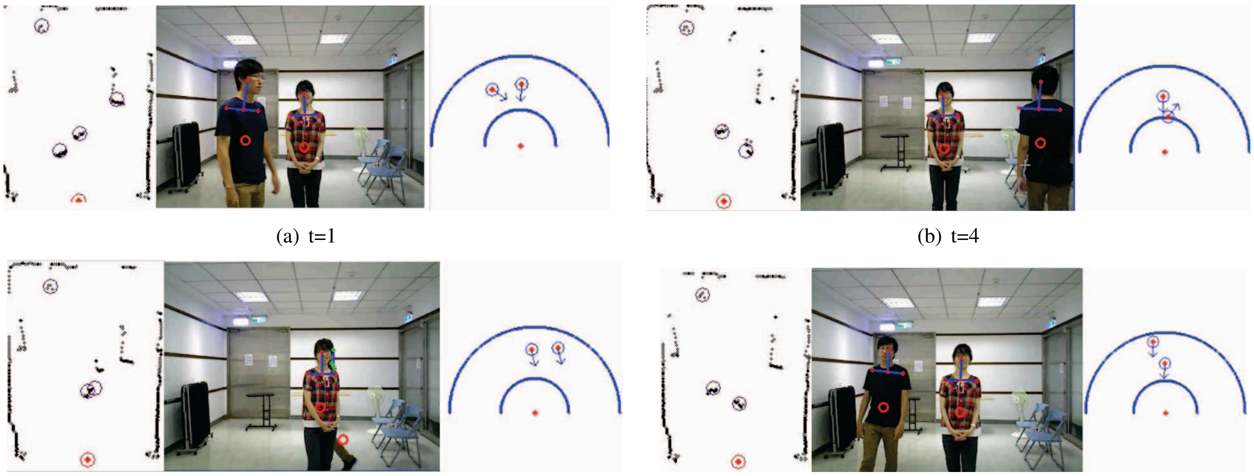
B. Tracking positions using Kalman filter

When the positions of people are located, we apply Kalman filters to track each person. Let the sensor model and prediction model be both assumed as Gaussian ones. Then, the Kalman filter model can be expressed as follows:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \alpha N(\mathbf{H}\mathbf{x_t} : \mathbf{z}_t, \Sigma_z) \int_{\mathbf{x}_{t-1}} N(\mathbf{x_t} : \mathbf{F}\mathbf{x}_{t-1}, \Sigma_x) \mathrm{d}\mathbf{x}_{t-1} \tag{1}$$

where $\mathbf{F}$ and Σ_x are matrices describing the linear transition model and the transition covariance, where $\mathbf{H}$ and Σ_z are the corresponding matrices for the sensor model. Now the mean and covariance of fusion result are used as the parameters in our overall sensor model, namely, $\mathbf{z}_t = \mu_{fusion}$ and $\Sigma_z = \sigma_{fusion}$ respectively.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE



(a) t=1

(b) t=4

(c) t=7

(d) t=10

Fig. 2. Snapshots of tracking with depth-related data fusion: 1) The right hand side snapshots show the estimation results including the positions and heading angles of peoples. 2)There are delays in some figures because the human subject walks with inconsistent speed occasionally. 3) The Kalman filter method deals with the occlusion as shown in Fig. 2(c). 4) The robot can keep track of the person when he comes up as shown in Fig. 2(d) and Fig. **??**

TABLE I
CONFUSION MATRIX AND ACCURACY IN IDENTIFYING THE POSITIONS OF SUBJECTS VIA DATA FUSION (LEGS ARE NOT OCCLUDED)

Fused number / True number	1	2	3	4	5	Accuracy
1	10	0	0	0	0	100%
2	1	9	0	0	0	90%
3	0	2	8	0	0	80%
4	0	1	2	7	0	70%
5	0	0	1	2	7	70%

C. Orientations of People

We define the angle of the human body relative to the robot coordinate system as follows:

$$\theta_{body} = \arctan(\frac{y_{Rshoulder} - y_{Lshoulder}}{x_{Rshoulder} - x_{Lshoulder}}) + \pi/2 \qquad (2)$$

where $[x_{Rshoulder}, y_{Rshoulder}]$ and $[x_{Lshoulder}, y_{Lshoulder},]$ are the 3D coordinates of the right shoulder joint and left shoulder joint of the person observed by the robot.

III. EXPERIMENTAL RESULTS

A. Experiment platform

We adopt the self-made office robot, called ARIO (Agile Robot In Office) as the experiment platform to demonstrate the effectiveness of the method. This robot is equipped with a RGB-D sensor (ASUS Xtion Pro) to detect torsos and shoulders. A high resolution color sensor is to detect faces. We also installed a SICK LMS 100 laser range finder located 40 centimeters above the floor to detect legs.

B. Results

Kalman filter with fusion of different depth-related data allows our robot to track the positions of people successfully. Figure. 2 demonstrates snapshots of one person who moves around other persons by snapshots. The right hand side snapshots in Fig. 2 show the estimation results including the positions and heading angles of peoples. There are delays in some figures because the human subject walks with inconsistent speed occasionally. The covariance intersection method can overcome the false positives of leg detection even under the circumstance where there are many leg-like objects circumstance as shown in left hand side snapshots in Fig. 2. The Kalman filter method deals with the occlusion as shown in Fig. 2(c), where the robot still keeps previous position of the person although it loses the fusion result. When the person comes up as shown in Figs. 2(d), the robot can keep track of the person and get new the positions and heading angles of this person. Table I presents the confusion matrix and accuracy of the data fusion process.

IV. CONCLUSION

In this paper, a real-time method was proposed that enables a robot to find the position and orientation of people. Experiment results have demonstrated the promise of the proposed method in efficiently knowing where people are.

REFERENCES

[1] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, vol. 30, 2006, pp. 33–40.

[2] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," Robotics and Autonomous Systems, vol. 42, no. 3-4, pp. 143–166, 2003.

[3] C.-T. Chou, J.-Y. Li, M.-F. Chang, and L. C. Fu, "Multi-robot cooperation based human tracking system using laser range finder," in Proceedings of IEEE International Conference on Robotics and Automation(ICRA), 2011, pp. 532–537.

[4] S. Julier and J. Uhlmann, Handbook of Multisensor Data Fusion. CRC Press, 2001, ch. General Decentralized Data Fusion with Covariance Intersection (CI).

AIoT Solution Survey and Comparison in Machine Learning on Low-cost Microcontroller

Hoang-The Pham[1], Minh-Anh Nguyen[1], Chi-Chia Sun[12]
[1]Department of Electrical Engineering
[2]Smart Machinery and Intelligent Manufacturing Research Center
National Formosa University, Yunlin, Taiwan

***Abstract*— Neural Networks, especially Convolutional Neural Network [1] are becoming increasingly popular in IoT edge devices today for executing data analytics right at the source without transmitting to Cloud Computing centers. So that it will be reduced latency as well as energy consumption for data communication. In this paper, we will compare CMSIS-NN and uTensor: low energy consumption microcontrollers. Most classification tasks have always-on, and real-time requirements, which limits the total number of operations per neural network inference. So that, with image classification model, microcontrollers will execute lower frame per second than GPU and embedded CPU. CMSIS-NN is a collection of efficient kernels which was developed to maximize the performance and minimize the memory footprint of Neural Network applications. Allow deploy machine learning models on ARM Cortex-M processors for intelligent IoT edge devices.**

I. INTRODUCTION

Connected devices or Internet of Things (IoT) have been rapidly increasing over the past few years and are predicted to reach 1 trillion across various market segments by 2035. As we known, the number of the IoT devices more and more increases, this will place a considerable burden on the network bandwidth, so that latency will be challenging when running the AIoT applications [2] in the future. Dependency on the cloud computing makes it harder to deploy AIoT application in areas with low and unreliable network connectivity. The solution for this problem is edge computing, data will not be transmitted to cloud, they will be processed and executed right at the source after collected by sensors of IoT devices. Thus, this solution will help us to reduce the latency as well as saving energy for data communication.

Deep Neural Network (DNN) [3] based solutions have performed very high accuracies for many complex applications such as computer vision, natural language processing, image classification, optical character recognition, object detection, and speech recognition, etc. Due to complex computation and hardware resource requirements, these executions of Neural Networks must be performed on cloud computing which has high performance server CPUs as well as GPUs. As mentioned above, it will add latency to the AIoT applications. Executing right at the source of data on small microcontrollers can reduce the overall latency and energy consumption of data communication between the IoT devices and the cloud. However, we have to deal with these challenges when deploy Neural Network model on microcontrollers in the edge side.

In this paper, we have surveyed CMSIS-NN [4][5], which is a collection of efficient neural network kernels developed to maximize the performance and minimize the memory footprint of neural networks on ARM Cortex-M7 processor. Furthermore, we will present the workflow to deploy neural networks on Cortex-M7 with and without CMSIS-NN. Caffe [6] is the deep learning framework which is made with expression, speed, and modularity in mind by Berkeley AI Research. We used Caffe for training the image classification models. The models trained will be quantized to reduce the size of the neural network and avoid floating point computations, that are more computationally expensive. Then, we used tools to convert the model weight to C++ code that can be compiled and run on the microcontroller. For the Cortex-M7's hardware, STM32F746ZGT6U [7] has been selected for comparing for performing image classification, especially neural networks model trained on CIFAR-10, MINIST, SVHN datasets.

II. CONTROL METHODS

II.1. Model with CMSIS-NN

CMSIS-NN is a collection of optimized neural network functions for ARM Cortex-M core microcontrollers enabling neural networks and machine learning being pushed into the end node of AIoT applications. It has implemented popular neural network layer types, such as convolution, depth separable, fully connected, pooling and activation (ReLU). Supporting a model trained with a popular framework such as TensorFlow, Caffe. The weights and biases will first be quantized to 8-bit or 16-bit integers then deployed to the microcontroller. The best performance was achieved by leveraging SIMD instructions features of the CPU to improve parallelism available for Cortex-M4 and Cortex-M7 core microcontrollers.

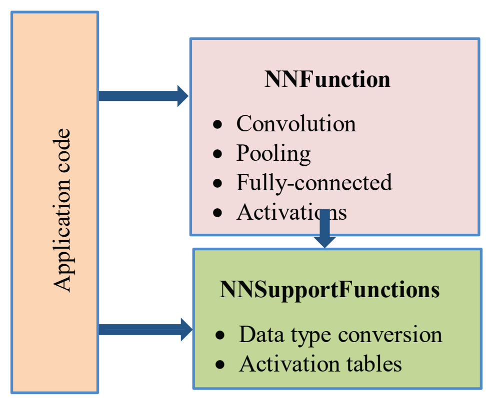


Figure. 1. Block Diagram CMSIS-NN

From pre-trained model CIFAR-10 by Caffe, the model will be translated to C++ file to be able to deploy on ARM. Figure 1 show how to build model with CMSIS-NN and deploy it on MCU is follow:

1) Quantize the model: Once we have the CIFAR-10 trained model, we need to optimize it for microcontroller. We use ARM quantization script to convert the model weights and 14 activations from 32-bit floating point to an 8-bit and fixed-point format. This work will reduce the size of the network and avoid floating point computations.
2) Convert the model: Then we need convert the model into C++ file that we can include it into image recognition application. Use generate script to get the quantization parameters and network graph connectivity and generates the code consisting of NN function calls.
3) Build image recognition application: We need include file to main

file and add classification capabilities to use neural network on ARM. Modify function and call run command to run neural network. We still need image as input of NN so we resize and translate picture to array can work in C++.

4) Deploy on ARM-Cortex M: Use MbedOS [8] compiler to generate binary file and upload to the MCU processor.

II.2. Model with MicroTensor

MicroTensor (uTensor) [9] ARM's early entrant into edge machine learning, takes TensorFlow models and compiles into highly efficient code for edge processing. uTensor converts machine learning models to readable and self-contained C++ source files, to simplify the integration with any embedded project. It is especially designed for low-power, constrained embedded devices, and it has deep roots in TensorFlow and MbedOS. From this merger, we have a great opportunity to bring uTensor's innovations to TensorFlow and ensure it is easy for all developers to use and support a wide range of ARM Cortex-M hardware. All of developers share a common vision to bring machine learning to the edge. We are looking forward to creating a state-of-the-art micro-inference framework together.

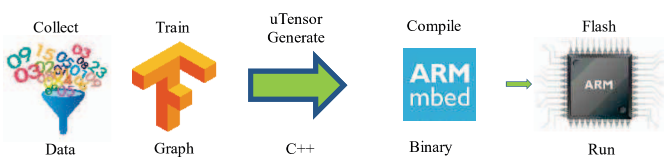


Figure. 2. Workflow of uTensor

This workflow from is Figure 2 is different from deploying neural network model with CMSIS-NN. Instead of training by Caffe, this model will be trained by TensorFlow [10]. The neural network models after trained will be convert to C++ code by tools called uTensor as mentioned above. Here is workflow for deploying Neural Network model CIFAR-10 to ARM Cortex-M7.

1) Training Neural Network model by TensorFlow： From CIFAR-10 dataset, we use TensorFlow and images from CIFAR-10 dataset to train the network. Then convert trained model to model file to use for uTensor-cli.
2) Translate the Neural Network model to C++ code by uTensor-Cli: Use uTensor-cli to identify the output nodes and generate the C++ files from CIFAR-10 model.
3) Compile and flash project to ARM Cortex-M by Mbed OS compiler: We need include file to main file and add classification capabilities to use neural network on ARM. Modify this function and call run command to run neural network. We still need image as input of NN so we resize and translate picture to array can work in C++.
4) Deploy on ARM Cortex-M: Use MbedOS compiler to generate binary file and burn to the board.

III. EXPERIMENTAL RESULTS

Next, we applied these two different ways to deploy neural network models with three datasets. Three common datasets that we used are MNIST (handwritten digits), CIFAR-10, Street View House Number (SVHN). Two models have not yet completed are SVHN without CMSIS-NN and MNIST with CMSIS-NN. The model MNIST with CMSIS-NN was not completed because of it has two Fully Connected layers which generated a lot of weights could not fit the microcontroller's memory. We could only complete measuring the accuracy and the performance of microcontroller when running neural network model.

Due to microcontroller's memory, we could not put in its memory many input images. Accuracy was measured by randomly run 100 samples and calculate the error rates. The performance is measured by calculate how long does it took the model to predict label. Before running the model, we initiated the time object then started counting the timer, let the model run, then stopped the timer and calculate how many seconds to complete the task. We calculated the performance by frames per second (FPS). From table 1, we can see model with uTensor gave us results of performance is better than model with Caffe & CMSIS-NN.

Table 1. Comparison between Caffe & CMSIS-NN and uTensor

	Caffe & CMSIS-NN		uTensor	
Model/ Datasets	Error Rate (%)	Performance	Error Rate (%)	Performance
CIFAR-10	26%	9.09 FPS	15%	2.22 FPS
MNIST	Out of memory	Out of memory	10%	7.69 FPS
SVHN	32%	8.33 FPS	Out of memory	Out of memory

IV. CONCLUSION

In this paper, we were running the neural network with pre-defined input data which is no reality when considering variety choices of sensors, camera, microphone, accelerometer all can be easily integrated with the microcontroller to acquire real-time data form the environment. There are endless possibilities when this neural network framework is leveraged to process those data and extract useful information. The Internet of Things is slowly permeating every aspect of our lives. By using CMSIS-NN, we can easily integrate Machine Learning on ARM then connect to AIoT system to make an intelligent AIoT system. There is an increasing interest in deploying the deep learning algorithms on low-power edge devices such as ARM Cortex-M microcontroller systems. CMSIS-NN can raise speed, and reduce power cost of system so it can help us easily use in many cases.

REFERENCES

[1] Keiron O'Shea1and Ryan Nash, "An Introduction to Convolutional Neural Networks," arXiv:1511.08458, 2015.

[2] Marjan Gusev, Schahram Dustdar, "Going Back to the Roots—The Evolution of Edge Computing, An IoT Perspective," IEEE Internet Computing, Vol. 22, No. 2, pp. 1-5, 2018.

[3] Michael A. Nielsen, "Neural Networks and Deep Learning," Determination Press, pp. 167-176, 2015.

[4] Liangzhen Lai, Naveen Suda, and Vikas Chandra, "CMSIS-NN: Efficient Neural Network Kernels for ARM Cortex-M CPUs," arXiv:1801.06601, 2018.

[5] Liangzhen Lai and Naveen Suda, "Enabling Deep Learning at the IoT Edge. International Conference on Computer-Aided Design," pp. 135:1-135:6, 2018.

[6] Delia Velasco-Montero, Jorge Femández-Bemi, Ricardo Carmona-Gálán, Angel Rodríguez-Vázquez, "On the Correlation of CNN Performance and Hardware Metrics for Visual Inference on a Low-Cost CPU-based Platform," IEEE International Conference on Systems, Signals and Image Processing, pp. 250-252, 2019

[7] "STM32F746ZG datasheet" by STMicroelectronics company https://www.st.com/resource/en/datasheet/stm32f746zg.pdf

[8] "Mbed OS" by Arm Limited available here https://os.mbed.com/

[9] Neil Tan, "uTensor- AI inference library based on Mbed and TensorFlow," https://github.com/uTensor/uTensor

[10] Tom Hope, Yehezkel S. Resheff and Itay Lieder, "Learning TensorFlow," pp. 1-21 O'Reilly Media, 2017.

Searching ROI for Object Detection based on CNN

Chia-Lin Wu
Dept. of Computer Science and Information Engineering, National Central University, Taiwan
106522062@cc.ncu.edu.tw

Chih-Yang Lin
Dept. of Electrical Engineering, Yuan-Ze University, Taiwan
andrewlin@saturn.yzu.edu.tw

Phanuvich Hirunsirisombut
Dept. of Computer Science and Information Engineering, National Central University, Taiwan
phanuvich.h@gmail.com

Hui-Fuang Ng
Dept. of Computer Science, FICT, Universiti Tunku Abdul Rahman, Malaysia
nghf@utar.edu.my

Timothy K. Shih
Dept. of Computer Science and Information Engineering, National Central University, Taiwan
timothykshih@gmail.com

***Abstract*—Several studies have explored the structural design of CNN to improve the network's performance since a well-designed feature extractor can benefit convolution-based tasks. Although CNNs are able to learn important patterns on raw images, images may contain unpredictable noise that can negatively influence the convolutional stage. Feature extraction cannot always accurately capture the desired features based solely on the input image, but including extra information could improve the result. This paper proposes a fusion input design to generate an additional feature that a CNN can use to provide extra ROI information. Whether a model can utilize the additional information is a determining factor that affects the performance improvement. The proposed method is tested on two public datasets with different structural designs. Overall, the results indicate that additional ROI information can deliver benefits to specific tasks.**

Keywords—Convolution Neural Network, Region of Interest, Deep Learning.

I. INTRODUCTION

When a CNN is being applied to a visual recognition task of general objects, some preprocessing steps such as normalization or resizing are usually performed. However, such operations can strip out useful information from the input that could be useful in completing the task. Although we want to eliminate the need for the preprocessing on the input, we also want to generate certain useful information for a CNN model without substantially impacting the main workflow. Moreover, the ability of a pre-trained model which can extract patterns and can be trained on another task with less effort may be compromised if there is a significant modification when applying it to a new task. For instance, if we append a new channel over an input image such as increasing the kernel size, the training result will get worse because the new parameter is not reliable.

Visual attention is a valuable concept that has been used in image and video captioning jobs [1]. A recurrent neural network is usually used as the backbone in the model [2,3,4]. The generation and masking of ROI within the model workflow provide the ability to select the necessary information in each of the training iterations. In this case, the feature extraction step of CNN runs only once, but the feature map remains in use several times without losing relevant features by employing a masking operation. However, for a feedforward network design like the traditional CNN based classifier, an incorrectly executed masking operation can negatively impact the result. Therefore, we propose a fusion input design that utilizes ROI information in the input stage. The ROI information is provided as additional input and fused with the original input image.

II. RELATED WORKS

CNNs are widely used in many visual tasks [4,5,6,7]. Convolution usually takes the role of feature extraction. In [5], a larger number of small filters are used in the convolution layers to capture more complex patterns. The size of the receptive field is directly related to how a CNN describes the input in an effective way since the depth has been found to benefit feature extraction.

The concept in [5,6] is to enlarge the gradient value within a backpropagation process via a proposed fusion operation shortcut. Since the gradient vanishes when the number of layer increases, using summation will produce a higher gradient value. In our work, we follow similar idea of fusing information from different groups. This is a significant and positive change to traditional CNNs. Another branch of thought, as outlined in [9], is to design an extra path that lets the model itself to weight the convolution generated information dynamically. The concept of gating each input channel is also applied in some areas of our work.

III. METHOD

A. The ROI mask

We assume that applying ROI information to a CNN model will improve feature extraction. A ROI can be any form that highlights the desired part of the input like a bounding box mask or a segmentation mask. Additional information is based on the data provided by the dataset. For instance, a bounding box mask has no pixel-level information, whereas a segmentation mask must fit to an object in the image. Therefore, bounding boxes can be more efficient if generated mask is applied.

B. Additional information path

Generally, CNN takes raw images as its input, so it is hard to append more channel-wise data to a pre-trained model and work properly. In details, convolution layers perform calculation using a fixed-size filter kernel, and the depth of a filter depends on the number of input channels. When it comes to a pre-trained model, assigning additional parameters and mask channels will make the early fine-tuning stage unstable, which can worsen the training result.

With regard to the input sample, it is less informative if we treat the binary mask as another input. Moreover, binary mask contains large edges at the ROI boundary that are not directly related to the classification task. Thus, the input mask is applied to the original image to produce a masked image before being entered into the CNN model. Even

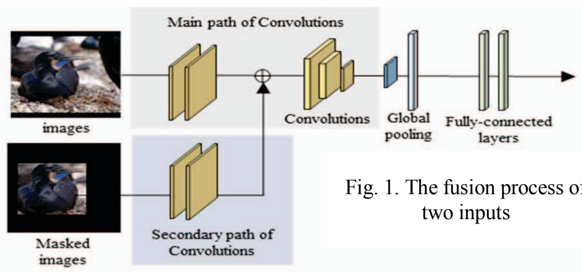


Fig. 1. The fusion process of two inputs

though inputting only a masked image seems fine for classification, the features outside the mask should not be ignored either. Both masked and raw images will be input into the model to allow the model to run different fusion approaches, as shown in Fig. 1.

The additional masked image can be processed by a secondary convolution path that has a smaller number of layers than the main path of the network. The output feature maps will be combined in a certain convolution stage. The secondary path has the same depth and hyper-parameter design as the convolution of the main path before the fusion stage. Applying the same parameter design on the secondary path aims to generate a similar level of features and enhance the fusion operation. The differences between approaches are based on the design choice of the fusion stage in the main path and whether the secondary path shares the same weights with the main convolution path.

Since the output feature maps of the two paths have the same size, the fusion operation can be done via direct summation or weighted summation. For the weighted summation approach, the weights are generated based on both feature maps.

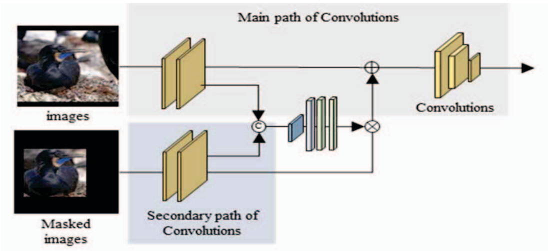


Fig. 1.Illustration of weighted fusion.

The weighted summation module works as shown in Fig. 2. The feature maps of the main path and the secondary path are concatenated and input to a fully connected network that generates the weights for the fusion operation after global pooling. The sub-network is stacked with two layers, following the design as in SENet [8]. The first layer reduces the vector dimension by a fixed factor k=16, followed by a ReLU activate function. The second layer generates weights which have the same channel number as the feature map of the addition path. After channel-wise weighting, the feature maps of the additional path and the main path are summed and then proceed to the next stage.

IV. Experiments

CUB200 [10] and Stanford-Dog [9] are the two datasets used in the experiments. Both datasets contain balanced training samples for each of the classes in the dataset, and provide bounding box data for the mask generation with size 448×488. The base model of the main convolutional path is ResNet50v2 [5,6]. The network in the secondary path has been narrowed down to only one convolution layer that has the same filter number and stride as the first convolution layer of the main path. Parameters of ImageNet [11] were adopted here for initializing the networks and tests were carried out to evaluate the impact of the proposed method on the training results.

Table I shows the performance comparison of various configurations. "*ROI*" denotes whether the additional ROI information is applied. "*Shared-weights*" indicates whether the parameters of the main path are shared with the secondary path. "*Weighted fusion*" indicates whether the proposed fusion operation is used. The results show that even though the improvement is little, additional ROI information can benefit feature extraction and overall classification. We can see that the improvement on the Stanford dog [9] dataset is greater, possibly because the task is more challenging since there are several similar patterns per instance. For the CUB200 [10] dataset, the object categories are less similar, therefore it is reasonable to expect less benefit coming from the ROI information.

TABLE I. The Comparison between different factor

Dataset	**Factors**			**Acc.**
	ROI	***Shared-weights***	***Weighted fusion***	
CUB200 [10]				80.7%
	✓			80.1%
	✓	✓		**81.%**
	✓		✓	80.4%
	✓	✓	✓	80.7%
Stanford dog [9]				78.7%
	✓			79.1%
	✓	✓		79.%
	✓		✓	79.2%
	✓	✓	✓	**79.4%**

V. Conclusion

We proposed a method of fusing additional information to a pre-trained CNN. Although the improvements are not substantial, a positive influence is observed on tasks that can benefit from additional ROI information. The secondary path is thus useful for such cases. In summary, the proposed method could produce greater performance improvement on a task where additional ROI information fits well with its requirement.

References

[1] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron C. Courville, Ruslan Salakhutdinov, Richard S. Zemel, and Yoshua Bengio, "Show, attend and tell: Neural image caption generation with visual attention." In ICML, 2015.

[2] Q. You, H. Jin, Z. Wang, C. Fang, and J. Luo. Image captioning with semantic attention. In CVPR, 2016.

[3] Chen, Xinlei and Zitnick, C Lawrence. Learning a recurrent visual representation for image caption generation. arXiv:1411.5654, 2014.

[4] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Identity mappings in deep residual networks. In ECCV, 2016.

[6] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016.

[7] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He. Aggregated residual transformations for deep neural networks. In CVPR, 2017.

[8] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. arXiv preprint arXiv:1709.01507, 2017.

[9] Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao and Li Fei-Fei. Novel dataset for Fine-Grained Image Categorization. First Workshop on Fine-Grained Visual Categorization (FGVC), In CVPR, 2011.

[10] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. In CVPR, 2009

RSSI Measurement with Channel Model Estimating for IoT Wide Range Localization using LoRa Communication

Yi-Cheng Lin[1], Chi-Chia Sun[12] *Member, IEEE*, and Kuo-Ting Huang[1]
[1]Department of Electrical Engineering
[2]Smart Machinery and Intelligent Manufacturing Research Center
National Formosa University, Yunlin, Taiwan

Abstract-- **Due to the technology of the Internet of Things (IoT) is rising in recent years, there are many popular applications in different fields. The development of network technology is growing for the trend toward wider coverage range and lower power consumption. LoRa system is one of the most popular IoT communication methods in Low Power Wide Area Network (LPWAN) as an auxiliary role of IoT to assist the transmission between devices in various applications. In order to provide more received information for the smart applications of IoT, the positioning technology to obtain the location information of the devices is necessary. However, GPS singles consumes too much energy for IOT devices, they are conflict to each. In this paper, we proposed a channel model estimating method to improve the conventional RSSI based distance measurement way in LoRa system accuracy up to same level of TDOA based system in wide rang localization applications.**

Keywords: IoT, LPWAN, LoRaWan, RSSI, Localization

I. Introduction

In this paper, we used the LoRa system for wide area localization experiments, which are characterized by long-distance communication, low power consumption and low cost. It is suitable for traffic signal control in smart cities, equipment data collection in smart factories, and plants monitoring in smart agriculture and other applications that need low data transfer rates and low power consumption. In these smart applications, it often combines positioning technology to obtain location information of its device for effectively use the received message for more in-depth use.

The GPS [1-2] system has highest positioning accuracy which is also the most popular choice for positioning on the markets. Meanwhile, the expensive hardware cost and the high power consumption are existed. Therefore, GPS are often used for navigation applications. In contrast to the cost, GPS is not suitable for applications that do not require real-time tracking and high accuracy of positioning, such as to record the stay duration and place of the staff, students or visitors to be present at their facilities or campus or farming in agriculture [3-4]. In the past, several Time Difference of Arrival (TDOA) based positioning methods with LoRa system are presented [5-6]. However, their calculation costs are relative high and the error is still more and less around 100 meters. We consider whether there are possibilities to develop positioning functions using RSSI method in LoRa wire range communication technology, it sure will be a more economical way. Next, we choose LoRa system to estimate the end-device location by RSSI value when it transmits data between LoRa gateway without extra equipment. Indeed, due to the low-power feature of LoRa, the RSSI is not sensitively varied to its distance change. Therefore, we create a related curve between RSSI and distance for the environment where exists LoRa system, named channel model estimating.

In experiments, the recognition of the RSSI value and the distance is sufficient, which is enough to be used as the positioning distance estimation in wide area application. Next, a channel model curve is established by the experiments. Then, the triangulation method is used and the self-adjustment error method is used to solve the problem that the actual positioning cannot be converged to a clear coordinate. The experimental results show that the RSSI positioning with estimation method can reach same accuracy level compared to TDOA method for the LoRa system at wide-range positioning application. This paper is organized as follows. Section II will brief introduce the proposed system solution for LoRa localization. The experimental results are shown in Section III while Section IV concludes the paper

II. System Architecture

The LoRa system is composed of end-device, gateway, network server and application. The end-device mainly includes a microcontroller(MCU) and a LoRa module. The LoRa module is initialized by MCU. It includes parameters such as frequency band, spread factor, and bandwidth, which are set by MCU using LoRa Command. The LoRa gateway includes Raspberry Pi embedded system and LoRa modules. The LoRa module is initialized by the embedded system. Then upload the data to the web server through the connection method of MQTT and store it. The user can access the data in the web server on the application side.

Since the established environmental channel model throughout measurements is not ideal as the distance increases, the RSSI value decreases, but the upper and lower fluctuations are formed. When the corresponding distance is erroneously selected, the estimation error of the accuracy of positioning will be influenced. Hence, this paper proposed an environmental channel model curve correction method to

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

compensation the distance estimation error caused by this condition. Following steps are the proposed correction method, the $RSSI_i$ values of N distance points on the established environmental channel model curve are sequentially compared, $1 \leq j \leq N$. The result of comparing the $RSSI_i$ value of the starting distance point i with the $RSSI_{(i+j)}$ value of the next distance point $i + j$ ($1 \leq j < N$, $i + j \leq N$) is called R,

$$R = RSSI_i - RSSI_{(i+1)}. \quad (1)$$

The channel model curve correction method is divided into 6 steps:

Step 1: Make $i = 1, j = 1$.

Step 2: Run Eq. (1). When $R \geq 0$, run Step 3. When $R < 0$, run Step 4.

Step 3: Save the RSSI value($RSSI_{(i+j)}$) of the next distance point($i + j$), then run Step 5.

Step 4: Discard the RSSI value($RSSI_{(i+j)}$) of the next distance point($i + j$), then run Step 6.

Step 5: When $i + j < N$, update the next distance point($i + j$) to the starting distance point i ($i = i + j$). Rerun Step 2, and end until $i + j = N$.

Step 6: When $i + j < N$, make $j = j + 1$, Continue to use the RSSI value($RSSI_{(i+j)}$) of the starting distance point i and the lower down distance point($i + j$) and Rerun Step 2 until the end of $i + j = N$.

Fig. 1 shows the steps in a flow chart manner. By using the environmental channel model curve correction method, the original curve is effectively corrected to a monotonically decreasing curve called the effective environmental channel model curve, as shown in Fig. 2. The positioning resolution has been corrected to more than 300 meters and the effective measurement distance is 1200 meters.

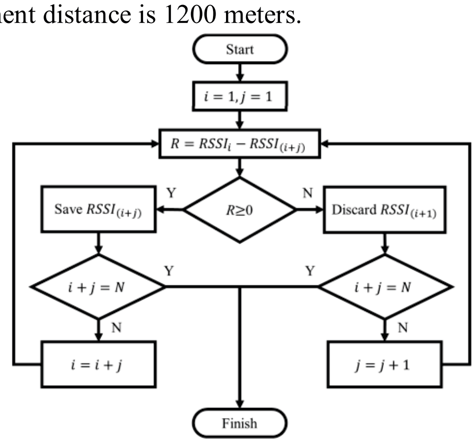


Fig. 1. Channel Model Curve Correction Method Flow Chart

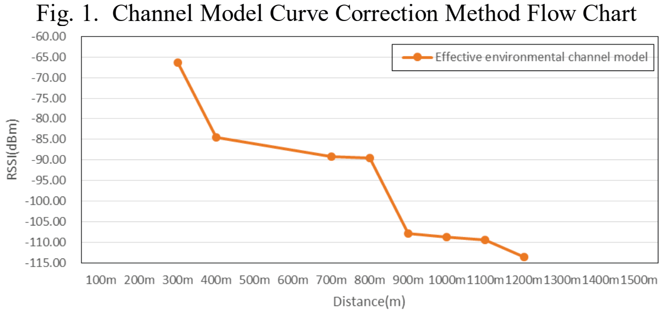


Fig. 2. Effective environmental channel model.

III. Experimental Results

Since the corrected curve has set the effective measurement experimental distance to 1,200 meters, three reference LoRa gateway, G0, G1 and G2 are placed as an enclosed triangle forming a positive triangle positioning area of 623,500m^2 area, which is used as the positioning area of this experiment.

From these nodes to be tested, the error of the positioning estimation result when the node is located in the central location is small. The average error of the eight nodes to be tested is only 173.61 m, and the error standard deviation is 72.19 meters as shown in Table 1. Compared to TDOA method [8-9], the proposed channel correction method can use the LoRa RSSI value to estimate the position localization in the large-scale outdoor area.

Table 1. Experimental Comparison

	Center of Positioning area	The TDOA error for the walking [5]	ALE - No outliers, Sample mean TDOA [6]
Median error	N/A	207 m	N/A
Average error	173.61 m	N/A	127 m
Error standard deviation	72.19 m	175 m	N/A

IV. Conclusion

In this paper, a channel model estimating method is proposed to improve the conventional RSSI based distance measurements at LoRaWan system, especially when using RSSI method in LoRa wireless communication technology, it can not only reduce the computation effort but also reduce hardware costs. The experimental results indicate that average accuracy up can improve to same level of TDOA based system in wide rang localization applications.

References

[1] M. Fang, L. Li, and W. Huang, “Research of Hybrid Positioning Based Vehicle Interactive Navigation System,” Multimedia Information Networking and Security, pp. 974-978, 2010.

[2] G. ANGEL and A. Brindha, “Real-Time Monitoring of GPS-Tracking Multifunctional Vehicle Path Control and Data Acquisition Based on ZigBee Multi-hop Mesh Network,” Recent Advancements in Electrical, Electronics and Control Engineering, pp. 398-400, 2011.

[3] F. Wu, J.-M. Redouté, and M. R. Yuce, “We-safe: A self-powered wearable IoT sensor network for safety applications based on LoRa,” IEEE Access, Vol. 6, pp. 40846-40853, 2018.

[4] N. Hayati, and M. Suryanegara, “The IoT LoRa system design for tracking and monitoring patient with mental disorder,” 2017 IEEE International Conference on Communication, Networks and Satellite, pp. 135-139, 2017.

[5] N. Podevijn, J. Trogh, A. Karaagac, J. Haxhibeqiri, J. Hoebeke, L. Martens, ... & W. Joseph,” TDoA-based outdoor positioning in a public LoRa network,” 12th European Conference on Antennas and Propagation, pp. 1-4, 2018.

[6] B. C. Fargas, & M. N. Petersen, “GPS-free geolocation using LoRa in low-power WANs,” 2017 global internet of things summit, pp. 1-6, 2017.

Blind Equalizer with Noise Reduction Function

Mitsuru Mashimo[1],Minoru Komatsu[2] and Hiroki Matsumoto[3]
[1,2,3] E-mail: [1]m1671042,[2]komatsu,[3]matsumoto@maebashi-it.ac.jp
Maebashi Institute of Technology 460–1 Kamisadori, Maebashi, Gumma, 371–0816 Japan

I. Introduction

Recently, in order to compensate for distortion of the received signals, the blind equalization method has been studied. The blind equalizer is designed by using only received signals. However, equalization precision is lower when noise is included in the received signals [1]. Therefore, in order to solve the problem, we propose a method to equalize accurately even in the noisy environment by using Total Least Squares (TLS). Specifically, the procedures of equalization with noise reduction function are as follows:

1) First, we estimate the channel accurately even in the noisy environment by using TLS. Next, we calculate the noise by using the result. Last, we subtract noise from the received signals.
2) We perform blind equalization on the received signals without noise.

The advantage of proposed method is evaluated by computer simulation.

II. Proposed method

In this section, we propose a method to equalize accurately even in the noisy environment.

A. Proposed model

Fig.1. shows the proposed model.

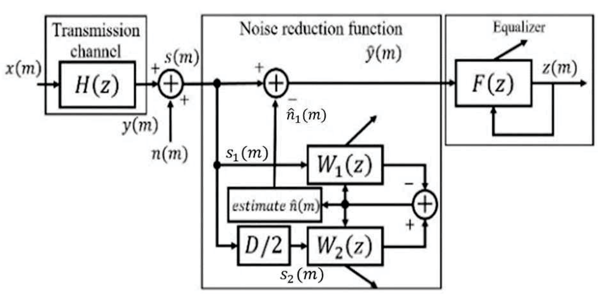


Fig. 1. Proposed model

The proposed model is consisted of three blocks (transmission channel, on-line noise reduction function, on-line blind equalizer). In the next subsection, we describe the noise reduction function.

B. Noise reduction function

In this subsection, we describe noise estimation by using the TLS method and noise reduction function.
First, in order to reduce noise, we describe the estimation of channel. We use TLS for channel estimation because this method can accurately estimate the channel even in noisy environment [2]. In the model, $x(m)$ is transmitted signals. $s(m)$ is received signals including channel output signals $y(m)$ and noise $n(m)$. Letting the estimation system of $H(z)$ be $W_1(z)$, parameter vector $\mathbf{w}_1(m)$ of $W_1(z)$ is given by

$$\mathbf{w}_1(m) = [w_{1(0)}(m), w_{1(1)}(m), \cdots, w_{1(M-1)}(m)]^T, \quad (1)$$

and letting the dummy system of $W_1(z)$ be $W_2(z)$, parameter vector $\mathbf{w}_2(m)$ of $W_2(z)$ is given by

$$\mathbf{w}_2(m) = [w_{2(0)}(m), w_{2(1)}(m), \cdots, w_{2(M-1)}(m)]^T. \quad (2)$$

where,M is the tap length of $\mathbf{w}_1(m)$ and $\mathbf{w}_2(m)$. Input signals vector $\mathbf{s}_1(m)$ of $W_1(z)$, input signals vector $\mathbf{s}_2(m)$ of $W_2(z)$ are given by

$$\mathbf{s}_1(m) = [s_1(m), s_1(m-1), \cdots, s_1(m-(M-1))]^T \quad (3)$$

$$\mathbf{s}_2(m) = [s_2(m), s_2(m-1), \cdots, s_2(m-(M-1))]^T, \quad (4)$$

where,$s_1(m)$, $s_2(m)$ are twice oversampling signals of $s(m)$. D is the delay of one sample and $\hat{n}(m)$ is the estimated noise. Specifically, the update equation of $\mathbf{w}(m)$ based on the TLS is given by

$$\mathbf{w}(m+1) = \mathbf{w}(m) - \mu((\mathbf{s}(m)\mathbf{s}^T(m)) + e^2(m)I)\mathbf{w}(m), \quad (5)$$

where,

$$\mathbf{w}(m) = [\mathbf{w}_2^T(m), \mathbf{w}_1^T(m)]^T, \quad (6)$$

$$\mathbf{s}(m) = [\mathbf{s}_2^T(m), -\mathbf{s}_1^T(m)]^T, \quad (7)$$

$$e(m) = \mathbf{s}^T(m)\mathbf{w}(m), \quad (8)$$

μ is step gain.
Next, based on the TLS [3], the noise $\hat{n}_1(m)$ is given by

$$\hat{n}_1(m) = ||\psi(m)|| sin\theta_3, \quad (9)$$

$$\theta_3 = \pi - (\theta_1 + \theta_2), \quad (10)$$

where,$\psi(m)$ is given by $\frac{\mathbf{s}^T(m)\mathbf{w}(m)}{||\mathbf{w}(m)||}$, θ_1 and θ_2 are given by

$$\theta_1 = cos^{-1}\frac{\mathbf{s}^T(m)\mathbf{w}(m)}{||\mathbf{s}(m)|| \cdot ||\mathbf{w}(m)||} \quad (11)$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

$$\theta_2 = tan^{-1} \frac{||\mathbf{s}_1(m)||}{||\mathbf{s}_2(m)||}. \tag{12}$$

Finally, equalizer input signals $\hat{y}(m)$ without noise is given by

$$\hat{y}(m) = s_1(m) - \hat{n}_1(m). \tag{13}$$

Letting equalizer input signals vector without noise be $\hat{\mathbf{y}}(m)$, $\hat{\mathbf{y}}(m)$ is given by

$$\hat{\mathbf{y}}(m) = [\hat{y}(m), \hat{y}(m-1), \cdots, \hat{y}(m-(L-1))]^T. \tag{14}$$

where, L is the tap length of the equalizer.

C. Blind deconvolution

In this subsection, we describe the adaptive blind equalization. We use the minimum entropy method for equalization because equalization precisition with this method is relatively high [4]. Letting parameter vector of equalizer $F(z)$ be $\mathbf{f}(m)$, $\mathbf{f}(m)$ is given by

$$\mathbf{f}(m) = [f_{(0)}(m), f_{(1)}(m), \cdots, f_{(L-1)}(m)]^T. \tag{15}$$

Letting regenerated signals be $z(m)$, $z(m)$ is given by

$$z(m) = \hat{\mathbf{y}}(m)\mathbf{f}^T(m). \tag{16}$$

The update equation of $\mathbf{f}(m)$ based on the minimum entropy method is given by

$$\mathbf{f}(m+1) = \mathbf{f}(m) - \alpha\hat{\mathbf{y}}(m)z(m)(|z(m)|^{q-2} - r_q), \tag{17}$$

where, α is step gain, $q = 4$ and $r_q = \frac{3}{(q+1)}$ [4].

D. Proposed algorithm

In this subsection, Table 1 shows the proposed algorithm of on-line equalization with noise reduction function.

TABLE I
PROPOSED ALGORITHM

Step0 Initialization
$m = 0$
$\mathbf{w}_1(m) = [0, \ldots, 0, 1, 0, \ldots, 0]^T$
$\mathbf{w}_2(m) = [0, \ldots, 0, 1, 0, \ldots, 0]^T$
$\mathbf{w}(m) = [\mathbf{w}_1{}^T(m), \mathbf{w}_2{}^T(m)]^T$
$\mathbf{f}(m) = [0, \ldots, 0, 1, 0, \ldots, 0]^T$
Step1 Noise reduction
$\mathrm{e(m)} = \mathbf{s}^T(m)\mathbf{w}(m)$
$\mathbf{w}(\mathrm{m}+1) = \mathbf{w}(\mathrm{m}) - \mu\big((\mathbf{s}(\mathrm{m})\mathbf{s}^T(m)) +$
$\mathrm{e}^2(m)\boldsymbol{I}\big)\mathbf{w}(m)$
estimating $\hat{n}_1(m)$ using $\mathbf{s}(m), \mathbf{w}(m)$
$\hat{y}(m) = s_1(m) - \hat{n}_1(m)$
Step2 Blind equalization
$\mathrm{z(m)} = \hat{\mathbf{y}}^T(m)\mathbf{f}(m)$
$\mathbf{f}(m+1) = \mathbf{f}(m) -$
$\alpha\hat{\mathbf{y}}(m)z(m)\{\lvert z(m)\rvert^{q-2} - r_q\}$
$m = m + 1$
Go to Step1

III. Computer simulation

A. Simulation conditions

In this simulation, we compare the equalization precision of the regenerated signals with the conventional method and the proposed method. The conditions are as follows:
The parameter vector $\mathbf{h}$ of $H(z)$ is $\mathbf{h} = [0.1, -0.3, 1.0, 0.4, -0.1]^T$. The distribution of the transmitted signals is uniform distribution and the distribution of the noise is Gaussian distribution. The SNR is 10dB. The tap length of estimation system for the channel, the dummy system and the equalizer are 5th-order, and the step gain α is 0.01. The algorithm is evaluated by using Mean Square Error (MSE). MSE is given by

$$MSE = 10log_{10}\frac{E[(x(m) - z(m))^2]}{E[(x(0) - z(0))^2]}. \tag{18}$$

B. Results and Discussion

Fig.2. shows the result.

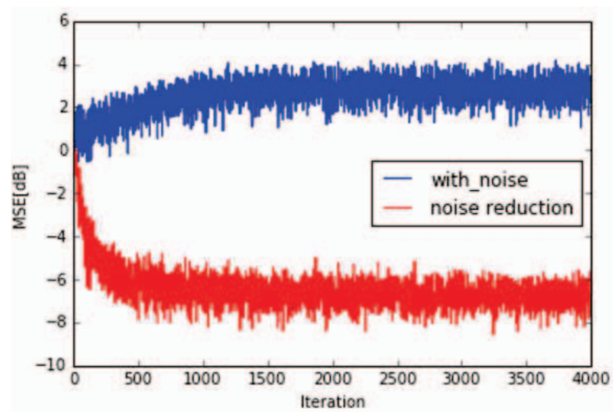


Fig. 2. Transmitted signal:uniform distribution,noise:gaussian distribution

From the Fig.2, the proposed method has high accuracy. In the proposed method, since the noise reduction function is effectively implemented, it is thought that equalization can be performed even in noisy environment.

IV. Conclusion

We proposed the adaptive blind equalizer with noise reduction function. In the proposed method, we confirmed that equalization precision is higher. The future works is to analyze equalization accuracy in noisy conditions.

References

[1] Y. Sato, ”Linear equalization theory, In Japanese” May 1990
[2] Carlos E. Davila, Member, IEEE ”An Efficient Recursive Total Least Squares Algorithm for FIR Adaptive Filtering” IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 42, NO. 2, FEBRUARY 1994
[3] T. Ogawa, H. Matsumoto, ”Blind Deconvolution Algorithm Using Observation Signal Including Noise, In Japanese” Journal of Signal Processing, Vol. 20, No. 3, pp. 105-112, 2016
[4] H. Matsumoto, T. Furukawa, ”A realization on faster convergence and higher reliability of the new blind deconvolution algorithm using the minimum entropy method, In Japanese” Trans. IEE of Japan, Vol. 122-C, No. 3, pp. 448-456, 2002

Verification of Regression Analysis of Muscle Fatigue Using Wrist EMG

Misato Matsushita
Graduate School of Advanced Technology and Science
Tokushima Univercity
Tokushima, Japan
c501937017@tokushima-u.ac.jp

Momoyo Ito
Graduate School of Advanced Technology and Science
Tokushima Univercity
Tokushima, Japan
momoito@is.tokushima-u.ac.jp

Shin-ichi Ito
Graduate School of Advanced Technology and Science
Tokushima Univercity
Tokushima, Japan
s.ito@tokushima-u.ac.jp

Minoru Fukumi
Graduate School of Advanced Technology and Science
Tokushima Univercity
Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

***Abstract*—Muscles can cause injury by training to improve physical performance. However, there are few ways to assess muscle fatigue currently. Therefore, in this paper, muscle fatigue is evaluated using surface EMG(ElectroMyoGram). For discrimination, we used linear regression analysis and support vector regression, and performed comparative verification.**

***Index Terms*—Biometrics, Wrist EMG, Regression analysis, SVR**

I. Introduction

Athlete and student club activities train to improve physical performance almost every day. Meanwhile, excessive strain can cause injury and as a result poor performance.

Presently, evaluation of muscle fatigue is made by subjective evaluation using visual analog scale (VAS) and questionnaires [1]. Several methods were presented based on biochemical evaluations, physiological evaluations, etc. to objectively evaluate muscle fatigue [2]. However, it is considered that those are not easy to use because they require medical specialized knowledge or need to visit a specialized institution.

Therefore, in this paper, a method is proposed to evaluate muscle fatigue easily in any environment with personal computers. This paper focuses on muscle action potentials that are easy to handle as a method of evaluation and could be handled anywhere. Most of the previous researches on muscle fatigue using electromyography were performed by attaching a sensor to a part with high muscle mass. However, measurement in a location with a large amount of muscle, burdens subjects and is contrary to the purpose of facilitating measurement. Therefore, we use wrist EMG(ElectroMyoGram) in this paper

II. Proposed method

A. Measurement

The measurement part measures the EMG(ElectroMyoGram) data used in the learning identification part. We use P-EMG plus and dry type sensors for measurement of EMG [3]. Measurement method uses surface electromyography examination. It can be easily measured without using expert knowledge.

Fig. 1. P-EMG plus

Fig. 2. Dry type sensor

Fig. 3. Measurement

As for measurement, sensors at the time of measurement are prevented from being dislocated by attaching around the wrist like a wristwatch.

B. Pre-processing

In this paper dry type sensors are used, but their characteristics are low in stability and preprocessing is therefore required.

1) Data extraction: About 3 seconds data are manually cut out of the part where EMG signals are stable state. The stable state indicates that the hand is bent and the movement is stable state. Thereafter, from the center point of the cut out data, 1024 points of data back and forth respectively are cut out for 2 seconds (2048 points). Then, overlap is made with a fixed width of 512 points and data is cut out to make it raw data.

2) Noise removal: Hum noise and drift noise were observed from EMG signals. Hum noise is mainly caused by electromagnetic induction, electrostatic induction, and leakage current. As a method of removing the hum noise, the power supply frequency band affected by the hum noise is the removal target. The value to be substituted for the value to be removed by the following formula is calculated. Let x(i) be the value to be removed.

$$x(i) = \{x(i-5)+x(i-4)+x(i-3) \\ +x(i+3)+x(i+4)+x(i+5)\}/6 \quad (1)$$

Therefore, we can remove clearly large values, which are a commercial frequency and its multiples.

Drift noise is caused by the fact that power lead wires shake with operation and electrodes float from the skin. Since this noise is characterized by affecting low frequencies, it is removed using a high pass filter. The high pass filter is a filter that has the function of passing frequencies higher than the cutoff frequency as it is and reducing low frequencies. Removal of frequencies below 20 Hz was carried out to remove drift noise while leaving muscle fatigue features.

C. Feature extraction

The pre-processed EMG signal is converted by Fast Fourier transform (FFT) to calculate power spectral values. FFT is a generic term for operations that perform discrete Fourier transform(DFT) processing at high speed. When the values of the data to be used differ significantly, distortion of the high frequency component occurs owing to the rapid change. In order to avoid this phenomenon, after using a Hamming window, frequency analysis by FFT is carried out to calculate power spectra. Characteristics of muscle fatigue usually appear in low frequency components [4]. The more you get tired, the lower the frequency band is. By using power spectral values, we can confirm features of muscle fatigue that cannot be confirmed in raw EMG data.

D. Learning identification

Regression analysis, which is a type of analysis in statistics, has the advantage of being able to model curve relationships. It is classified into linear regression and nonlinear regression. Linear regression is one in which the relationship between the objective variable and the explanatory variable is linear. It determines a regression equation using the least squares method, which is a method of finding the regression equation so that the sum of squares of the estimated error values is minimized. On the other hand, in nonlinear regression, the relationship between the objective variable and the explanatory variable is nonlinear. we used support vector regression in non-linear analysis. In support vector regression, linear analysis is performed after mapping to a multidimensional space using a non-linear function called a kernel. In this paper, we used RBF kernel(Radial basis function kernel). In addition, We performed multiple regression analysis to use many features for identification.

III. Experiment and Result

A. Experiment

Subjects of experiment were three men and a woman in their twenties. We left for more than 2 days so that there was no influence between experiments. Experiment alternated between measurement and muscle training. Measurement repeated identification operation and weakness operation for 5 seconds for 10 times. The discrimination operation was measured with the finger bent. Strength training was carried out using a hand grip. The subjects held the handgrip to the limit and recorded the number of times. The discrimination operation was carried out 4 times a day for 3 days. Below is data used for verification.

TABLE I
Experimental data

days	3 days
stages	4 stages
Number of channels	8 ch
Number of data	360

B. Result

There is a problem with versatility on individual difference in muscle fatigue. Therefore, we made learning identification by each personal model. Verification was performed from two results of regression analysis. The result of discrimination using linear regression analysis is indicated by ”Dataset 1”, and the result of discrimination using support vector regression is ”Dataset 2”. Below is the result.

TABLE II
Result of regression analysis

	A	B	C	D	Avg.
Dataset 1	43.4	20.66	39.83	49.99	38.47
Dataset 2	67.69	70.17	77.26	71.74	71.74

IV. Consideration

In this paper, we were able to confirm the usefulness of support vector regression. As a factor that the non-linear discrimination rate was better than the linear one, it is mentioned that the non-linear error was acceptable. In linear regression, the regression equation was determined using the least squares method, but since support vector regression uses the kernel function to determine the regression equation, it is considered that it could be fitted to a state of fatigue by linear regression. In the future, we would like to increase the number of data by increasing the load to increase identification accuracy. Furthermore, in this paper, the values of all measured channels were used as they were for identification, but by adding weight to the channel measuring data of flexor muscles loaded with strength training, we think that identification accuracy can be improved. In addition, the number of data can be increased by providing the overlap of data for improvement of identification accuracy.

References

[1] Kim, Seong-Yeol and Koo, Sung-Ja, (2016) “Effect of duration of smartphone use on muscle fatigue and pain caused by forward head posture in adults”, Journal of physical therapy science, Vol.28, No.6, pp.6

[2] Vassão, PG and Toma, RL and Antunes, HKM and Tucci, HT and Renno, ACM, (2016) “Effects of photobiomodulation on the fatigue level in elderly women: an isokinetic dynamometry evaluation”, Lasers in medical science, Vol.31, No.2, pp.275-282

[3] “P-EMG plus”(October 31, 2018), Oisaka Electronic Equipment Ltd., February 4, 2019, “http://www.oisaka.co.jp/p-emgplus.html”(in Japanese)

[4] Tomohiro Kizuka, Tadashi Masuda, Tohru Kiryu, and Tsugutake Sadoyama, (2006) “Biomechanism LIbrary Practical Usage of Surface Electromyogram,” Tokyo: Tokyo Denki University Press(in Japanese)

Distant Sound Source Suppression Based on Multichannel Nonnegative Matrix Factorization with Bases Distance Maximization Penalty

Kazuma Takiguchi
Grad. Sch. of Eng. Sci.
Osaka University
Osaka, Japan
takiguchi@sip.sys.es.osaka-u.ac.jp

Arata Kawamura
Info. Sci. and Eng.
Kyoto Sangyo University
Kyoto, Japan
kawamura@cc.kyoto-su.ac.jp

Youji Iiguni
Grad. Sch. of Eng. Sci.
Osaka University
Osaka, Japan
iiguni@sys.es.osaka-u.ac.jp

***Abstract*—In this research, we address distant sound source suppression based on Multichannel Nonnegative Matrix Factorization (MNMF), and propose a new penalized method. A conventional method based on MNMF separates an observed signal into a target signal and other distant sound sources. Unfortunately, MNMF often degrades the separation performance owing to the basis-sharing problem. Our penalized method forces the target basis to become different from the nontarget one. Experimental results show that the proposed method can improve the separation capability more than the conventional one.**

***Index Terms*—nonnegative matrix factorization, distant sound source suppression, semi-supervised method**

I. Introduction

In order to prevent the sound quality deterioration of conversational speech, it is required to suppress the noises mixed with the target signal. In this research, we separate signals observed by the microphone array into a target signal and distant noises, and suppress only the distant noises. Conventionally, Multichannel Nonnegative Matrix Factorization (MNMF) [1] is used as the distant sound source suppression method. MNMF decomposes an observed signal into a basis matrix and an activation matrix. The basis matrix involves gain patterns for each microphones in the observed signal by frequency, and the activation matrix involves the time-varying gain corresponding to each basis.

In addition, Semi-Supervised MNMF (SSM) [1] has been proposed, which utilizes a distant sound source as supervision for a priori training. In SSM, an observed signal is separated by using the supervised distant basis and the target basis. SSM incurs a risk of degrading the separation performance owing to the sharing of bases in the supervised distant basis and the target basis [2]. To solve this problem, we propose a new penalized SSM (PSSM), which employs a penalty term based on bases distance maximization in the cost function to force the target basis to become different from the supervised distant basis. In the following, we assume that the observed signals consist of a single target sound source and multiple distant ones. In this case, SSM can have only two bases where one is the target basis and the other is the nontarget basis for distant sound sources.

II. Proposed Method

A. Decomposition Model of SSM

The following equation represents the decomposition model of SSM:

$$\boldsymbol{Y}(\omega) \approx \boldsymbol{B}(\omega)\boldsymbol{A}(\omega), \tag{1}$$

where $\boldsymbol{Y}(\omega)(\in \mathbb{R}_{\geq 0}^{M\times T})$ is an observed spectrogram, $\boldsymbol{B}(\omega)(\in \mathbb{R}_{\geq 0}^{M\times K})$ is the bases matrix that includes the target basis $\boldsymbol{B}^{\mathbf{S}}(\omega)(\in \mathbb{R}_{\geq 0}^{M\times 1})$ and the supervised distant basis $\boldsymbol{B}^{\mathbf{D}}(\omega)(\in \mathbb{R}_{\geq 0}^{M\times 1})$, and $\boldsymbol{A}(\omega)(\in \mathbb{R}_{\geq 0}^{K\times T})$ is the activation matrix that corresponds to $\boldsymbol{B}(\omega)$. Also, ω is frequency bin, M is the number of microphones, T is the number of frames of the observed signal, and K is the number of sources ($= 2$). In the following, the symbol ω representing frequency is omitted because we always discuss in frequency bins. In SSM, the supervised distant basis $\boldsymbol{B}^{\mathbf{D}}$ is trained in advance via a distant sound source. After fixing $\boldsymbol{B}^{\mathbf{D}}$, the matrices $\boldsymbol{B}^{\mathbf{S}}$ and $\boldsymbol{A}$ are optimized.

B. Cost Functions

In this section, we propose PSSM algorithm. Hereafter, we denote the entries of the nonnegative matrices $\boldsymbol{Y}$, $\boldsymbol{B}$, $\boldsymbol{B}^{\mathbf{S}}$, $\boldsymbol{B}^{\mathbf{D}}$, and $\boldsymbol{A}$ as $y_{m,t}$, $b_{m,k}$, $b^{\mathbf{S}}_m$, $b^{\mathbf{D}}_m$, and $a_{k,t}$, repectively. The cost function of SSM is determined by generalized Kullback-Leibler divergence ($KL\text{-}divergence$) [3] between $\boldsymbol{Y}$ and $\boldsymbol{BA}$. In this study, we propose the use of the following generalized cost function:

$$\mathcal{J} = \mathcal{D}_{KL}(\boldsymbol{Y} \parallel \boldsymbol{BA}) \tag{2}$$

In PSSM, to avoid the sharing of bases, we make $\boldsymbol{B}^{\mathbf{S}}$ as different as possible from $\boldsymbol{B}^{\mathbf{D}}$. We consider the maximization of Euclidean distance between these bases , which is given by

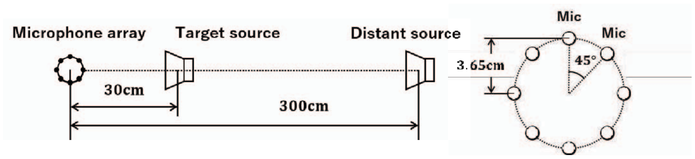
$$\arg\max_{b^{\mathbf{S}}_m} \sum_m (b^{\mathbf{S}}_m - b^{\mathbf{D}}_m)^2. \tag{3}$$

This maximization corresponds to the minimization of this reciprocal. Hence, the cost function with the bases distance maximization penalty is given by

$$\mathcal{L} = \mathcal{J} + \frac{\mu}{\sum_m (b^{\mathbf{S}}_m - b^{\mathbf{D}}_m)^2}, \tag{4}$$

TABLE I
EXPERIMENTAL CONDITIONS

Sampling frequency	16,000 Hz
Frame length	4,096 samples
Frame shift	2,048 samples
Reverberation time	0.33 s
Number of MNMF iterations	300
Channels	8 ch
Signal to distant noise ratio	−5, 0 dB



(a) (b)

Fig. 1. (a) Recording condition and (b) the microphone array used in (a).

where μ is the nonnegative weighting parameter.

C. Auxiliary Functions and Update Rules

In this section, we derive the update rules based on the cost function (4), similarly to [2], [3]. Since it is difficult to analytically derive the optimal $\boldsymbol{B}^{\mathbf{S}}$ and $\boldsymbol{A}$, we define auxiliary function $\mathcal{L}^+$, $\mathcal{J}^+$, and $\mathcal{P}^+$ that represent the upper bound of $\mathcal{L}$, $\mathcal{J}$, and the penalty term in (4), respectively. Applying Jensen's inequality to the penalty term, we obtain

$$\frac{\mu}{\sum_m (b_m^{\mathbf{S}} - b_m^{\mathbf{D}})^2} \leq \mu \sum_m \frac{\phi_m^2}{(b_m^{\mathbf{S}} - b_m^{\mathbf{D}})^2} \equiv \mathcal{P}^+, \tag{5}$$

where ϕ_m (≥ 0) is an auxiliary variable that satisfies $\sum_m \phi_m = 1$ and the equality in (5) holds if and only if the auxiliary variable is set to

$$\phi_m = \frac{(b_m^{\mathbf{S}} - b_m^{\mathbf{D}})^2}{\sum_{m'} (b_{m'}^{\mathbf{S}} - b_{m'}^{\mathbf{D}})^2}. \tag{6}$$

Using (4) and (5), we can define the upper bound function $\mathcal{L}^+$ as

$$\mathcal{L}^+ = \mathcal{J}^+ + \mathcal{P}^+, \tag{7}$$

The update rules with respect to each variable are determined by setting the gradient of the cost function (7) to zero. By solving $\partial \mathcal{L}^+ / \partial b_m^{\mathbf{S}} = 0$, we can obtain the update rules of $b_m^{\mathbf{S}}$ and $a_{k,t}$ with the bases distance maximization penalty as

$$b_m^{\mathbf{S}} \leftarrow b_m^{\mathbf{S}} \left(\frac{\sum_t y_{m,t} a_t^{\mathbf{S}} (\sum_k b_{m,k} a_{k,t})^{-1}}{\sum_t a_t^{\mathbf{S}} - 2\mu (b_m^{\mathbf{S}} - b_m^{\mathbf{D}}) \{\sum_{m'} (b_{m'}^{\mathbf{S}} - b_{m'}^{\mathbf{D}})^2\}^2} \right), \tag{8}$$

$$a_{k,t} \leftarrow a_{k,t} \left(\frac{\sum_m y_{m,t} b_{m,k} (\sum_k b_{m,k} a_{k,t})^{-1}}{\sum_m b_{m,k}} \right), \tag{9}$$

where $a_t^{\mathbf{S}}$ is a entry of matrix $\boldsymbol{A}^{\mathbf{S}} (\in \mathbb{R}_{\geq 0}^{1 \times T})$ which is the target activation included in $\mathbf{A}$. Since $b_m^{\mathbf{S}}$ is nonnegative, the denominator on the right side of (8) has to be nonnegative. Before each update, nonnegative variable μ is determined by

TABLE II
AVERAGE SCORES OF SDR OF 9 OBSERVED SIGNALS IN EACH SNR CASE

Method	SNR = −5 [dB]	SNR = 0 [dB]
SSM	−3.5	2.5
PSSM	−1.2	3.1

$$\mu \leftarrow \min_m \frac{\sum_t a_t^{\mathbf{S}}}{2(b_m^{\mathbf{S}} - b_m^{\mathbf{D}}) \{\sum_{m'} (b_{m'}^{\mathbf{S}} - b_{m'}^{\mathbf{D}})^2\}^2}. \tag{10}$$

III. EXPERIMENT

To confirm the efficacy of PSSM, we compared the applicability of PSSM and SSM ($\mu = 0$ in (8)) for separating mixture signals into the target and the distant sound source. The experimental conditions used are shown in Table I. We prepared 3 kinds of target speech signals with female and male speech, and 3 kinds of environmental sound sources with car, machine and babble noises. The observed signals were obtained by recording two sources selected from the target and environmental signals. Therefore we prepared 9 combinations of observed signals. When recording the observed signals, the target signal is placed in nearby position from a microphone array (30cm), and the environment one is in distant position (300cm), as shown in Fig. 1(a). Fig. 1(b) shows a circular microphone array we use, which has 8ch microphones. Moreover, the target signal to distant noise ratio (SNR) of the observed signal is −5[dB] and 0[dB].

As the evaluation score, we used the signal to distortion ratio (SDR) [4] that indicates the quality of the separated target sound. Table II shows the average scores of SDR of 9 observed signals in two conditions that SNR is −5[dB] and 0[dB]. From this result, we can confirm that our proposed method can achieve higher separation accuracy compared with the conventional method.

IV. CONCLUSION

In this study, we propose a new penalized SSM that forces the target basis to become different from the distant basis trained in advance. From the experimental result, it can be confirmed that the proposed method increases the separation performance compared with the conventional method.

REFERENCES

[1] Y. Murase, N. Ono, S. Miyabe, T. Yamada and S. Makino, "Far-noise suppression by transfer-function-gain non-negative matrix factorization in ad hoc microphonearray," Acoustical Society of Japan, volume 73, issue 9, pp. 563–570, 2017.

[2] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi and K. Kondo, "Music signal separation based on supervised nonnegative matrix factorization with orthogonality and maximum-divergence penalties," IEICE Trans. Fundamentals, vol. E97-A, no. 5, pp. 1113-1118, 2014.

[3] H. Kameoka, "Non-negative matrix factorization and its variants for audio signal processing," in Applied Matrix and Tensor Variate Data Analysis, T. Sakata (Ed.), Springer Japan, 2016.

[4] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," IEEE Trans. Audio, Speech, Lang. Process., vol. 14, no. 4, pp. 14621469, Jul. 2006.

Alveolar fricative consonants detection with easily interpretable feature for speech training

1st Kouki Furuya
Grad. Sch. Eng. Sci.,
Osaka University
Osaka, Japan
furuya@sip.sys.es.osaka-u.ac.jp

2nd Arata Kawamura
Info. Sci. and Eng.,
Kyoto Sangyo University
Kyoto, Japan
kawamura@cc.kyoto-su.ac.jp

3rd Youji Iiguni
Grad. Sch. Eng. Sci.,
Osaka University
Osaka, Japan
iiguni@sys.es.osaka-u.ac.jp

***Abstract*—In this study, we conducted feature extraction for speech training by hearing-loss people alone. We targeted alveolar frictional consonant (/s, z/), which is particularly error prone, and divided its features into three parts: fricativeness, articulation point, and voicing. The indicators created by these features are easily interpretable and help correct incorrect pronunciation.**

***Index Terms*—Speech Training, Phoneme Detection, Fricativeness, Articulation Point, Voicing**

I. Introduction

The incidence of hearing loss in newborns is 0.133% [1]. Congenital hearing loss can cause speech problems. Speech disorders are improved with the help of a speech therapist, but support systems for speech therapy that do not require specialists are also being studied [2], [3]. However, some computer-based systems require people to assist in training [4]. There is also a system for practicing alone, but there are many that only show how speech is actually perceived and do not teach what mistakes are being made [5]. Therefore we developed a speech training system that can be trained by itself and presents improvement methods. In this method, alveolar fricative consonant (/s/ and /z/) is detected using three easily interpretable feature quantities. In this case, only the /s/ and /z/ utterances with many mistakes were targeted, but we plan to support other consonants in the future.

II. Method

A. Characteristics of alveolar fricative consonants

To prepare interpretable features, consider Phonetic features of alveolar fricative consonants. They have the following Phonetic features.

- They are fricative consonant. (Fricativeness)
- They are alveolar consonant. (Articulation point)
- /s/ is unvoiced, and /z/ is voiced consonant. (Voicing)

Create an index that represents each feature.

B. Fricativeness

/s/ and /z/ are fricative consonants, so they has high energy of high-frequency component. Therefore, the energy of the high frequency component can be used to determine the fricative consonant. The following $F(t)$ is used as an index of friction. Note that t is a frame number.

$$F(t) = \frac{1}{f_{\mathrm{high}}^{\mathrm{F}} - f_{\mathrm{low}}^{\mathrm{F}} + 1} \sum_{i=f_{\mathrm{low}}^{\mathrm{F}}}^{f_{\mathrm{high}}^{\mathrm{F}}} \log |S(t,i)| \tag{1}$$

$S(t,i)$ represents the t-th frame and the i-th frequency component, and $|S(t,i)|$ is its amplitude. $f_{\mathrm{low}}^{\mathrm{F}}$ and $f_{\mathrm{high}}^{\mathrm{F}}$ are fixed threshold frequencies representing the range of high frequency components.

In the investigation of Li et al., the frequency serving as a clue to recognize /s/ and /z/ became high frequency components including around 8 kHz, and the frequency used as a clue of /ʃ, ʒ/ is 2-4 kHz [6]. So we use $f_{\mathrm{low}}^{\mathrm{F}} = 2$ kHz, and $f_{\mathrm{high}}^{\mathrm{F}}$ is maximum frequency.

C. Articulation point

Resonance of alveolar fricative consonant occurs in the space between teeth and lips. This space is short so the resonance frequency is high. Therefore, a spectral gradient can be used as a feature representing the shape of the amplitude spectrum. Jesus et al. utilized spectral gradients around the peak frequency to analyze consonants [7]. In this research, since /s/ and /z/ are targeted, the interval for calculating the gradient is a fixed range. The slopes of the log amplitude spectra from 0 Hz to $f_{\mathrm{low}}^{\mathrm{G}}$ Hz and $f_{\mathrm{high}}^{\mathrm{G}}$ Hz to 20 kHz are $G_{\mathrm{low}}(t)$ and $G_{\mathrm{high}}(t)$, respectively. Their units are dB/kHz. In [7], threshold frequencies are 6 kHz for /s, z/ and 4 kHz for /ʃ, ʒ/, so we use the same values: $f_{\mathrm{low}}^{\mathrm{G}} = 4$ kHz, $f_{\mathrm{high}}^{\mathrm{G}} = 6$ kHz. These gradients are calculated by simple regression. In /s, z/, the values of both $G_{\mathrm{low}}(t)$ and $G_{\mathrm{high}}(t)$ decrease, and in the other consonants, the value of either or both increases. Therefore, alveolar fricative consonant can be detected using the following $G(t)$.

$$G(t) = -\left(G_{\mathrm{low}}(t) + \alpha G_{\mathrm{high}}(t)\right) \tag{2}$$

α is a constant that adjusts the scale of $G_{\mathrm{low}}(t)$ and $G_{\mathrm{high}}(t)$.

D. Voicing

/z/ is a voiced sound with vocal cord vibration and /s/ is an unvoiced sound without it. Voiced speech has a harmonic structure whose fundamental frequency is the frequency of

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

vocal cord vibration. That is, the amplitude spectrum of voiced speech has peaks at equal intervals at integral multiples of the fundamental frequency. The following $V(t)$ is used as an index to detect such structures. $V(t)$ is a feature that fixes the fundamental frequency f_0 and adds the amplitude up to its N-th harmonic. In this research, $N = 5$. In order to reduce the influence of volume, the average value of amplitude is subtracted.

$$V(t) = \max_{f^{\mathrm{V}}_{\mathrm{low}} \le f_0 \le f^{\mathrm{V}}_{\mathrm{high}}} \sum_{i=1}^{N} (|S(t, if_0)| - m(t)) \tag{3}$$

$$m(t) = \frac{1}{Nf^{\mathrm{V}}_{\mathrm{high}} + 1} \sum_{i=0}^{Nf^{\mathrm{V}}_{\mathrm{high}}} |S(t, i)| \tag{4}$$

$f^{\mathrm{V}}_{\mathrm{low}}, f^{\mathrm{V}}_{\mathrm{high}}$ represents a possible range as a fundamental frequency of human voice. In the investigation of Titze et al., achievable fundamental frequency ranges of human are 90-450 Hz for male and 120-800 Hz for female [8]. So we use $f^{\mathrm{V}}_{\mathrm{low}} = 90$ Hz, $f^{\mathrm{V}}_{\mathrm{high}} = 800$ Hz.

E. s-indicator and z-indicator

The s-indicator and z-indicator are configured using the above three indices. After the three indicators are calculated from the frequency spectrum of audio input, each is clipped with an appropriate threshold and normalized to 0 or more and 1 or less.

The clipping threshold of $[-50, -25]$ was used for $F(t)$, $[22, 27]$ for $G(t)$, and $[0.5, 2.5]$ for $V(t)$, respectively. /s/ is friction sound when gum sound is silent and /z/ is friction sound when gum sound is vocal. Therefore, s-indicator $I_s(t)$ and z-indicator $I_z(t)$ are configured by the following expressions.

$$I_s(t) = \min\{F(t), G(t), 1 - V(t)\} \tag{5}$$

$$I_z(t) = \min\{F(t), G(t), V(t)\} \tag{6}$$

The value of $I_s(t)$ is higher when /s/ is pronounced, and the value of $I_z(t)$ is higher for /z/. If neither is pronounced, both values will be smaller. These indicators can explain the reason when they become smaller than intended. For example, when the value of $I_s(t)$ is small even though /s/ is pronounced, the breath or narrowing is insufficient if $F(t)$ is small, the positions of teeth or tongue is bad if $G(t)$ is small, and vocal cord vibration has occurred if $V(t)$ is large.

III. Experiment and Result

In order to confirm whether $I_s(t), I_z(t)$ work properly, experiments were performed to calculate their values for speech uttering various phonemes. Fig. 1 is an example of the value of each index. The voice used is one in which /s, z, ʃ, ʒ, ç, ɸ, h, u/ is pronounced in order. The sampling rate is 44.1kHz. Fig. 1 shows that the value of $F(t)$ is large except for the vowel (/u/), the value of $G(t)$ is large only with /s/ and /z/, and the value of $V(t)$ is large only with voiced sound (/z, ʒ, u/). Also, the value of $I_s(t)$ is increased by /s/ alone, and the value of $I_z(t)$ is increased by /z/ alone.

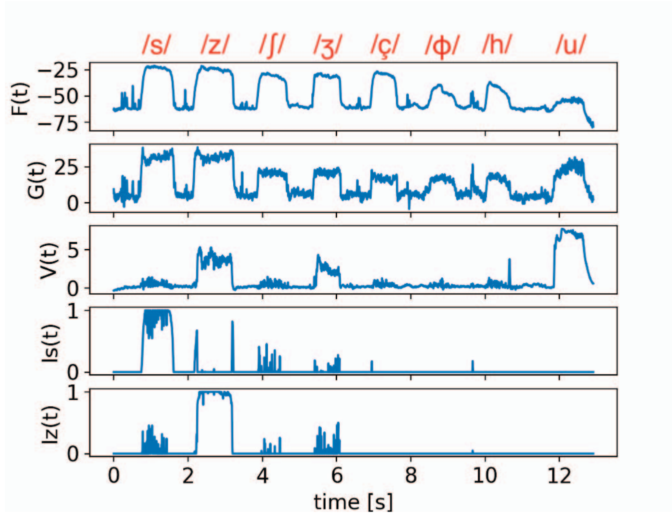


Fig. 1. The values of $F(t), G(t), V(t), I_s(t), I_z(t)$.

IV. Conclusion

We divided the characteristics of alveolar frictional consonant into fricativeness, articulation point, voicing, and created an index to measure them. Thereby, we created s-indicator and z-indicator with feature values that can be interpreted easily, and confirmed that they work properly. This index is used for speech training of hearing loss people. As a future work, we intend to create features that can be interpreted for other friction consonants. In addition, we want to apply to the field of the speech training and to carry out a demonstration experiment.

References

[1] C. C. Morton and W. E. Nance, "Newborn hearing screening—a silent revolution," *New England Journal of Medicine*, vol. 354, no. 20, pp. 2151–2164, 2006.

[2] F. R. Adams, H. Crepy, D. Jameson, and J. Thatcher, "Ibm products for persons with disabilities," in *1989 IEEE Global Telecommunications Conference and Exhibition'Communications Technology for the 1990s and Beyond'*. IEEE, 1989, pp. 980–984.

[3] K. Vicsi, P. Roach, A. Öster, Z. Kacic, P. Barczikay, A. Tantos, F. Csatári, Z. Bakcsi, and A. Sfakianaki, "A multimedia, multilingual teaching and training system for children with speech disorders," *International Journal of speech technology*, vol. 3, no. 3-4, pp. 289–300, 2000.

[4] O. Bälter, O. Engwall, A.-M. Öster, and H. Kjellström, "Wizard-of-oz test of artur: a computer-based speech training system with articulation correction," in *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2005, pp. 36–43.

[5] A. Grossinho, I. Guimaraes, J. Magalhaes, and S. Cavaco, "Robust phoneme recognition for a speech therapy environment," in *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 2016, pp. 1–7.

[6] F. Li, A. Trevino, A. Menon, and J. B. Allen, "A psychoacoustic method for studying the necessary and sufficient perceptual cues of american english fricative consonants in noise," *The Journal of the Acoustical Society of America*, vol. 132, no. 4, pp. 2663–2675, 2012.

[7] L. M. Jesus and C. H. Shadle, "A parametric study of the spectral characteristics of european portuguese fricatives," *Journal of Phonetics*, vol. 30, no. 3, pp. 437–464, 2002.

[8] I. Titze, T. Riede, and T. Mau, "Predicting achievable fundamental frequency ranges in vocalization across species," *PLoS computational biology*, vol. 12, no. 6, p. e1004907, 2016.

Adaptive line enhancer-based beat noise suppression for FM radio in motor vehicle

1st Kosuke Hasada
Grad. Sch. of Eng. Sci.
Osaka University
Osaka, Japan
hasada@sip.sys.es.osaka-u.ac.jp

2nd Arata Kawamura
Info. Sci. and Eng.
Kyoto Sangyo University
Kyoto, Japan
kawamura@cc.kyoto-su.ac.jp

3rd Youji Iiguni
Grad. Sch. of Eng. Sci.
Osaka University
Osaka, Japan
iiguni@sys.es.osaka-u.ac.jp

***Abstract*—In FM-radio on motor vehicles, there exists an interference called as a beat noise which is caused by equipped electronic control units. We propose a method which suppresses a beat noise using VAD (Voice Activity Detection) and ALE (Adaptive Line Enhancer). The ALE adaptively eliminates the beat noise in noise-only segments obtained by VAD. The ALE whose coefficients are automatically fixed when the variance of the ALE estimation error becomes sufficiently small in noise-only segments. In speech existence segments, the fixed ALE is used. Since the proposed ALE is automatically fixed after convergence, it can reduce speech degradation caused by misjudgement of VAD. Unfortunately, the ALE often adds a pseudo beat noise in speech existence segments. Thus, when the variance of the output signal is greater than the variance of the input signal, we remove the ALE. Simulation results show that the proposed method improves SNR in comparison to a conventional method.**

***Index Terms*—FM-radio, Beat Noise, Adaptive Line Enhancer**

I. Introduction

An FM radio receiver in a motor vehicle is an important device that provides the driver with useful information. Recently, many electronic devices exist in a motor vehicle and they generate electromagnetic noise which interferes with FM radio [1], [2]. Such interference is called as a beat noise. A beat noise degrades the sound quality of FM radio. This paper focuses on a software approach to suppress the beat noise.

We conventionally proposed to suppress a beat noise using VAD (Voice Activity Detection) and two ALEs (Adaptive Line Enhancers) [3]. With the energy of an input signal to the ALE, the VAD judges a noise-only segment or a speech segment which includes the speech signal. In the conventional method, the first ALE adaptively eliminates the beat noise in the noise-only segments. The first ALE coefficients are copied to the second ALE at certain intervals in the noise-only segments. In the speech segments, the second ALE is used as a fixed filter. Since the second ALE coefficients are copied from the first ALE in noise-only segment, the second ALE does not work well when the first ALE coefficients are copied before convergence. The main problem is that the coefficient update of the second ALE is done independently with the convergence of the first ALE. To avoid such undesired situation, the ALE coefficients used in the speech segments should be related with the convergence of the ALE coefficients.

In this paper, we improve the beat noise suppression capability by using a single ALE, while the conventional method requires two ALEs. In the proposed method, the ALE adaptively eliminates the beat noise in the noise-only segments, and its coefficients are automatically fixed when the ALE estimation error variance becomes sufficiently small. In the speech segments, the ALE is used where the ALE coefficients are not updated to eliminate the beat noise and let through the speech signal. Unfortunately, the ALE often increases noise for some reasons, e.g., the beat noise frequency is changed, the ALE coefficients are not appropriate. Hence, when the ALE output variance is greater than the input variance, we remove the ALE.

II. Beat noise suppression using ALE

Let $m(t)$ be a broadcasting signal at continuous time t. We have the FM signal $x_{FM}(t)$ as [3]

$$x_{\mathrm{FM}}(t) = A_c \cos\left\{2\pi f_c t + \beta_f \int_0^t m(\tau)d\tau\right\}, \qquad (1)$$

where A_c and f_c are the amplitude and frequency of the carrier signal, respectively. The parameter β_f is a constant to adjust the power of $m(t)$.

In the FM demodulation process, the two signals are obtained from FM signal as [3]

$$I(n) = \frac{A_c}{2}\cos\{\sum_{i=0}^{n} m(i)\} + \sum_{p=1}^{P}\frac{A_p}{2}\cos\{\theta_p(n)\} + \omega(n), \qquad (2)$$

$$Q(n) = \frac{A_c}{2}\sin\{\sum_{i=0}^{n} m(i)\} + \sum_{p=1}^{P}\frac{A_p}{2}\sin\{\theta_p(n)\} + \omega(n), \qquad (3)$$

$$\theta_p(n) = 2\pi(f_p - f_c)n + \varphi_p, \qquad (4)$$

where $I(n)$ and $Q(n)$ denote In-phase signal (I signal) and Quadrature-phase signal (Q signal) respectively. A_p, f_p, ϕ_p denote amplitude, frequency, and initial phase of the pth beat noise, respectively. Here, the beat noise is modeled as a sinusoid. The number of the beat noise is P. $\omega(n)$ represents the additive white Gaussian noise. We remove the respective noise signals from I and Q signal.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Similar to [3], we use the VAD to judge noise-only segments and speech segments based on the energy of I signal. The proposed method updates the ALE coefficients when the variance of the ALE estimation error is large in the noise-only segment.

Let a variance of white Gaussian noise be σ_w^2. We calculate the variance of the ALE estimation error as

$$\bar{\sigma}_e^2(n) = \lambda\bar{\sigma}_e^2(n-1) + (1-\lambda)e^2(n), \quad (5)$$

$$e(n) = I(n) - I_{ALE}(n), \quad (6)$$

$$\lambda = 1 - \frac{1}{2K}, \quad (7)$$

where $I_{ALE}(n)$ is the ALE output of I signal and $e(n)$ is the ALE estimation error. K denotes the filter order of the ALE. When the ALE completely removes the beat noise in the noise-only segments, we have $\sigma_e^2(n) = \sigma_w^2$ where $\sigma_e^2(n)$ is the true variance of the ALE estimation error. Hence, we stop the update of the ALE coefficients when $\bar{\sigma}_e^2(n) \leq \sigma_w^2$, where we assume that σ_w^2 is known.

The energy-based VAD may not absolutely catch the first part of an utterance as the speech segment, because the energy of the first part is usually small. In such part, $\bar{\sigma}_e^2(n) > \sigma_w^2$ may be satisfied, since the input signal involves speech signal. But, it is better to not update the ALE in such part.

We attempt to give the ALE the robustness against the VAD misjudgement. In the proposed method, to update the ALE again in the noise-only segments, it requires to continuously satisfy $\bar{\sigma}_e^2(n) > \sigma_w^2$ in a certain time length T_a. The time length T_a should be set greater than the time length of the VAD misjudgement on the first part of utterance.

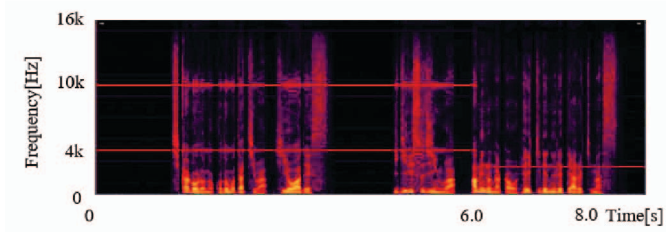
In speech segments, we never update the ALE. When the beat noise frequencies are not changed, the beat noise suppression is appropriately performed. Unfortunately, when the beat noise frequencies are changed in the speech segments, the ALE works to add a pseudo beat noise as an adverse effect.

To solve this problem, we remove the ALE when $\bar{\sigma}_e^2(n)$ is greater than the variance of $I(n)$, $\sigma_I^2(n)$. Thereby, we can avoid the undesired effect by the ALE, although it cannot eliminate the beat noise whose frequencies are changed in the speech segments. The remained beat noise is eliminated by the ALE which works in the next noise-only segments.

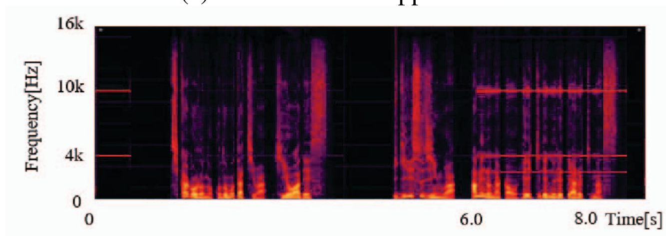
III. Simulation

We carried out the simulations to confirm the effectiveness of the proposed method by comparing with the conventional method under a virtual environment. The sampling frequency 32kHz is employed. We put the beat noise frequencies as 4kHz and 10kHz for 0s-6s. The number of the beat noise is changed from two to one at 6s, and the changed beat noise frequency is 2.5kHz. We put the filter order as $K = 4000$,$T_a = 0.2$s, the carrier to noise ratio as 90dB, and σ_w^2 as 4.5×10^{-13}.

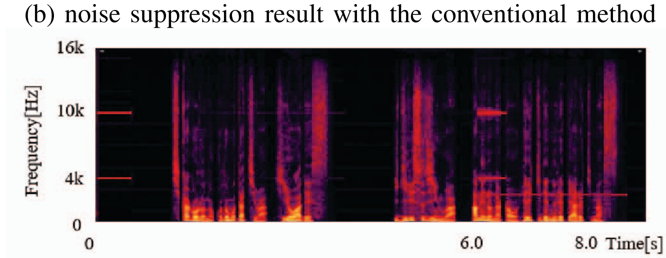
The spectrograms of simulation results are shown in Fig. 1, where (a) shows the demodulated signal without noise suppression, and (b) shows the noise suppression result with the conventional method, (c) shows the result with the proposed method. From Fig. 1 (b), we see that three beat noise signals



(a) without noise suppression



(b) noise suppression result with the conventional method



(c) noise suppression result with the proposed method

Fig. 1. Spectrograms of demodulated signals

exist from 6s to the next noise-only segment. From Fig. 1 (c), we see that the two beat noise signals are eliminated, although the beat noise of 2.5kHz remains.

We evaluate the noise suppression performance by using the signal to noise ratio(SNR). SNR is objective evaluation and a large value denotes a good result. With the conventional method, the SNR of the demodulated signal is 19.0dB. With the proposed method, the SNR is 23.6 dB and improved 4.6 dB from the conventional method.

IV. Conclusion

In this paper, we proposed a beat noise suppressor using the ALE whose coefficients are updated when the ALE estimation error variance is greater than the variance of white Gaussian noise in noise-only segments. The simulation results show that the proposed method improves SNR of 4.6 in comparison to the conventional method.

References

[1] Y. Mabuchi, A. Nakamura, T. Hayashi, K. Ichikawa, T. Uno, and H. Mizuno, "A feasibility study for reducing common-mode current on the wire harness connected to electronic control units," IEICE Trans. Electron., vol.J89–C, no.11, pp.854–865, Nov. 2006.

[2] T. Maeno, T. Unou, and K. Katoh, "Reduction of conductive noise currents through wire–harnesses from electronic equipment for vehicles," IEICE Trans. Commun., vol.J90–B, no.4, pp.437–441, Apr. 2007.

[3] K. Hasada, A. Kawamura and Y. Iiguni, "Adaptive beat noise estimation for FM radio in motor vehicle ," IEICE Technical Report., vol.117, no.516, pp.291–296, May 2018.

Adaptive Direct Blind Equalization under Noisy Environment

1st Minoru Komatsu
Suwa University of Science
Chino, Japan
m_komatsu@rs.sus.ac.jp

2nd Nari Tanabe
Suwa University of Science
Chino, Japan
nari@rs.sus.ac.jp

3rd Toshihiro Furukawa
Tokyo University of Science
Katsushika-ku, Japan
tofurukawa@ed.tus.ac.jp

***Abstract*—This paper proposes adaptive direct blind equalization under noisy environment. There is direct blind equalization corresponding to noisy environment using Rayleigh Quotient [2] as one of the conventional method of direct blind equalization method. It is known that the conventional method is higher precision even when the observed signals is included the noises. However, the conventional method has the problem that there is many computational complexity because this method have to calculate inverse matrix. In this paper, for solving this problem, we propose adaptive direct blind equalization under noisy environment. The features of proposed method are (i) realization of high performance and stable equalizer estimation with Data Least square(DLS) and (ii) reduction of computational complexity using gradient method. We show the effectiveness of the proposed method using computer simulation.**

***Index Terms*—Direct Blind Equalization, Data Least Square, Zero-Forcing, Gradient method, Noisy environment**

I. Introduction

Recently, in digital communication, we need the communication technology which fast sending and receiving to digital data. In wireless digital communication, the transmitted signals are distorted by the reflection and diffraction. This is called Fading, and the solution of this problem is studied actively. There is direct blind equalization corresponding to noisy environment using Rayleigh Quotient [2] as one of the method for solving this problem. It is known that the conventional method is higher precision even when the observed signals is included the noises. However, the conventional method has the problem that there is many computational complexity because this method have to calculate inverse matrix. In this paper, for reducing the computational complexity, we propose adaptive direct blind equalization under noisy environment. We show the effectiveness of the proposed method using computer simulation.

II. Preparation

In this section, we describe about the system model of digital communication and blind ZF equalization. We use system model in single input multiple output. Here, we apply L-times Over-sampling to the received signals. The observed signals vectors $\mathbf{x}_N(n)$ with which N is stacking number are as follows:

$$\begin{aligned}\mathbf{x}_N(n) &= [\ \mathbf{x}(n) \quad \mathbf{x}(n-1) \quad \cdots \mathbf{x}(n-N+1)\ ]^T \\ &= \mathcal{H}\mathbf{s}_N(n) + \mathbf{v}_N(n),\end{aligned} \tag{1}$$

where $\mathcal{H}$ is block toeplitz matrix consist of channel vectors, $\mathbf{s}_N(n)$ are the transmitted signals vectors and $\mathbf{v}_N(n)$ are the noises vectors. We discribe Zero-Forcing(ZF) criterion. ZF criterion is satisfied by

$$\mathbf{A}^H\mathbf{g} = \mathbf{e}_{LN+1} = [0\cdots0 \ \ 1 \ \ 0\cdots0]^T \ , \tag{2}$$

where

$$\mathbf{A}^H \ = [\mathbf{R}_{11} - \mathbf{W}_{12}\mathbf{R}_{22}^{+}\mathbf{W}_{21} \ \vdots\ \mathbf{u} \ \vdots\ \mathbf{W}_{12}\mathbf{R}_{22}^{+}\mathbf{W}_{21}]^H$$

and, each elements are as follows:

$$\left.\begin{array}{ll}\mathbf{R}_{11} & = \mathbf{E}[\mathbf{x}_N(n)\mathbf{x}_N^H(n)]\ , \\ \mathbf{R}_{22} & = \mathbf{E}[\mathbf{x}_M(n-d)\mathbf{x}_M^H(n-d)]\ , \\ \mathbf{R}_{12} & = \mathbf{E}[\mathbf{x}_N(n)\mathbf{x}_M^H(n-d)]\ , \\ \mathbf{R}_{21} & = \mathbf{R}_{12}^T, \\ \mathbf{W}_{12} & = \mathbf{E}[\mathbf{x}_N(n-1)\mathbf{x}_N^H(n-d)]\ , \\ \mathbf{W}_{21} & = \mathbf{W}_{12}^T, \\ \mathbf{u} & : The\ right\ singluar\ vector \\ & \quad corresponding\ to\ maximum\ singluar\ value\ of \\ & \quad \mathbf{R}_{12}\mathbf{R}_{22}^{+}\mathbf{R}_{21} - \mathbf{W}_{12}\mathbf{R}_{22}^{+}\mathbf{W}_{21}\end{array}\right\}$$

and $\mathbf{e}_{LN+1}$ is the vector that $LN+1$th element is 1 and the other elements are 0. Here, we use matrix $\mathbf{A}^H$ instead of $\mathcal{H}$. In (2), using pseudo channel matrix of left pseudo inverse matrix $\mathbf{A}^{+}$, $\mathbf{g}$ is as follows:

$$\mathbf{g} = \mathbf{A}^{+}\mathbf{e}_{LN+1}. \tag{3}$$

When the observed signals $\mathbf{x}_N(n)$ are not included the noises, It is known that we obtain the high recovery signals using (3). However, because the conventional method is not considered the noises which are included the observed signals, the recovery performance of signals is lower when the observed signals $\mathbf{x}_N(n)$ are included the noises. Next section, to solving this problem, we propose direct blind equalization method with suppressing the noises.

III. Proposed method

In this section, we propose Adaptive Direct Blind Equalization under noisy environment. The signals of recovery performance is lower in case of the observed signals is included the noises, because $\mathbf{A}^H$ is included he noises. To solving this problem, we need construct new model which was able to suppress the noises in $\mathbf{A}^H$. In Fig.1, we show the

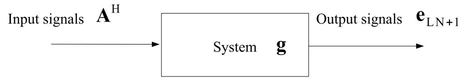


Fig. 1. input-output model for suppressing noises

TABLE I
THE COMPARISON OF COMPUTATIONAL COMPLEXITY

<table>
<tr><th></th><th>conventional method</th><th>proposed method</th></tr>
<tr><td>step.1(A)</td><td colspan="2">$(LN)^3 + (LN)^2 + (LM)^2$
$+(LN) \times (LM) \times (LN + LM + 1)$</td></tr>
<tr><td>step.2(g)</td><td>$2(LN)^3$
$+2(2LN + 1)$
$\times(LN)^2$
$+2(LN)$
$\times(2LN + 1)^2$
$+2(LN)^2 - (LN) + 4$</td><td>$2(LN)^3 + 2(LN)^2$
$+10(LN) + 6$</td></tr>
<tr><td>total</td><td>$15(LN)^3$
$+(LM + 12) \times (LN)^2$
$+(LN + 1) \times (LM)^2$
$+(LM + 1) \times (LN) + 4$</td><td>$4(LN)^3$
$+(LM + 7) \times (LN)^2$
$+(LN + 1) \times (LM)^2$
$+(LM + 11) \times (LN) + 6$</td></tr>
</table>

model of input-output system for suppressing noises. From Fig.1, we can assume input of system to $\mathbf{A}^H$. Here, there is DLS algorithm as the method which is suppressed noises of $\mathbf{A}^H$. By applying this algorithm to Direct Blind equalization method, we can obtain direct blind equalization method with suppressing the noises. The cost function of DLS algorithm is as follow:

$$J_{DLS} \quad = \|\Delta \mathbf{A}^H\|_F^2 = \frac{\tilde{\mathbf{g}}^H \mathbf{C}^H \mathbf{C} \tilde{\mathbf{g}}}{\tilde{\mathbf{g}}^H \mathbf{D} \tilde{\mathbf{g}}}, \tag{4}$$

where

$$\left.\begin{array}{ll} \tilde{\mathbf{g}} & = [\mathbf{g}^T \ \ -1]^T, \\ \mathbf{C} & = [\mathbf{A}^T \vdots \, \mathbf{e}_{LN+1}], \\ \mathbf{D} & = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}. \end{array}\right\} \tag{5}$$

Using gradient method to (4), the update of equalizer $\mathbf{g}$ suppressing the noises is as follow:

$$\begin{aligned} \mathbf{g}_n = & \mathbf{g}_{n-1} \\ & - \alpha[\{(\mathbf{A}^T)^H \mathbf{A}^T - \lambda \mathbf{I}\}\mathbf{g}_{n-1} - (\mathbf{A}^T)^H \mathbf{e}_{LN+1}], \end{aligned} \tag{6}$$

where

$$\lambda = \frac{\tilde{\mathbf{g}}^H \mathbf{C}^H \mathbf{C} \tilde{\mathbf{g}}}{\tilde{\mathbf{g}}^H \mathbf{D} \tilde{\mathbf{g}}}. \tag{7}$$

IV. Simulations

In this section, we evaluate the proposed method using computational complexity and Bit Error Ratio (BER).

A. computational complexity

In this section, we evaluate the computational complexity of proposed method. We show the computational complexity of conventional method and proposed method in Table I. Fig.2 shows computational complexity by diagram. As a result, we confirm that the proposed method is lower computational complexity and lower calculate time than the conventional method. The computational complexity is lower because the proposed method calculate the equalizer based on gradient method.

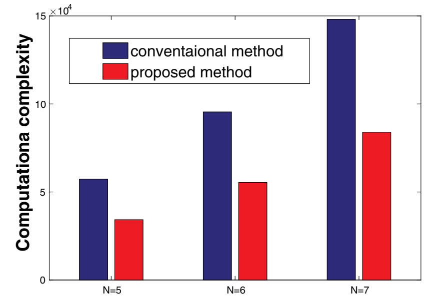


Fig. 2. comparison of computational complexity

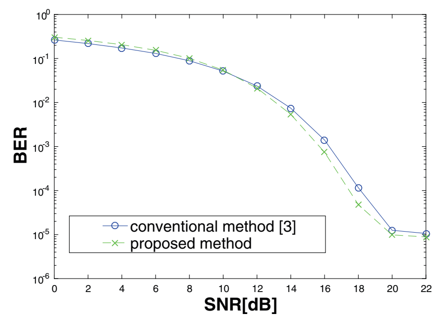


Fig. 3. Bit Error Ratio

B. Bit Error Ratio

In this section, we evaluate the proposed method using BER. Fig.3 shows simulation results. As a result, we confirm that the performance of proposed method is nearly equal the performance of conventional method. We prove that the proposed method has stable higher recovery performance of signals and few computational complexity because of a result of Fig.2 and Fig.3.

V. Conclusions

In this paper, we proposed adaptive Direct Blind Equalization under Noisy Environment and comfirmed the effectiveness of proposed method.

References

[1] Xiaohua Li and Howard Fan: "Direct Estimation of Blind ZF Equalizers Based on Second-Order Statistics", IEEE TRANSACTIONS ON SIGNAL PROCESSING, , VOL. 48, NO. 8, pp.2211-2218 (2000-8) "The Data Least Squares Problem and Channel Equalization", IEE Proc,Vis image signal Process., Vol. 147, No.3 (1993-6)

[2] Komatsu M, Tanabe N and Furukawa T: "Direct Blind Equalization Corresponding to Noisy Environment using Rayleigh Quotient ", International Colloquium on Signal Processing & Its Applications (CSPA), 2019 International Symposium, , (2019-03)

Low-power Bandgap Reference with Soft Startup for Energy Scavenging Applications

Nicodimus Retdian
Global Course on Engineering and Science
Shibaura Institute of Technology
Tokyo, Japan
nico@shibaura-it.ac.jp

***Abstract*—Band-gap reference (BGR) is one of the important and indispensable analog circuits in integrated circuits.The design of BGR has been widely known. In most cases, the power supply voltage is assumed to be a constant value with a variation of $\pm 10\%$. However, in the case of energy scavenging, the input voltage, which will be converted to the power supply voltage, varies in a wide range. Linear regulator such as Low Drop-Out Regulator (LDO) is used to produce the power supply inside the chip, however, the BGR cannot use LDO output as it power supply voltage since LDO uses BGR output as its reference voltage. In near field communication (NFC) systems, the input voltage varies with the distance between the reader and device. This work assumes the implementation of BGR in an NFC system where the supply voltage of BGR varies from 0V to 5V. A BGR start-up mechanism for a wide input voltage range is proposed. The proposed start-up mechanism is implemented in 0.18um CMOS process and verified using Spectre simulation. The proposed BGR operates at 1.35∼5.5V input voltage range. The temperature coefficient at a 1.7-V input voltage and a temperature range from -25 to 80°C is 21ppm/$^\circ$C.**

***Index Terms*—band-gap reference, energy scavenging, near field communication, start-up circuit, temperature compensation**

I. Introduction

Band-Gap Reference (BGR) produces a reference voltage independent of Process, Voltage and Temperature (PVT) variations. In an energy scavenging system such as wireless power transfer, the input power supply voltage of the system may varies in wide range thus it is necessary to design a BGR which can tolerate wide variation of input voltage. Power supple insensitive BGR uses self biasing technique which requires start-up mechanism. As an example, Banba's BGR [1] uses POR signal produced by the LDO regulator to start the BGR. However, this approach is not feasible in an energy scavenging system where BGR should be ready before the LDO operates.

This paper proposes a BGR dedicated for energy scavenging system with a soft start-up mechanism. The proposed BGR is implemented using enhancement type MOS devices with thick oxide in a standard CMOS process.

II. Banba's BGR

Assuming a first order temperature characteristic, the output voltage V_{ref} is given by

$$V_{ref} = \frac{R_3}{R_1 R_2}\left[R_1 \Delta V(1+\alpha T) + R_2 V_a(1+\beta T)\right] \qquad (1)$$

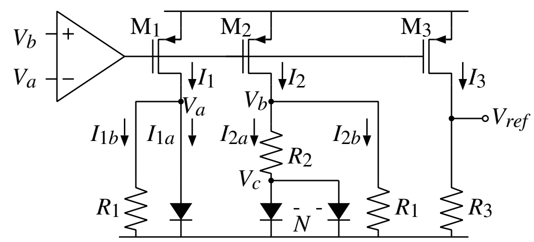


Fig. 1. Conventional BGR by Banba et.al. [1]

where, $\Delta V = V_T \ln(N), \alpha = 0.086\text{mV}/^\circ\text{C}, \beta = -2\text{mV}/^\circ\text{C}$. In order to make the temperature coefficient of V_{ref} equal to zero,

$$\frac{R_1}{R_2} = -\frac{V_a \beta}{\Delta V \alpha} = -\frac{\beta \ln(I_{1a}/I_{s0})}{\alpha \ln(N)} \approx 23.26 \ \ln(\frac{I_{1a}}{I_{s0}} - N).$$

Note that I_{1a}/I_{s0} should be larger than N to get a positive resistance ratio. Banba's BGR uses a Power On Reset (POR) signal to start it up. However, POR signal is commonly activated when LDO's output reaches a given threshold. Thus, it is not possible to use POR signal to start the BGR in an energy scavenging system. Furthermore, Banba's BGR uses depletion type MOS transistor as the input stage of the opamp to reduce the power supply voltage.

III. Proposed BGR

The proposed BGR is shown in Fig.2. The proposed BGR utilizes the Banba's BGR core, but R_1 is splited into two resistors R_{1a} and R_{1b} and the input of opamp is taken from the node between R_{1a} and R_{1b}. This will lower the common-mode input voltage of the opamp and thus allowing the use of enhancement type MOS transistors at the input of the opamp.

The opamp uses a self bias topology and need a start-up mechanism. Banba's BGR uses a POR signal to forcefully pull V_e down and start the circuit. However, this start-up technique is not feasible in an energy scavenging system. $M_a, M_c, M_d, R_a \sim R_d$ creates a self start-up circuit for the proposed BGR. During the start-up, $V_a = 0$V and R_d will pull V_g to V_{IN}. It will turn M_d on and pull V_e down to 0V. As a result currents will start to flow in the circuit and pull V_a up until M_c turns on and pulls V_g down near 0V such that M_d is turned off.

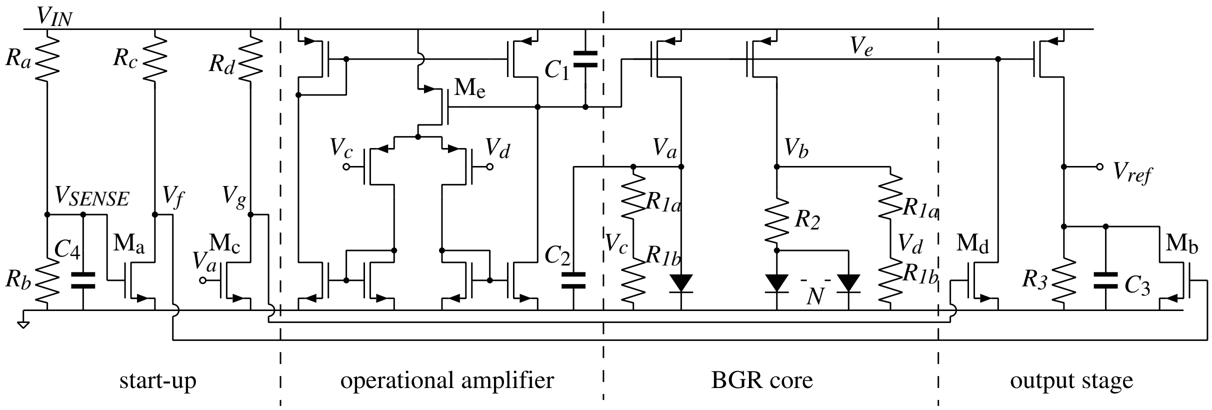


Fig. 2. Proposed BGR

The output voltage V_{ref} is held to 0V by M_b until V_{IN} reaches a sufficient threshold voltage to turn the LDO on. The value of the threshold voltage V_{tl} is given by

$$V_{tl} = \frac{R_a + R_b}{R_b} V_{thn} \quad (2)$$

where V_{thn} is the threshold voltage of the nMOS transistor. The proposed BGR uses thick oxide device (3.3V) to allow an input voltage up to 5.5V.

IV. Simulation Results

The proposed circuit is simulated using Spectre and 0.18um CMOS process parameters. The specifications for the BGR are summarized in Tab.I. Figure 3 shows the output voltage characteristic when the input voltage is swept from 0 to 5.5V. It also shows that the minimum input voltage of proposed BGR is 1.35V. For a typical 1.8V LDO regulator, the minimum operation voltage will be 1.7V. The output voltage variation for input voltage from 1.7 to 5.5V is approximately only 8mV.

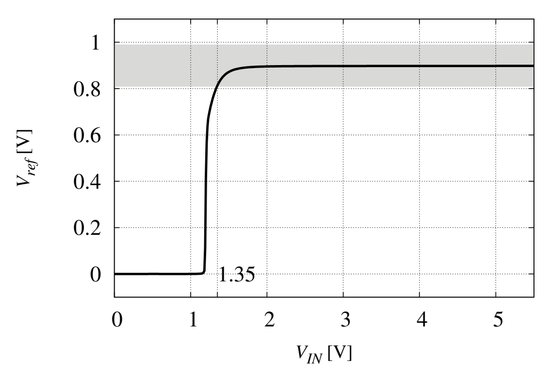


Fig. 3. Simulated DC characteristic ($V_{IN} = 0 \sim 5.5$V)

TABLE I
BGR specifications

Parameter	Value
V_{IN}	~5.5V
V_{ref}	0.9V±10%
Bias current	≤10uA
Temperature range	–25~80°C

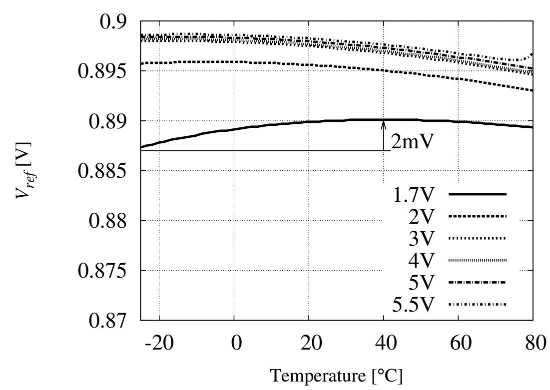


Fig. 4. Simulated temperature characteristic ($V_{IN} = 1.7 \sim 5.5$V)

Figure 4 shows the simulated temperature characteristic of the proposed BGR for input voltages from 1.7 to 5.5V. At $V_{IN} = 1.7$V, the output voltage variation is only 2mV. The average temperature coefficient of V_{ref} is approximately 21ppm/°C. The typical bias current is 5.8uA at $V_{IN} = 1.7$V and 8.9uA at $V_{IN} = 5.5$V. The minimum PSRR at 27.12MHz is 37dB.

V. Conclusions

A BGR start-up mechanism for a wide input voltage range is proposed. The proposed start-up mechanism is implemented in 0.18um CMOS process and verified using Spectre simulation. The proposed BGR operates at 1.35~5.5V input voltage range. The temperature coefficient at a 1.7-V input voltage and a temperature range from -25 to 80°C is 21ppm/°C.

References

[1] H. Banba,et.al, ”A CMOS Bandgap Refernce Circuit with Sub-1-V Operation”, IEEE J. Solid-State Circuits, vol.34, no.5, pp.670-674, May 1999

A broadband low noise amplifier for high performance wireless microphones

Shui-Chuen Lin
Department of Electronic Engineering
National Chin-Yi University of Technology
Taichung, Taiwan, R.O.C.
lsc@ncut.edu.tw

Chang-Lang Tai
Department of Electronic Engineering
National Chin-Yi University of Technology
Taichung, Taiwan, R.O.C.
tai50626@yahoo.com.tw

San-Fu Wang
Department of Electronic Engineering
National Chin-Yi University of Technology
Taichung, Taiwan, R.O.C.
sf_wang@ncut.edu.tw

Yitsen Ku
Department of Electical Engineering
California State University Fullerton
California, U.S.A.
ORCID: 0000-0002-6141-187x

Zi-Wei Wang
Department of Electronic Engineering
National Chin-Yi University of Technology
Taichung, Taiwan, R.O.C.
bobsam14@gmail.com

Shih-chun Liu
Department of Electronic Engineering
National Chin-Yi University of Technology
Taichung, Taiwan, R.O.C.
qazqazg8@gmail.com

***Abstract*—In this paper, a new CMOS broadband low-voltage low-noise amplifier (LNA) is proposed with an operating frequency range of 100MHz–2.5GHz with current reuse, resistor feedback, inductor feedback, mirror bias, inductor peaking and source inductive degeneration techniques. The proposed LNA uses inductor feedback and gain boosted technique to increase its high frequency gain to achieve high gain and wide bandwidth. It uses source inductive degeneration technique to reduce its noise figure and improve its input matching performance. Moreover, it also uses current reuse and mirror bias techniques to reduce its power consumption. Therefore, the LNA with high gain, wide bandwidth, good gain flatness performance and low power consumption by proposed method. The simulation results show that the power gain is around 10dB (s21), the NF is less than 2.5dB, the input reflection coefficient (S11) is less than -10.5 and the 1dB compression point (P1dB) is about -11dBm. The LNA consumes maximum power at about 5.5mW with a 1V power supply. The proposed LNA is implemented in a TSMC 0.18um RF COMS process.**

Keywords*— *low noise amplifier, low-power, gain flatness, mirror bias technique, broadband.

I. INTRODUCTION

The application of a wireless communication is becoming more and more important and popular, especially in portable wireless electronic products [1]. However, when more and more wireless electronic products are used, this means that mutual interference between electronic products will become more and more serious. Wireless microphone is one of the typical applications of wireless electronic products, and they also encounter mutual interference problems. A good solution to the above problem is to integrate several different specifications into the same terminal. For example, if 400MHz, 900MHz, and 2.4GHz are integrated into the same terminal, data transmission can be performed using the 900MHz band or the 400MHz band, and the 2.4GHz band encounters severe interference. This approach can effectively solve the interference problem. Broadband receivers are required when integrating multiple different specifications into the same terminal, while wideband low noise amplifiers (LNAs) are an integral part of wideband receivers [2].

Low noise amplifier is one of the key circuit blocks in radio receiver system, and is widely used in these wireless electronic products to amplify weak signal from the antenna. Therefore, the LNA is usually added directly to the first stage of the receive path, which usually dominates the NF and operating bandwidth in the receiver.

Moreover, in order to increase the use time of wireless electronic products and reduce the cost of wireless electronic products, monolithic radio frequency integrated chip (RFIC) with low supply voltage and low power consumption has been increasingly concerned [3-4]. Therefore, the paper will design a broadband LNA.

II. PROPOSED CIRCUIT

In this work, a new CMOS broadband low noise amplifier (LNA) with current reuse, resistor feedback, inductor feedback, mirror bias, inductor peaking and source inductive degeneration techniques is proposed. Fig. 1 shows the proposed broadband LNA circuit.

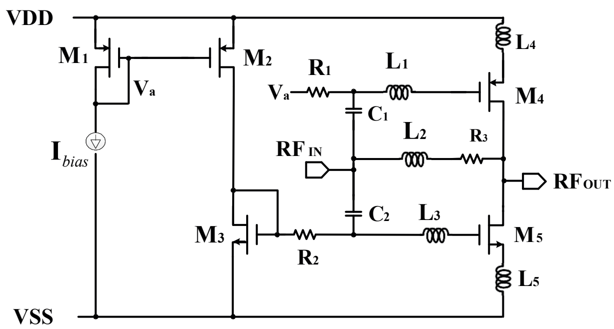


Fig. 1. The proposed wideband LNA.

The proposed broadband LNA circuit integrates inductor peaking and inductor shunt feedback techniques to achieve the requirements of high gain at high frequency, where L_1 and L_3 are used to implement the inductor peaking technique, and the L_2 is used to implement the inductor shunt feedback technique [5]. Moreover, the proposed broadband LNA circuit integrates source inductive degeneration and resistor shunt feedback techniques to achieve broadband input matching requirements [6-7], where L_4 and L_5 are used to implement source inductive degeneration technique, and the R_3 is used to implement resistor shunt feedback technique.

In addition, M_1, M_2 and M_3 are used to generate mirror bias voltages for M_4 and M_5, and C_1 and C_2 are used to isolate the feedback DC voltage and bias voltages. Finally, the current of the PMOS transistor, M_4, is reused by the NMOS transistor, M_5, so the power consumption of the proposed LNA is decreased. Based on the above principles, the proposed LNA has a method of high gain, wide bandwidth, good gain flatness performance, and low power consumption.

III. Simulation Results

The broadband LNA is designed using a TSMC 0.18 µm RF CMOS process. Fig. 2 through Fig. 6 illustrate the simulation results of input reflection coefficient, power gain, noise figure and 1 dB compression point.

Fig. 2 shows the simulation results for the input reflection coefficient (S11), which verifies the circuit input impedance of Fig. 1, where the circuit input impedance is close to 50Ω in the band between 100 MHz and 2500MHz. Fig. 3 shows the simulation result of the LNA power gain (S21), which is better than 10.2 dB in the bandwidth, and has a gain variable less than 2dB. Fig. 4 shows the simulation result for the overall LNA noise figure, and the noise figure is below 2.8 dB in the bandwidth.

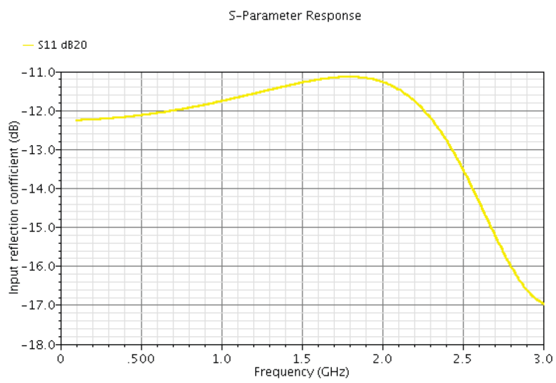


Fig. 2. Input reflection coefficient of proposed wideband LNA.

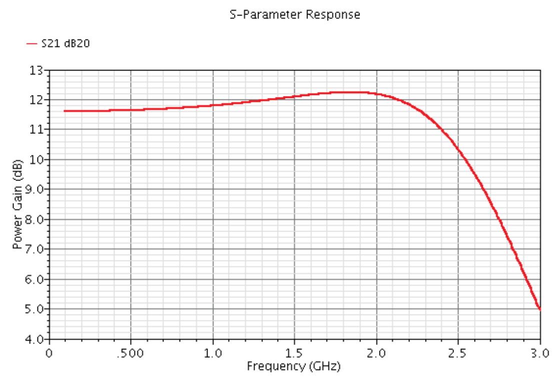


Fig. 3. Power gain of proposed wideband LNA.

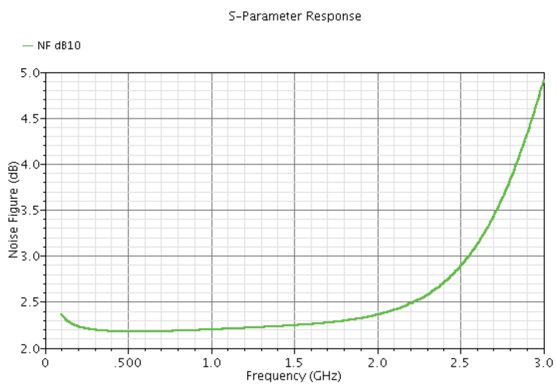


Fig. 4. Noise figure of proposed wideband LNA.

In order to check the linearity requirement, the 1-dB compression point (P1dB) performance of the LNA is -11.92 dB at 1000 MHz operation frequency, and the simulation result is shown in Fig. 5. The simulation results show that the proposed broadband LNA circuit can reduce the noise figure and improve the gain performance. Moreover, it has good power consumption performance.

The broadband low noise amplifiers can transmit data using the frequency band with less interference noise. For example, it can use the 900 MHz band or the 400 MHz band for data transmission and severe interference in the 2.4GHz band. Therefore, it is suitable for wireless microphone applications. The LNA consumes about 5.5mW of the power from 1 V power supply. Table 1 compares the proposed wideband LNA with the other wideband LNAs.

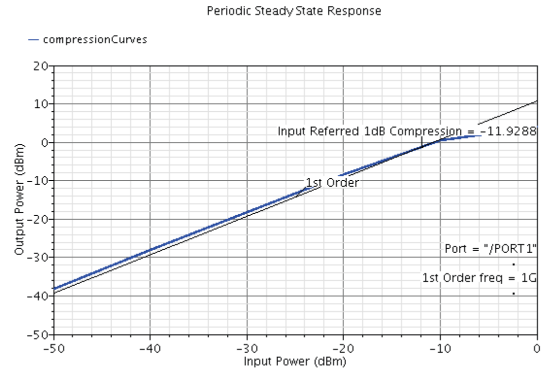


Fig. 5. The simulated 1-dB compression point of proposed LNA.

TABLE I. The Comparisons of the Proposed Wideband LNA and the Other Wideband LNAs.

	[2] 2013	[7] 2011	This work
CMOS Process	0.18µm	0.18µm	0.18µm
BW (GHz)	0.7~2.7	0.47~3	0.1~2.5
Supply Voltage (V)	1.8	1.8	1.0
S11 (dB)	-10	<-10	<-11.2
Gain (dB)	14~17	11.3~13.6	10.2~12.4
NF (dB)	<2.5	<2.5	<2.8
P1dB (dBm)	-8.5	-13	-11.92
Gain flatness	3	2.3	<2
Power (mW)	13.5	27	5.5

References

[1] M. Parvizi, K. Allidina, and M. N. El-Gamal, "A Sub-mW, Ultra-Low-Voltage, Wideband Low-Noise Amplifier Design Technique," Trans. on Very Large Scale Integration (VLSI) Systems, Vol. 23, No. 6, pp. 1111-1122, 2015.

[2] O. A. Hidayov, N. H. Nam, G. Yoon, S. K. Han, and S. G. Lee, "0.7-2.7 GHz wideband CMOS low-noise amplifier for LTE application." Electronic Letter, Vol. 49, No. 23, pp. 1433-1435, 2013.

[3] F. Chen, W. Zhang, W. Rhee, J. Kim, Kim, D. and Z. Wang, "A 3.8-mW 3.5–4-GHz Regenerative FM-UWB Receiver with Enhanced Linearity by Utilizing a Wideband LNA and Dual Bandpass Filters," IEEE Trans. Microw. Theory Techn., Vol. 61, No. 9, pp. 3350-3359, 2013.

[4] Y. S. Hwang, S. F. Wang, and J. J. Chen "Design Method for CMOS Wide-band Low Noise Amplifier for Mobile TV Application," IEICE Electronics Express, Vol. 6, No. 24, pp. 1721–1725, 2009.

[5] S. F. Chao, J. J. Kuo, C. L. Lin, M. D. Tsai, and H. A. Wang, "dc-11.5 GHz low-power wideband amplifier using splitting-load inductive peaking technique." IEEE Microw. Wireless Compon. Letter, Vol. 18, No. 7, pp. 482-484, 2008.

[6] C. Feng, X. P. Yu, Z. H. Lu, W. M. Lim, and W. Q. Sui, "3–10 GHz self-biased resistive-feedback LNA with inductive source degeneration," Electronic Letter, Vol. 49, No. 6, pp. 387-388, 2013.

[7] S. F. Wang, Y. S. Hwang, S. C. Yan, and J. J. Chen, "A New CMOS Wideband Low Noise Amplifier with Gain Control," Integration - the VLSI Journal, Vol. 44, No. 2, pp.136-143, 2011.

Investigation of Hybrid ADC Combined with First-order Feedforward Incremental and SAR ADCs

Yuichiro Kobayashi
Department of Electrical Engineering
Graduate School of Science and Technology
Tokyo University of Science
Noda-shi, Chiba, Japan

Tatsuji Matsuura, Ryo Kishida, Akira Hyogo
Department of Electrical Engineering
Faculty of Science and Technology
Tokyo University of Science
Noda-shi, Chiba, Japan

***Abstract*—We propose a hybrid Analog to Digital Converter (ADC) combined with the first-stage as a first-order feedforward incremental (FF I-) ADC and the second-stage as a SAR (Successive Approximation Register) ADC as high-precision and high-speed ADCs. As a result of simulation, the proposed hybrid ADC correctly converts data even if capacitance in FF path of the first-order I-ADC has 10% error.**

***Keywords*—Hybrid ADC, Incremental ADC, Feedforward, SAR ADC**

I. Introduction

An Analog to Digital Converter (ADC) is used to measure the remaining amount of a lithium-ion battery in Battery Management System (BMS) [1]. An ADC with higher resolution is required to prevent overcharge and discharge of the battery. This paper aims 16-bit resolution, 50 µs conversion speed ADC as a high precision ADC for DC (Direct Current) voltage measurements of the BMS.

When I- (incremental) ADC is used for 16-bit resolution, conversion speed is very slow (2^{16} clock counts are needed) [2]. We propose a hybrid ADC of I- and SAR (Successive Approximation Resister) ADC [3]. The SAR ADC compensate for conversion speed in the hybrid ADC. When the first-stage resolution is 8 bits, this architecture achieves 256 times faster conversion speed compared to only I-ADC. However, it is difficult to transfer the quantization error only correctly to the second stage in a conventional feedback I-ADC [2].

In the first-order I-ADC, if the FF path is equipped, the integrator output has only the quantization error at the end of the I-ADC conversion. Therefore, as for the first stage of a hybrid ADC, first order I-ADC with FF path is better.

However, if there is the FF path has capacitance error, the integrator output is not only quantization error but it includes a part of the input signal. The errors in the first-stage must be corrected. So, we equipped a 1-bit enlarged conversion range or extension region [$-2V_{ref1}$, $+2V_{ref1}$] in the second-stage ADC to suppress the error. This is confirmed by a simulation using MATLAB/Simulink.

II. Proposed hybrid ADC

In the first-stage ADC as shown in Fig. 1, the integrator output voltage $V_{intout}(z)$ includes only quantization error when FF path weighting capacitor is correct.

Output voltage is

$$V_{out}(z) = V_{in}(z) + (1 - z^{-1})Q(z) \ , \qquad (1)$$

where $Q(z)$ is quantization error.

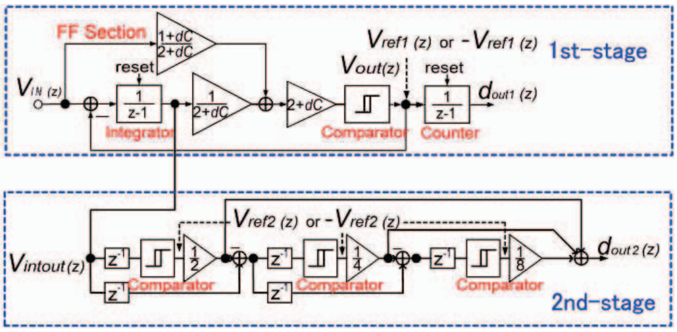


Fig. 1 Block diagram of the proposed hybrid ADC

Table 1 Simulation conditions

Input voltage V_{IN}	-1.0 to 1.0 V (100µV increments)
1st-stage clock count CL_1	5
2nd-stage clock count CL_2	4
1st-stage reference voltage V_{ref1}	1.0 V
2nd-stage reference voltage V_{ref2}	2.0 V

From Eq. (1), integrator output voltage is

$$V_{intout}(z) = \frac{z^{-1}}{1 - z^{-1}}\left(V_{in}(z) - V_{out}(z)\right)$$
$$= -z^{-1}Q(z) \ . \qquad (2)$$

Eq. (2) shows that the integrator output includes only the quantization error. The second stage ADC in Fig. 1 converts this quantization error.

Next, digital output of hybrid ADC d_{outall} is

$$d_{outall} = d_{out1} + d_{out2}, \qquad (3)$$

where d_{out1} and d_{out2} is digital output of first- and second-stage ADC. As shown in Eq. (3), the digital output of the hybrid ADC is a combination of the digital outputs of both stages. Further, if the capacitor of FF path has capacitance error, the integrator output shown in Eq. (2) is not only quantization error but it includes a part of the input signal. So, we equipped a 1-bit extension region [$-2V_{ref1}$, $+2V_{ref1}$] in the second-stage SAR ADC to compensate the error in the first-stage ADC's digital output.

III. Simulation

A. Simulation conditions

For easy confirmation of principle, both the first- and second-stage ADCs n-bit and m-bit resolutions are 2-bit resolution this time. The second-stage SAR ADC has 1-bit enlarged conversion range of [$-2V_{ref1}$, $+2V_{ref1}$] with reference voltage (V_{ref2}) to convert voltage of over first stage range [-V_{ref1}, V_{ref1}] and compensate the capacitance error in the first-stage ADC.

We simulate DC analysis with 0 or 10 % capacitance error dC in the first stage of FF using MATLAB/Simulink.

Where dC is capacitance error of FF section. And first-stage clock count is

$$CL_1 = 2^n + 1\,. \qquad (4)$$

B. Simulation results

Fig. 2 shows the simulation result without capacitance error and with 10 % capacitance error. The result of no capacitance error is equivalent to ideal output, which increments every 100 µV input voltage. Error occurs in the result of 10 % capacitance error. Fig. 3 shows the simulation result of the integrator output in the first-stage ADC. This result is equivalent to the digital output in the first-stage ADC as shown in Fig. 2. The maximum integrator output is 1.1 V due to 10 % capacitance error. Fig. 4 shows the digital output of the second stage SAR ADC. Extension region in SAR ADC converts voltage of over 1 V and compensate the capacitance error in the first-stage ADC. Fig. 5 shows the digital output of the hybrid ADC. The input voltage is correctly converted to the digital voltage regardless of the capacitance error. If an extension region is provided in the second stage, it is possible to correct a capacitance error in the digital output of the first-order FF I-ADC.

We also simulate when the extension region is not provided in the second stage, a digital output error of the hybrid ADC occurs due to the capacitance error.

IV. Conclusion

In this paper, we considered an ADC capable of achieving a resolution of 16 bits and a conversion speed of 50 µs and proposed a hybrid ADC in which the first stage is a first-order FF I-ADC and the second stage is a SAR ADC. When the first-stage is 8 bits, this architecture can achieve 256 times faster conversion speed compared to an I-ADC alone.

Next, by adopting the first-order FF I-ADC in the first stage, it was found that an error occurs in the digital output due to a capacitance error in the FF section. It was found that this error can be corrected by providing an extension region in the second stage SAR ADC. A/D conversion can be correctly performed in the proposed hybrid ADC.

Acknowledgement

This work is supported by JSPS Grant-in-Aid for Scientific Research (C) 17K06400.

References

[1] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, “Analog-to-digital converters,” IEEE Signal Processing Magazine, vol. 22, no. 6, pp. 69-77, Nov. 2005.

[2] J. Márkus, J. Silva, and G. C. Tames, “Theory and Applications of Incremental Converters,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 51, no. 4, pp. 678-690, Apr. 2004.

[3] P. Harpe, “Successive Approximation Analog-to-Digital Converters: Improving Power Efficiency and Conversion Speed,” IEEE Solid-State Circuits Magazine, vol.8, pp. 64-73, Nov. 2016.

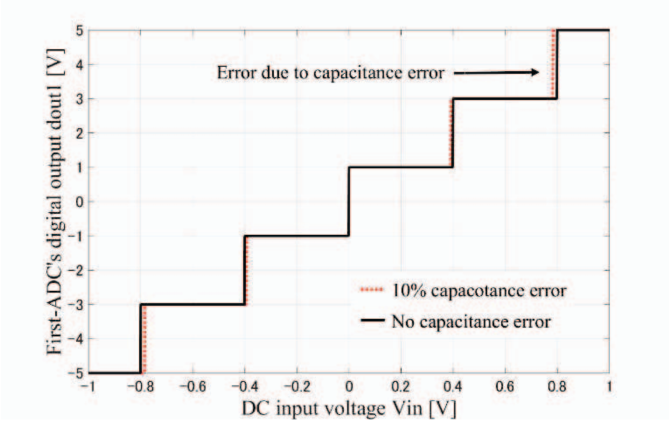


Fig. 2 Digital output of the first stage ADC ($\mathrm{dout1}(z)$)

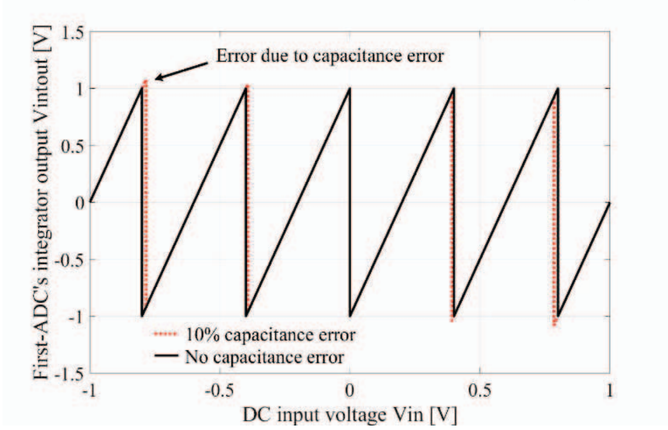


Fig. 3 Integrator output of the first-stage ADC ($V_{\mathrm{intout}}(z)$)

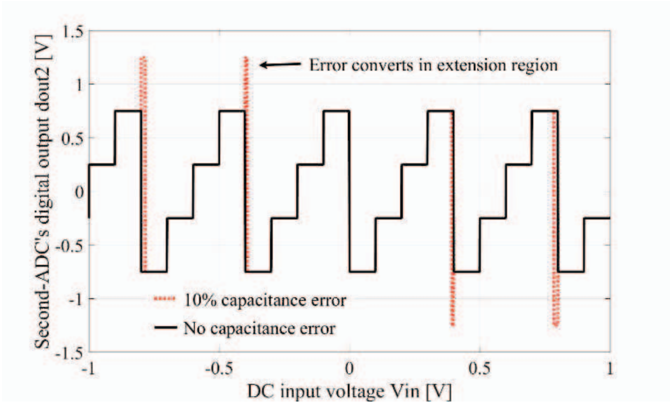


Fig. 4 Digital output of the second-stage ADC ($\mathrm{dout2}(z)$)

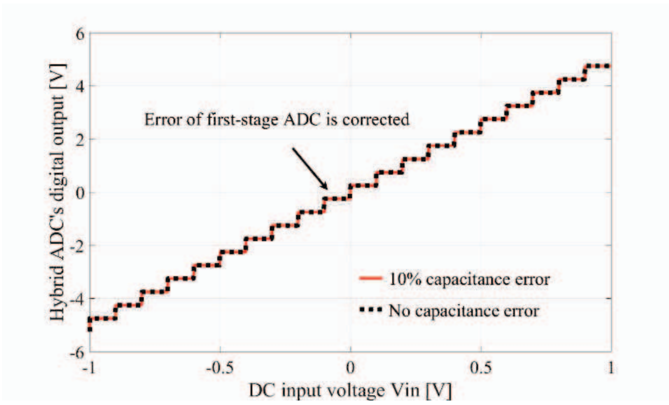


Fig. 5 Digital output of the hybrid ADC

Examination of Incremental ADC with SAR ADC to Reduce Conversion Time with High Accuracy

Yuichiro Unno
Department of Elecrical Engineering
Tokyo University of Science
Chiba, Japan
unno@vlsi.ee.noda.tus.ac.jp

Tatsuji Matsuura, Ryo Kishida, Akira Hyogo
Department of Elecrical Engineering
Tokyo University of Science
Chiba, Japan

***Abstract*—We propose a hybrid ADC that decomposes the upper bits with an Incremental Analog to Digital Converter (I-ADC), and the lower bits with a Successive Approximation Register ADC (SAR ADC). The proposed hybrid ADC is investigated by circuit simulation at the transistor level. The ADC accurately converts the date and reduces conversion time.**

***Keywords*— hybrid ADC, incremental ADC, SAR ADC, switched-capacitor circuit.**

I. Introduction

In recent Internet of Things (IoT) eras, higher precision Analog to Digital Converters (ADCs) are required. Incremental ADCs (I-ADCs) are useful for higher precision conversion [1]. However, the I-ADC is lower conversion speed since 2^{n} clock counts are required for n-bit resolution. The conversion time becomes longer with higher resolution. In order to reduce the conversion time, I-ADC and Successive Approximation Register ADC (SAR ADC) are used in high- and low-order bits respectively. Of n-bit resolution, if n_1-bit is converted by I-ADC and n_2-bit is converted by SAR ADC, the conversion time can be reduced to $2^{n_1}+n_2$ clock counts. In this paper, we design the hybrid ADC at the transistor level and confirm its operation by circuit simulation.

II. Design of Incremental and SAR ADC

A. Incremental and SAR ADC operation

Figs. 1 and 2 show a circuit diagram of the designed ADC and clock timing respectively. In this architecture, C_f of I-ADC is composed of binary weighted capacitor array for lower bit SAR determination. The operation of the I-ADC decomposing the upper bits (n_1 bits) is as follows.

1)Turn on $\emptyset_{\mathrm{RS}}$ to reset charge in the capacitor C_f of integrator.
2)Sampling: By turning on $\emptyset_1$, C_s and C_{DAC} store charge from V_{in} and V_{ref} respectively.
3) Integration and DAC subtraction: Turn off $\emptyset_1$ and turn on $\emptyset_2$. C_f stores the sum of charge in C_s and $2C_{DAC}$ because left plate voltages in those capacitors become V_{cm} and $-V_{ref}$.
4) Steps 2) and 3) are repeated by 2^{n_1}. When the integrated voltage exceeds V_{cm}, the output of the comparator $\emptyset_D$ becomes high, and the input to C_{DAC} during sampling changes to $-V_{ref}$. At the same time, the number of counter is incremented. The counted number becomes the digital output code of the I-ADC.

Residual signal remains in C_f after 2^{n_1} by I-ADC operation. The operation of the SAR ADC to decompose the lower bit (n_2 $bits$) is the following procedure.

1) Most significant bit of SAR ADC from residual of I-ADC is determined by 1/2 C_f and switch S_1 at the first clock of SAR ADC. SAR logic controls all switches (S_1-S_3) by comparison results of the comparator.
2) 2nd bit is determined by 1/4 C_f and S_2 at the second clock.
3) Determine all bits to the least significant bit by n_2.

B. Calculation method of digital output

The digital output code D_{out} is calculated. The integrator output of the I-ADC after 2^{n_1} clock ($V_{int\_out}[2^{n_1}]$) is

$$V_{int\_out}[2^{n_1}] = \frac{1}{2}(2^{n_1}V_{in} + \sum_{k=1}^{2^{n_1}} d_k V_{ref}) \qquad (1)$$

Here, d_k ($k = 1,2,3 \ldots 2^{n_1}$) is the determination result of the comparator of I-ADC and it is 1 or -1.

The voltage value of the top plate of capacitor array $\mathrm{C_f}$ after $2^{n_1}+n_2$ clock ($V_{out}[2^{n_1}+n_2]$) is calculated by (2).

$$V_{out}[2^{n_1}+n_2] = V_{int\_out}[2^{n_1}] + \sum_{i=1}^{n_2} \frac{d_i}{2^i} V_{ref} \qquad (2)$$

Here, d_i ($i = 1,2,3 \ldots n_2$) is the judgment result 1 or -1 of the SAR ADC's comparator after i clock. From (1) and (2), the digital output value D_{out} for V_{in} is calculated by (3).

$$D_{out} = (\sum_{k=1}^{2^{n_1}} d_k - 1)2^{n_2} + \sum_{i=1}^{n_2} d_i\, 2^{n_2-i} \qquad (3)$$

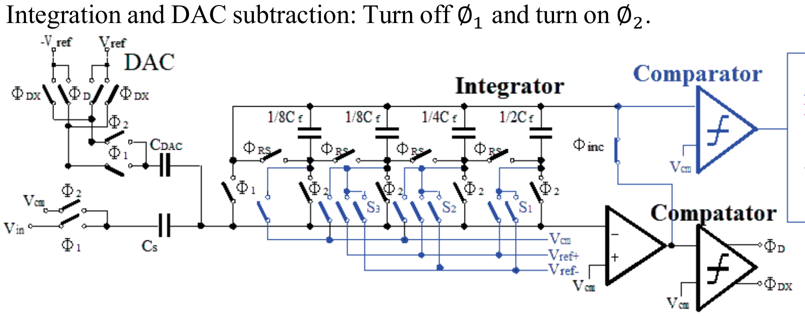


Fig. 1. Configuration of I- ADC＋SAR ADC

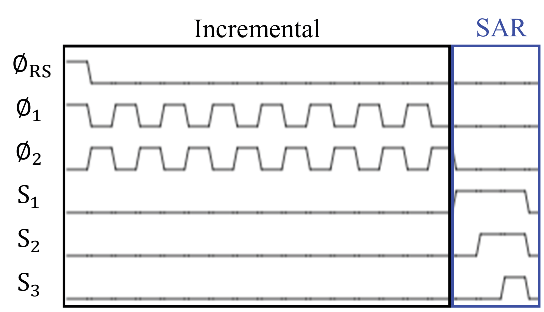


Fig. 2. Clock chart

III. SIMULATION

A. Simulation conditions

The ADC as shown in Fig. 1 is simulated using the Cadence Spectre circuit simulator. The number of both high and low order bits are 3-bits. The input voltage ($V_{in} = V_{in+} - V_{in-}$) is -1.4 to 1.4 V by 193 steps. The integrator output of I-ADC and the digital output value are confirmed. Table 1 shows simulation conditions and device parameters. Switches and comparator are composed in transistor level with 0.15 µm process. Operational amplifier is used to ideal transconductance amplifier with voltage controlled current source.

B. Simulation result

Fig. 3 shows the voltage waveform of the top plate of C_f when $V_{in} = 0.8$ V. d_k and d_i determine whether V_{out} is larger or less than 0. According to (3), $D_{out} = 50$.

Fig. 4 shows the signal remaining in C_f after the end of incremental operation. By first SAR comparison, V_{int_out} is +/-(1/2)V_{ref} shift and V_{in_sar} is generated. After that, normal SAR operation determines the lower bits.

Fig. 5 shows D_{out} for changing V_{in}. D_{out}(inc)and D_{out}(SAR) are digital output code of I- and SAR ADC respectively. The final output code D_{out} is generated by these output codes are combined. As a result of comparing the simulation value and the theoretical value, it was confirmed that input voltage Vin is correctly converted.

IV. CONCLUSION

In order to reduce the conversion speed of high precision ADC, we propose the hybrid ADC with incremental and SAR ADC. We designed it at the transistor level and confirmed that AD conversion is possible correctly by circuit simulation.

ACKNOWLEDGIMENT

This work is supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Cadence Design Systems, Inc and JSPS Grant-in-Aid for Scientific Research (C) 17K06400.

TABLE I. Simulation parameters

Components	Symbol	Value
Power supply voltage	V_{DD}	1.8 V
Common voltage	V_{cm}	0.9 V
Clock frequency	f_{CLK}	10 MHz
Reference voltage (+)	V_{ref+}	0.7 + 0.9 V
Reference voltage (−)	V_{ref-}	−0.7 + 0.9 V
Sampling capacitance	C_s	1.0 pF
DAC capacitance	C_{DAC}	0.5 pF
Feedback capacitance	C_f	2.0 pF

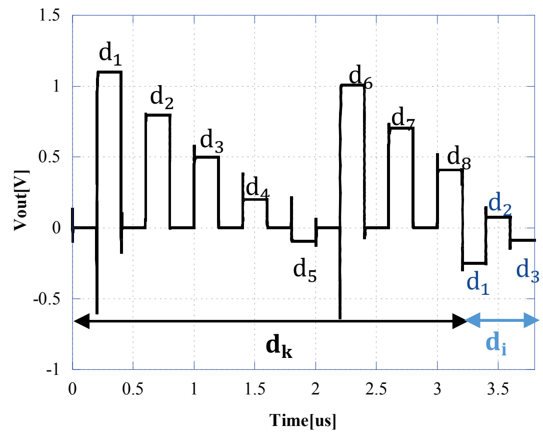


Fig.3 The voltage of the top plate of C_f at $V_{in} = 0.8$ V

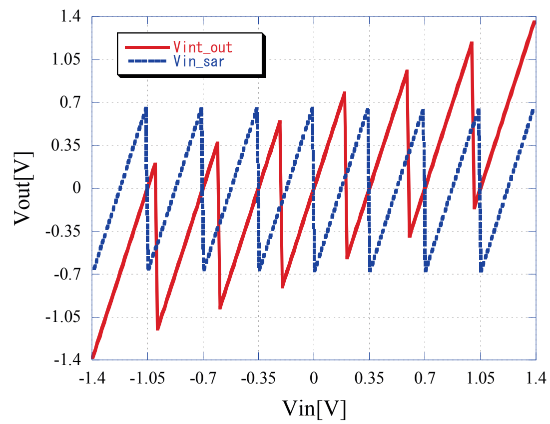


Fig. 4. Signal remaining in C_f after the end of incremental

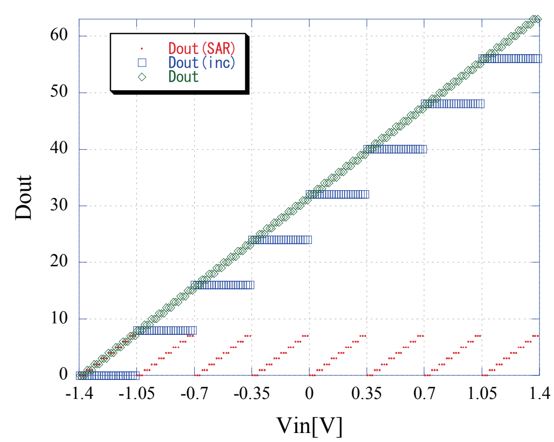


Fig.5. Simulation result of D_{out} for V_{in}

REFERENCES

[1] J. Márkus, J. Silva, and G.C.Tames. "Theory and Applications of Incremental Converters," *IEEE Translations on Circuits and Systems I: Regular Papers*, vol. 51, no. 4, pp. 678-690, April 2004.

[2] J. L. McCreary, P. R. Gray, "All-MOS charge redistribution analog-to-digital conversion techniues—Part I, " *IEEE J. Solid-State Circuits*, vol. SC-10, no. 6, pp. 371-379, Dec. 1975.

2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)

A plus type CC-based current-mode universal biquadratic circuit

Takao Tsukutani
Professor Emeritus, National Institute of Technology, Matsue College
Matsue, Japan
t.tsukutani@sea.chukai.ne.jp

Noboru Yabuki
Department of Integrated Science and Technology
National Institute of Technology, Tsuyama College
Tsuyama, Japan
yabuki@tsuyama-ct.ac.jp

Kazuharu Hashitsume
Advanced Studies on Teaching Profession, Graduate Schools of Education
Shimane University
Matsue, Japan
hashitsume@edu.shimane-u.ac.jp

***Abstract*— This paper introduces a current-mode universal biquadratic circuit using only plus type current conveyors (CCs) (i.e. differential voltage current conveyors (DVCCs) and second generation current conveyors (CCIIs)). The biquadratic circuit enables low-pass (LP), band-pass (BP), high-pass (HP), band-stop (BS) and all-pass (AP) responses by the selection and/or addition of the input and output currents without any component matching constraints. Moreover the circuit parameters ω_0 and Q can be set orthogonally adjusting the circuit components. A design example is given together with simulation responses by PSPICE.**

***Keywords*— analog circuit, biquadratic responses, current conveyors, CMOS technology**

I. INTRODUCTION

High performance active circuits have received much attention. The circuit designs using active devices such as the CCs, operational trans-conductance amplifiers (OTAs), operational trans-resistance amplifiers (OTRAs), etc. have been reported in the literature [1]-[3]. A CC is a very useful active device, and CC-based circuit is suitable for high frequency operation. There are some kinds of CCs (e.g. CCII, third generation current conveyor (CCIII), DVCC, etc.). The plus type CCs are composed of simpler circuit configuration than the minus type ones. Hence they have wide frequency operation and low power performance compared with the minus type CCs.

The biquadratic circuit is a convenient second-order function block. Several biquadratic circuits using the CCs have been discussed previously [4],[5]. However the plus type CC-based biquadratic circuit has not yet been studied sufficiently.

This paper introduces a current-mode universal biquadratic circuit using only the plus type CCs (i.e. DVCCs and CCIIs) and grounded passive components. First we show a basic current-mode biquadratic circuit, and then typical current-mode circuit is consisted of using the basic current-mode one. The circuit enables low-pass (LP), band-pass (BP), high-pass (HP), band-stop (BS) and all-pass (AP) responses by the selection and/or addition of the input and output currents with no component matching constraints. Moreover the circuit parameters ω_0 and Q can be set orthogonally adjusting the circuit components.

A design example is given with PSPICE simulation, and the circuit workability is confirmed.

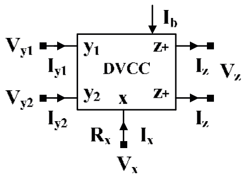
II. DVCC AND CCII

The symbols of the plus type DVCC and CCII are given in Fig.1. The plus type DVCC and CCII are characterized by the following terminal equations:

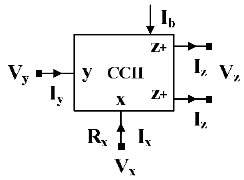
$$V_x = V_{y1} - V_{y2} - I_x R_x, \quad I_z = I_x \tag{1}$$

$$V_x = V_y - I_x R_x, \quad I_z = I_x \tag{2}$$

where R_x denotes the parasitic resistance at the x-terminal.
Figure 2 shows the DVCC [5] and CCII [1] with MOS transistors. They can be composed of less transistors than the minus type ones.

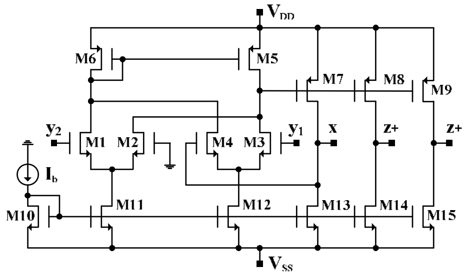


(a) DVCC

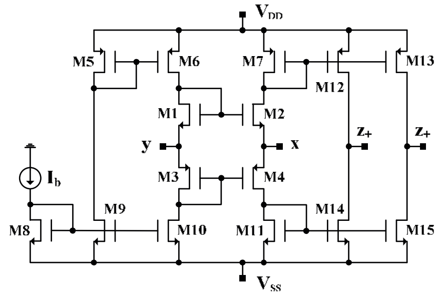


(b) CCII

Fig.1. Symbols of CCs.



(a) DVCC



(b) CCII

Fig.2. Plus type CCs with MOS transistors.

III. CC-BASED CURRENT-MODE BIQUADRATIC CIRCUIT

Figure 3 shows a basic current-mode biquadratic circuit configuration. In this circuit, all the x-terminals of the CCs are connected to grounded resistors for minimizing the parasitic effects.
The current outputs $I_{o1}(s)$ and $I_{o2}(s)$ are given by:

$$I_{o1}(s) = -\frac{(s^2 + 1/C_1C_2R_2R_3)I_{in1}(s) - (1/C_1R_1)sI_{in2}(s)}{s^2 + (1/C_1R_1)s + 1/C_1C_2R_2R_3} \tag{3}$$

$$I_{o2}(s) = -\frac{-(1/C_1C_2R_2R_3)I_{in1}(s) + (s^2 + (1/C_1R_1)s)I_{in2}(s)}{s^2 + (1/C_1R_1)s + 1/C_1C_2R_2R_3} \tag{4}$$

This circuit enables the LP, BP and BS responses by selection of the input and output currents as follows:

$$T_{LP}(s) = \frac{I_{o2}(s)}{I_{in1}(s)} = \frac{1/C_1C_2R_2R_3}{s^2 + (1/C_1R_1)s + 1/C_1C_2R_2R_3} \tag{5}$$

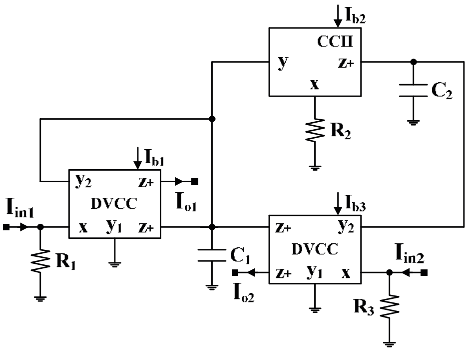


Fig.3. Basic current-mode biquadratic circuit.

$$T_{BP}(s)=\frac{I_{o1}(s)}{I_{in2}(s)}=\frac{(1/C_1R_1)s}{s^2+(1/C_1R_1)s+1/C_1C_2R_2R_3} \tag{6}$$

$$T_{BS}(s)=\frac{I_{o1}(s)}{I_{in1}(s)}=-\frac{s^2+1/C_1C_2R_2R_3}{s^2+(1/C_1R_1)s+1/C_1C_2R_2R_3} \tag{7}$$

Moreover the HP response can be achieved by the current addition of $I_{HP}(s)=I_{o1}(s)+I_{o2}(s)$, and the AP response is performed selecting the input current $I_{in}(s)=I_{in1}(s)=I_{in2}(s)$. The circuit transfer functions are given as:

$$T_{HP}(s)=\frac{I_{HP}(s)}{I_{in1}(s)}=-\frac{s^2}{s^2+(1/C_1R_1)s+1/C_1C_2R_2R_3} \tag{8}$$

$$T_{AP}(s)=\frac{I_{o1}(s)}{I_{in}(s)}=-\frac{s^2-(1/C_1R_1)s+1/C_1C_2R_2R_3}{s^2+(1/C_1R_1)s+1/C_1C_2R_2R_3} \tag{9}$$

Thus five standard circuit transfer functions can be obtained by choosing the circuit currents.

The typical current-mode biquadratic circuit is consisted of using the basic current-mode one shown in Fig.4.

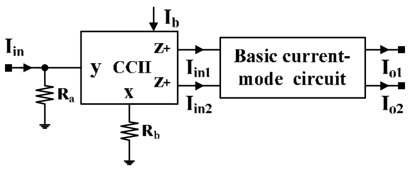


Fig.4. Typical current-mode biquadratic circuit.

The circuit parameters ω_0, Q and H are represented as below:

$$\omega_0=\sqrt{\frac{1}{C_1C_2R_2R_3}},\quad Q=R_1\sqrt{\frac{C_1}{C_2R_2R_3}},\quad H=\frac{R_a}{R_b} \tag{10}$$

The circuit parameter ω_0 and Q can be set orthogonally according to the circuit components, and meanwhile the parameter H is set independently.

IV. DESIGN EXAMPLE AND SIMULATION RESPONSES

We verified the circuit operation using PSPICE simulation program. As an example, we tried to achieve a current-mode circuit with f_0 ($=\omega_0/2\pi$)=1MHz, Q=1.0 and H=1.0. To achieve the specification above, we set the circuit resistors listed in table I.

TABLE I. CIRCUIT COMPONENTS

x	LP	BP	HP	BS	AP
R_1 (kΩ)	10.5	11.5	12	11.2	11
R_2 (kΩ)	10.5	11.5	12	11.2	11
R_3 (kΩ)	10.5	11.5	12	11.2	11
R_a (kΩ)	15	14.8	14.5	15	15.2
R_b (kΩ)	10	10	10	10	10

Figure 5 shows the simulation responses. Figure 5 (a) shows the LP, BP, HP and BS responses, and the AP response is shown in Fig.5 (b). In the figures, the marks signify the simulation responses, and the continuous lines show the theoretical responses. This can be viewed as an excellent result over a wide frequency range. Here we set that the capacitors, input current, bias currents and DC supply voltages were $C_1=C_2=12$pF, $I_{in}=10\mu$A, $I_{b1}=I_{b2}=I_{b3}=I_b=10\mu$A and $V_{DD}=-V_{SS}=1.2$V. The power dissipation was 0.618mW.

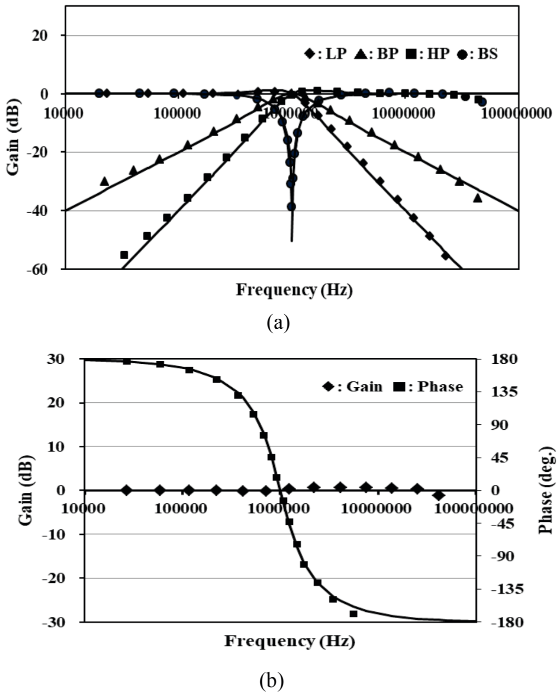


Fig.5. Simulation responses.

In this simulation, the MOS sizes of the DVCC were set as 20μm/0.5μm (M1 to M4), 30μm/2μm (M5 to M9), 10μm/2μm (M10 to M15), and the aspect ratios were 20μm/1μm (M1 to M4) and 10μm/1μm (others) in the CCII. And we used the model parameters from MOSIS 0.5μm.

V. CONCLUSION

A current-mode universal biquadratic circuit employing plus type CCs and grounded passive components has been proposed. We have demonstrated that the circuit can achieve five circuit responses by selecting and/or adding the input and output currents without the component matching constraints. The achievement example has been given together with simulation results by PSPICE. The simulation responses were appropriate enough over a wide frequency range.

The non-idealities (i.e. voltage and current tracking errors) of the CC affect the circuit performances. The solution for this will be discussed in the future.

REFERENCES

[1] A. Fabre, et al., "High frequency applications based on a new current controlled conveyor," IEEE Trans. Cir. Syst., vol.43, pp.82-91, 1996.
[2] M.T. Abuelma'atti, et al., "A novel mixed-mode OTA-C universal filter," Int. J. Electron., vol.92, pp.375-383, 2005.
[3] C. Cakir, et al., "Novel allpass filter configuration employing single OTRA,"IEEE Trans. Cir. Syst., vol.52, pp.122-125, 2005.
[4] M.A. Ibrahim, et al., "A 22.5MHz current-mode KHN-biquad using differential voltage current conveyor and grounded passive elements", Int. J. Electron. Comm., vol.59, pp.311-318, 2005.
[5] T. Tsukutani, et al., "Novel current-mode biquadratic circuit using only plus type DO-DVCCs and grounded passive components," Int. J. Electron., vol.94, pp.1137-1146, 2007.

Multi-Output Octagonal MOSFET for the Common Device of Both Sensor and Circuit Design

Tomochika Harada
Graduate School of Science and Engineering, Yamagata University
Yonezawa, Yamagata, 992-8510, Japan
E-mail: tharada@yz.yamagata-u.ac.jp

***Abstract*—In this paper, I propose the multi-output Octagonal MOSFET for the common device of both sensor and circuit design. In previous works, one sensor device designs using a resistor, capacitor, etc. So, it is very difficult to design sensors and circuits together using the common device. However, multi-sensing operation and normal circuit operation are both available for proposed octagonal MOSFET. Moreover, reconnecting the output terminals of octagonal MOSFET, the size ratio, W/L, of it can change and I realize it can easily design both sensors and circuits using only one kind of common device.**

***Index Terms*—octagonal MOSFET, multi-sensor, multi-output MOSFET, circuits and sensors design**

I. Introduction

IC and LSI have been used in many electronic devices, such as a car, mobile phone, etc. In IC and LSI, some sensors and circuits are available [1]- [3]. In previous works, the sensor device designs using a resistor, capacitor, etc. So, it is very difficult to design sensors and circuits together using the common device. However, multi-sensing operation and normal circuit operation are both available for proposed octagonal MOSFET [4] [5]. By making the MOSFET sensor structure an octagonal shape, the electrode arrangement becomes a line symmetrical and point symmetrical, and an effect of suppressing deviation of terminals at the time of manufacturing can be expected. It has been shown that it can be used as a stress detection element or a temperature detection element as a previous study [4] and it can be expected to be utilized as a multi-sensor [5].

In this paper, multi-output Octagonal MOSFET for the common device of both sensor and circuit design is proposed. Especially, I evaluate the changing the inversion voltage with the size ratio W/L of MOSFET using the inverter circuit.

II. Octagonal MOSFET

Fig.1 shows the device layout and microphotograph of the octagonal MOSFET. This device is designed and fabricated by the $2.0\mu m$ 1-poly 2-metal CMOS process. In general, MOSFET has only Gate, Drain, Source, and Bulk electrode. But, octagonal MOSFET has additional 6 output electrodes, which are located along the side of a regular octagon radially. These electrodes, which are O1 to O8 in fig.1, can be defined source, drain, one pair of hall effect detection outputs, which is perpendicular to drain current direction, and the other pair of temperature sensing output, which is parallel to drain current direction, arbitrarily.

Multi-output octagonal MOSFET has three types of operating modes.

1) Operation as the normal MOSFET.
2) Differential operation between two opposite sides for the detection of the magnetic field and the stress.
3) Differential operation only a single side for the detection of temperature, current, etc.

Using these modes, both sensors and circuits can be designed using only this device.

III. Measurement results and conclusion

For evaluating, the simple inverter, which is composed of only octagonal MOSFETs, is designed and fabricated. In fig.2, Vdd and GND are connected to p315 and n45 terminals, respectively. However, this device has 8 terminals. Thus, if these terminals can be reconnected to Vdd or GND, the size ratio, W/L, of MOSFET is easy to change without redesigning the device size. The inversion threshold voltage is shifted because the W/L of nMOS and pMOS are changed. Other methods are connecting two or four terminals to GND. Fig.3 and 4 show the result of the input-output characteristics of the octagonal MOSFET inverter and the result around the threshold voltage of inversion in the inverter, respectively. From these results, the maximum changing size ratio can be 3.47 times as large as the normal inverter.

This device operates as the multi-sensor in previous research [5]. For example, fig.5 shows the electrical measurement method. V_{TEMP} and V_{HALL} are temperature and magnetic field output voltage, respectively. Fig.6 shows the results of the multi-sensing operation. T is the temperature of fig.5 measured using thermocouple gauge, and the other data are from voltages of hall effect and temperature measurement terminals. From the result of temperature, ΔV_{TEMP} and T are almost the same behavior. Thus, both hall effect and temperature detection are available in the proposed device.

From these results, both sensors and circuits can be designed by only 1 kind of the proposed common device, multi-output octagonal MOSFET.

Acknowledgment

This IC chip was fabricated at the Semiconductor Center of Kitakyushu Science and Research Park. Besides this work was partly supported by the Center of Innovation Program from the Japan Science and Technology Agency (JST).

References

[1] X. Zhao, J. Cao, Y. Song, D. Wen, Q. Lin, L. Tian, "Fabrication and Characteristics of the 2D Magnetic Sensor based on the MOSFET Hall Device", Key Engineering Materials, Vols.609-610, pp.1066-1071, 2014

[2] Baltes, Brand, Fedder, Hierold, Korvink, and Tabata, "CMOS-MEMS", WILEY-VCH Verlag GmbH & Co. KGaA, pp.279-299, 2005

[3] M. Doelle, C. Peters, P. Ruther, O. Paul, "Piezo-FET Stress-Sensor Arrays for Wire-Bonding Characterization," JOURNAL OF MICRO-ELECTROMECHANICAL SYSTEMS, VOL. 15, NO.1, pp.120-130, Feb.2006

[4] T.Harada, K.Kaiwa, and Y.Yamazaki, "No Reattaching and 8 Directions Detectable Octagonal MOSFET Stress Sensor",2015 International Conference on Solid State Devices and Materials(SSDM), F-6-4,pp.848-849, Sep.30, 2015

[5] T.Harada, "Parallel Sensing Operation using Octagonal MOSFET", 2017 International Symposium on Electronics and Smart Devices (ISESD), pp.278–281, Oct.2017

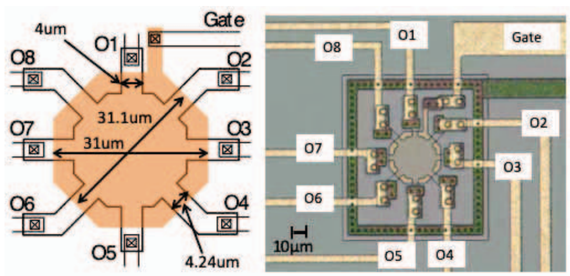


Fig. 1. The layout and microphotograph of the octagonal MOSFET

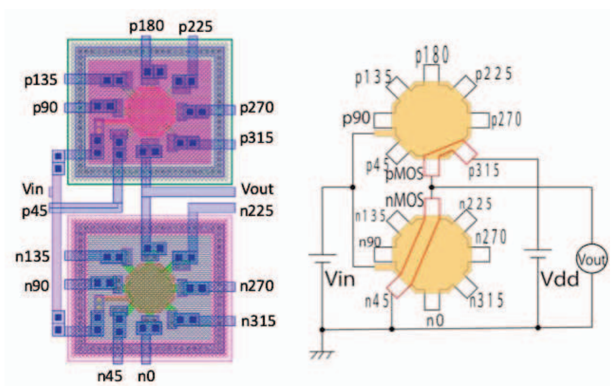


Fig. 2. The layout and schematics of the octagonal MOSFET inverter

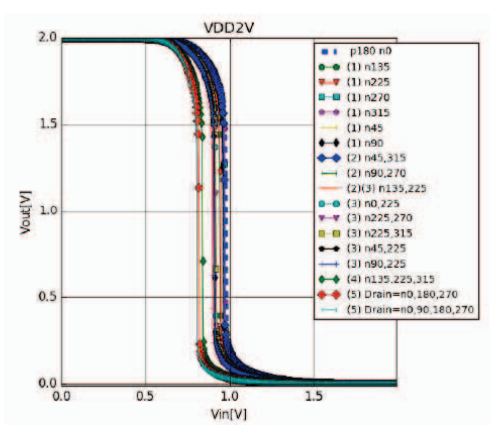


Fig. 3. The input-output characteristics of the octagonal MOSFET inverter

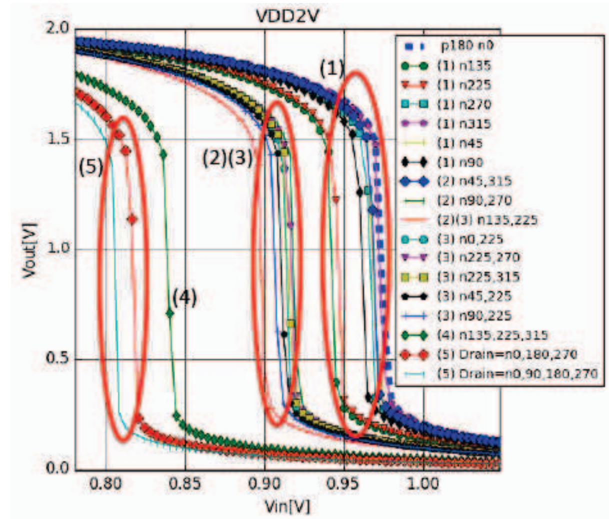


Fig. 4. The results of the input-output characteristics expanded around the threshold voltage of the inversion

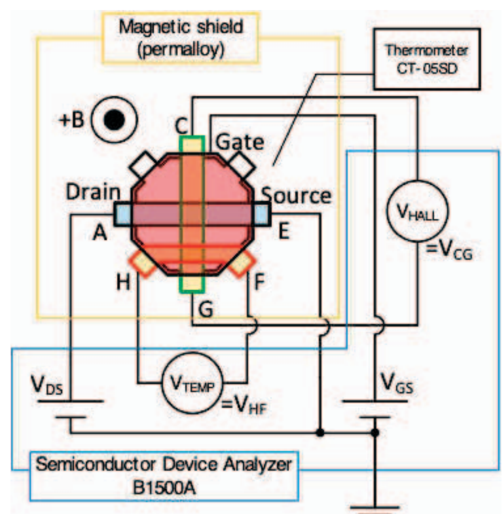


Fig. 5. Octagonal MOSFET for the sensor operation [5]

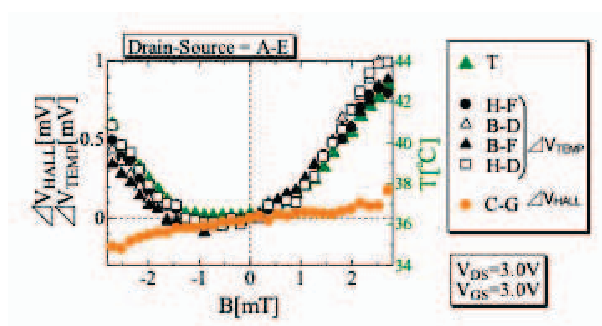


Fig. 6. The measurement results of multi-sensing operation [5]

Design of 10 GHz CMOS Optoelectronic Receiver Analog Front-End in Low-Cost 0.18 μm CMOS Technology

Xiangyu Chen
Graduate School of Engineering
Gifu University, Japan
1-1 Yanagido, Gifu-shi, 501-1193
Email: xiangyuchenCN@hotmail.com

Yasuhuiro Takahashi
Dept. of Electr., Electro. and Comput. Eng.
Faculty of Engineering, Gifu University, Japan
1-1 Yanagido, Gifu-shi, 501-1193
Email: yasut@gifu-u.ac.jp

***Abstract*—This paper presents an analog front-end circuit for an optical receiver, which consists of a trans-impedance amplifier (TIA), a limiting amplifier (LA) and an output buffer fabricated in a low-cost 0.18 μm CMOS technology. The introduced TIA uses a floating active inductor (FAI) based on gyrator-C structure, which is why it can increase the bandwidth while occupying a smaller chip area. The proposed transimpedance amplifier achieves a transimpedance gain of 41 dBΩ and -3 dB frequency of 10 GHz with 0.1 pF total input capacitance. In addition, the TIA with the post limiting amplifier increases the gain to over 70 dBΩ and the bandwidth is still above 10 GHz.**

I. INTRODUCTION

In recent years, the amount of information exchanging has been increasing, and the optical fiber communication systems and networks have been developing rapidly. Optical fiber communication with light wave as carrier and optical fiber as the transmission medium has become one of the important ways for people to exchange data information. The analog front-end, the most critical module of optical receiver, consists of a trans-impedance amplifier (TIA) and a limiting amplifier (LA) and aims to amplify the weak current received from the photo-diode to a recognizable voltage with digital swing.

In this design, the proposed TIA can provide -3 dB bandwidth greater than 10 GHz and transimpedance gain of 41 dBΩ. Due to the use of FAI, the occupied area of the TIA is greatly reduced, almost 18.8% of the conventional TIA area in Reference (1), and the floating active inductor does not increase power dissipation. In order to make the gain more desirable, the TIA must followed by additional amplifying stages that boost the signal swing to logical levels. In our proposed front-end circuit, a modified Cherry-Hooper amplifier with resistive loads is used, which cascades an output buffer.

II. TRANSIMPEDANCE AMPLIFIER

The conventional TIA is shown in Fig. 1, where L_1 and L_2 are spiral inductors. In this design, we change L_2 from a spiral inductor to a floating active inductor, which allows the TIA to increase the bandwidth while reducing the chip area occupied by the TIA. In addition, unlike the conventional TIA in reference (1), we assume that L_1 is constructed by bonding wire in this work. It is well known that in general process, the inductance value of 1 mm bonding wire is between 2 nH and 2.5 nH. In order to reduce the negative impact of the parasitic inductance of bonding wire and reduce the area, we use the bonding wire as the inductor L_1, which is not on-chip. So we cannot find L_1 in the layout diagram as shown in Fig. 2 (a). Fig. 2 (b) shows the post-layout frequency response of the proposed TIA, and we can see from the simulation results that no matter what value L_1 takes between 2 nH and 3 nH, the performance of this design is in an acceptable range. Fig. 2 shows that the occupied area of only TIA core is 180 μm $\times$ 118 μm.

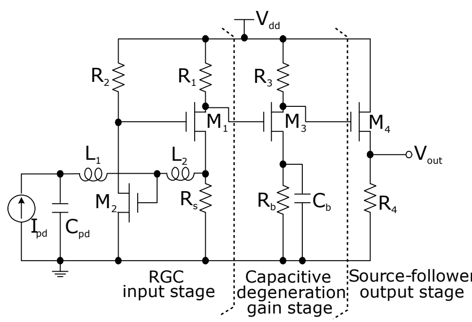
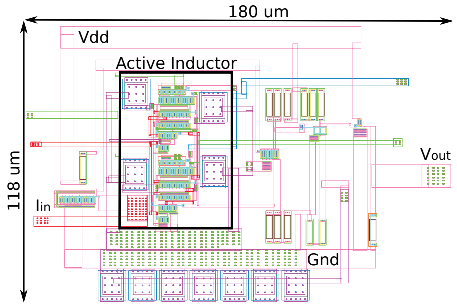


Figure 1. Schematic of the conventional TIA.

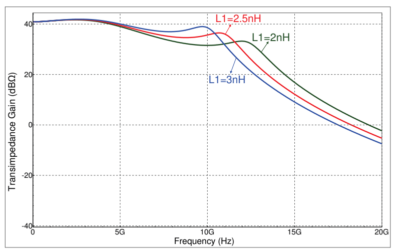
An active inductor is constructed with a small number of transistors for the implementation of the the gyrator. The gyrator consists of two back-to-back connected transconductors [2]. A gyrator-C network is said to be lossless when both the input and output impedances of the transconductor of the system are infinite, and the transconductances of the transconductors are constant.

III. PROPOSED ANALOG FRONT-END AND SIMULATION RESULTS

Figure 3 depicts a schematic diagram of the proposed analog front-end network. In this design, an unbalanced pseudo-differential TIA with one photodetector is used, as shown in



(a)



(b)

Figure 2. (a) Layout of the proposed TIA. (b) Post-layout simulated frequency response of the proposed TIA.

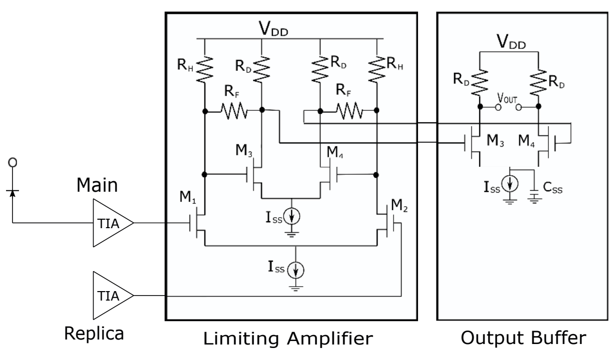


Figure 3. Schematic of the proposed analog front-end.

Fig. 3. This part consists of a single-ended main TIA and a matching replica TIA (a.k.a dummy TIA). The replica TIA simply produces a DC voltage that tracks the dark level of the voltage over process, voltage and temperature. To alleviate the gain-headroom trade-off, we use the modified Cherry-Hooper amplifier, where the resistors provide part or all of the bias current of the input differential pair. In this LA, R_H must be much greater than the input resistance of the second stage to avoid degarding the gain. In order to achieve the matching with the subsequent circuits and achieve the maximum power output, the LA must use the output buffer unit to achieve output impedance matching and improve the driving capability of the circuit.

The simulated frequency response of the analog front-end is shown in Fig. 4. The -3 dB frequency of this circuit is 10 GHz and the conversion gain is about 70 dBΩ.

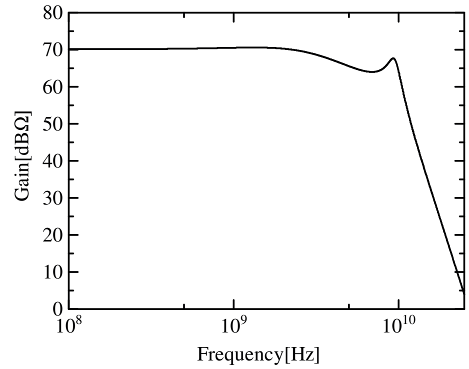


Figure 4. Frequency response of the proposed analog front-end.

IV. Conclusion

This paper presented the design of a broadband optical receiver analog front-end circuit in a low-cost 0.18 μm CMOS process. It contains an unbalanced psedudo-differential TIA with floating active inductor, a LA of modified Cherry-Hooper amplifier, and a differential output buffer stage. The proposed AFE circuit achieves -3 dB frequency of 10 GHz and gain of 70 dBΩ. Instead of using spiral inductors, active inductors are utilized in this design. Thus, the chip area can be greatly reduced. The proposed architecture is suitable for low-cost application.

Acknowledgment

This work was supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Cadence Corporation and Synopsys, Inc. The VLSI chip in this study has been fabricated in the chip fabrication program of VDEC, the University of Tokyo in collaboration with Rohm Corporation and Toppan Printing Corporation.

We would also like to acknowledge the financial support by the China Scholarship Council (CSC) through the award of a Ph.D. scholarship (No. 201808050092) to Xiangyu Chen, which is under the State Scholarship Fund.

References

[1] Z. Lu, K. S. Yeo, and J. G. Ma, "Broad-band design techniques for transimpedance amplifiers," *IEEE Trans. Circuit and Systems-I: Regular Papers*, vol. 54, no. 3, pp. 590–600 (2007–3).

[2] F. Yuan, *CMOS Active Inductors and Transformers*, Springer, New York, 2008.

[3] F. Mahmoudi, and C. A. Salama, "8 GHz 1V, CMOS quadrature down-converter for wireless applications," *Analog Integr. Circ. Sig. Process.*, vol. 48, pp. 185–197 (2006).

[4] Z. Lu, K. S. Yeo *et al.*, "Design of a CMOS broadband transimpedance amplifier with active feedback," *IEEE Trans. Very Large Scale Integration (VLSI) Systems.*, vol. 18, no. 3, pp. 461–472 (2010–3).

[5] B. Razavi, *Design of integrated circuits for optical communications*, WILEY, USA, 2012.

[6] E. Sackinger, *Analysis and Design of Transimpedance Amplifiers for Optical Receivers*, WILEY, USA, 2018.

[7] H. Kimura, et al., "A 28 Gb/s 560 mW multi-standard serdes with single-stage analog front-end and 14-Tap decision feedback equalizer in 28 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 3091–3103 (2014–12).

Comparison of Face Recognition Loss Functions

Hung-Yi Wu, Min-Hsiang Chang, Gee-Sern Jison Hsu
Artificial Vision Laboratory
National Taiwan University of Science and Technology
jison@mail.ntust.edu.tw

***Abstract*—Significant progresses have been made to face recognition algorithms in recent years. The progresses include the improvements of the solutions and the availability of more challenging databases. As the performance on previous benchmark databases, such as MPIE and LFW, saturates, more challenging databases are emerging and keep driving the advancements of face recognition solutions. The loss function considered in the solution usually plays the most critical role. We compare a few latest loss functions with the same feature embedding network by evaluating their performance on two recently-released databases, the IARPA Janus BenchmarkB (IJB-B) [1], and IARPA Janus BenchmarkC (IJB-C) [2], and highlight the directions for the next phase research.**

***Index Terms*—Face recognition, loss function, deep learning, convolutional neural network**

I. Introduction

Deep learning approaches have been advancing the face recognition technology in an unprecedented pace. A typical deep learning approach includes a deep convolutional neural network (CNN) and a database for the training and testing of the CNN. The loss function is a core part of the CNN that can determine how well the targeted task, e.g., classification or regression, is solved. The softmax function that computes the cross-entropy loss is commonly used when solving a classification problem. The improvements or modifications to the softmax loss is an active research with the goal of finding more effective loss functions.

We have selected several latest loss functions, including the Triplet Loss [3], the Center Loss [4], the Marginal Loss [5] and the Angular Softmax Loss [6], and compared their performance on two challenging databases, the IARPA Janus BenchmarkB (IJB-B) [1] and IARPA Janus BenchmarkC (IJB-C) [2]. The comparison is based on the same CNN architecture and only the loss function is replaced by each specific one. Note that although the performances of the selected loss functions are reported in their papers, many are studied on other databases instead of the IJB-B and IJB-C.

II. SELECTED LOSS FUNCTIONS

The loss functions selected in this study include the Triplet Loss [3], the Center Loss [4], the Marginal Loss [5] and the Angular Softmax Loss [6]. As these loss functions consider the Softmax Loss as a core reference, we introduce the Softmax Loss first and the others follow. The Softmax Loss function can be written as follows:

$$L_1 = -\frac{1}{N_b}\sum_{i=1}^{N_b}\log\frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^{n} e^{W_j^T x_i + b_j}} \tag{1}$$

where $x_i \in \mathbb{R}^d$ denotes the deep feature of the i-th sample, belonging to the y_i-th class. $W_j \in \mathbb{R}^d$ denotes the j-th column of the weight $W \in \mathbb{R}^{d\times n}$ and $b_j \in \mathbb{R}^n$ is the bias term. The batch size and the class number are N_b and n, respectively. The softmax loss is widely used in deep face recognition [7]. However, the softmax loss function does not explicitly optimize the feature embedding to enforce higher similarity for intra-class samples and diversity for inter-class samples, which motivates the developments of other loss functions.

A. Triplet Loss

The goal of the triplet loss is to ensure that the deep feature vector $f_i^a = f(x_i^a)$ of an anchor image x_i^a is close to the feature $f_i^p = f(x_i^p)$ of a positive data x_i^p, which has the same identity as of x_i^a, and is far away from $f_i^n = f(x_i^n)$ of a negative data x_i^n that has a different identity than x_i^a. The triplet loss is implemented in the FaceNet [3], a face recognition system made by Google, as: $L = \frac{1}{N}\sum_i^N max(\|f_i^a - f_i^p\|^2 + \alpha - \|f_i^a - f_i^n\|^2, 0)$, where α is the margin enforced between the positive and negative pairs.

B. Center Loss

Center loss [4] was proposed to improve the softmax loss for face verification. It learns a center for the features of each class and meanwhile tries to pull the deep features of the same class close to the corresponding center. Given the deep feature $[x_i]$ in a batch, the center loss can be computed as: $L_c = \frac{1}{2}\sum_{i=1}^{N}\|x_i - c_{y_i}\|_2^2$, where $c_{y_i} \in \mathbb{R}^d$ is the center of class y_i. During training, the center loss encourages instances of the same classes to be closer to a learnable class center. However, since the parametric centers are updated at each iteration based on a mini-batch instead of the whole dataset, which is very unstable, it has to be under the joint supervision of the softmax loss during training. Therefore the combined loss is formulated as: $L = L_S + \lambda L_c$, where L_S is the softmax loss, which is widely employed when solving typical classification problems, and λ is a hyper-parameter that balances the two losses.

C. Marginal Loss

The Marginal Loss function [5] was proposed to simultaneously maximize the inter-class distances and minimize the intra-class variations, both being desired features of a loss function. The Margin Loss function focuses on the marginal samples and is computed as follows,

$$L_M = \frac{1}{m^2 - m}\sum_{i,j,i\neq j}^{m}\left(\xi - y_{ij}\left(\theta - \left\|\frac{x_i}{\|x_i\|} - \frac{x_j}{\|x_j\|}\right\|_2^2\right)\right) \tag{2}$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

The term $y_{ij} \epsilon \{\pm 1\}$ indicates whether the faces x_i and x_j are from the same class or not, θ is the distance threshold to distinguish whether the faces are from the same person/class, and ξ is the error margin besides the classification hyperplane [5]. The final Marginal Loss function is defined as the joint supervision with regular Cross-Entropy (Softmax) Loss function and is given as follows: $L = L_S + \lambda L_M$, where L_S is the Softmax Loss. The hyper-parameter λ balances the two losses. The coupling with the cross-entropy loss provides separable features and prevents the loss from degrading to zeros [5].

D. Angular Softmax Loss

The Angular Softmax Loss was proposed to improve some issues with the softmax loss [6]. The issue of the bias $b_j = 0$ is handled by transforming the logit [8] as $W_j^T x_i = \|W_j\| \, \|x_i\| \cos\theta_j$, where θ_j is the angle between the weight W_j and the feature x_i. The issue with the individual weight $\|W_j\| = 1$ is handled by taking l_2 normalization to make the predictions only depend on the angle between the feature vector and the weight. To make it discriminative, the authors generalize it to the following Angular Softmax (A-Softmax) Loss, and call their solution "SphereFace".

$$L_3 = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{e^{\|x_i\| \cos(m\theta_{y_i})}}{e^{\|x_i\| \cos(m\theta_{y_i})} + \sum_{j=1, j\neq y_i}^{n} e^{\|x_i\| \cos\theta_j}}. \tag{3}$$

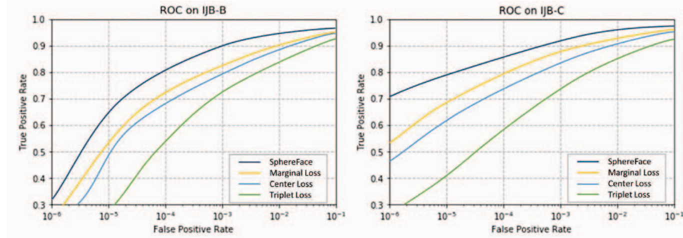
where θ_{y_i} has to be in the range of $[0, \frac{\pi}{m}]$.

III. Experiments

The MS-Celeb-1M dataset [9] can be the largest face recognition dataset so far and contains about 10M images for 100K subjects. However, this dataset has many mislabeled images, and we need to clean it beforehand. We have extracted 5.8M face images from 85K subjects by some semi-automatic processing for training.

The IJB-B [1] and IJB-C [2] can be two of the most mainstream and most challenging face databases. IJB-B has 76.8K face images from 1,845 different individuals. IJB-C has 148.8K face images from 3,531 different individuals. These two datasets contain face images of different conditions regardless of subject conditions (pose, expression, occlusion) or acquisition conditions (illumination, standoff, etc.). We tested the selected loss functions on both databases. For preprocessing the training and testing data, we employ the MTCNN [10] to detect the facial areas and landmarks. The 5 detected landmarks (two eyes, nose and two mouth corners) are used to crop the whole facial images, and each cropped face is normalized to 64×64. For a fair comparison, we employ the same CNN architecture, ResNet100 [11], as the embedding network. We follow the best settings reported in [3], [4], [5], [6], and compute the cosine distance of two features to obtain the similarity score.

The evaluation results are shown in Fig. 1 as the ROC (Receiver Operating Characteristic) curves. The SphereFace with A-Softmax loss outperforms all other loss functions on both databases. The second best is the Marginal Loss, then the



(a) ROC for IJB-B (b) ROC for IJB-C

Fig. 1. Performance for 1:1 verification on the IJB-B (left) and IJB-C (right) datasets.

Center Loo, and then the Triplet Loss. At false positive rate 10^{-4}, the SphereFace gives true positive rate 81% (86%), the Marginal Loss 73% (80%), the Center Loss 68% (74%) and the Triplet Loss just 54% (58%) on the IJB-B (IJB-C).

IV. Conclusion

Verified on two latest benchmark datasets, we have compared the performances of several state-of-the-art loss functions embedded in the same CNN networks. The difference between the best and the worse can be more than 20% in the 1:1 verification, revealing the strength of a properly designed loss function. It can be interesting to analyze the data failed to be handled by one but successfully handled by the other, extending our understandings toward the influences of the loss functions. The analysis may also help design better loss functions to deal with the data hard to handle by all selected loss functions.

Reference

[1] C. Whitelam, E. Taborsky, A. Blanton, B. Maze, J. Adams, T. Miller, N. Kalka, A. K. Jain, J. A. Duncan, K. Allen *et al.*, "Iarpa janus benchmark-b face dataset," in *CVPRW*, 2017.

[2] B. Maze, J. Adams, J. A. Duncan, N. Kalka, T. Miller, C. Otto, A. K. Jain, W. T. Niggel, J. Anderson, J. Cheney *et al.*, "Iarpa janus benchmark-c: Face dataset and protocol," in *ICB*. IEEE, 2018.

[3] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *CVPR*, 2015.

[4] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *ECCV*. Springer, 2016.

[5] J. Deng, Y. Zhou, and S. Zafeiriou, "Marginal loss for deep face recognition," in *CVPRW*, 2017.

[6] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "Sphereface: Deep hypersphere embedding for face recognition," in *CVPR*, 2017.

[7] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age," in *FG*. IEEE, 2018.

[8] G. Pereyra, G. Tucker, J. Chorowski, Ł. Kaiser, and G. Hinton, "Regularizing neural networks by penalizing confident output distributions," *arXiv preprint arXiv:1701.06548*, 2017.

[9] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "Ms-celeb-1m: A dataset and benchmark for large-scale face recognition," in *ECCV*. Springer, 2016.

[10] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.

Detection of Dangerous Objects By Pan-tilt Camera

Yurika Fujii
Tokushima Universithy
Tokushima, Japan
c501837044@tokushima-u.ac.jp

Minoru Fukumi
Tokushima Universithy
Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

Shin-ichi Ito
Tokushima Universithy
Tokushima, Japan
s.ito@tokushima-u.ac.jp

Momoyo Ito
Tokushima Universithy
Tokushima, Japan
momoito@is.tokushima-u.ac.jp

***Abstract*—Security cameras have increased in public facilities. The number of crimes has decreased by security cameras, but we will have too much data of cameras. In this paper, we have aim of security improvement. First, we search whether there are humans in images by OpenPose. We then obtain position of human's hands. Finally, we detect dangerous objects around the hands by image classification.**

***Index Terms*—VGG, OpenPose, security camera, pan-tilt camera, transfer learning**

I. Introduction

In public facilities, security cameras have increased. The reasons are crime prevention, protecting people and offering data after a happening. Actually, there has been a report [1] that crimes were decreased by security cameras as many as possible. From these, it is important that we install security cameras in public areas. However, we expect to increase too much data of movie by increasing cameras too.

This paper explains managing security cameras and improvement of security using data of cameras. When a human has dangerous objects, we define the situation as dangerous. First, we detect humans in camera's images. Because only tools are not dangerous. Second, we detect hands of humans by skeleton detection. We cut out the rectangle around the position of hands. If the dangerous objects are in the rectangle, the situation is dangerous. In this paper, we define what "dangerous objects" is a kitchen knife (knife). The reason is that knives are tools held by hand and have the ability to kill people.

II. measurement

A. *OpenPose* [2]

OpenPose is an algorithm that detects skeleton data of human by deep learning. OpenPose can detect skeleton in 2D images. The aim of this paper is to detect dangerous situations in images by security cameras. We think it is suitable for this aim, therefore we use OpenPose. First, we detect skeleton by OpenPose. Fig. 1 shows an example of skeleteton detection. By such a skeleton detection, we judge whether or not there are humans in a scene image. We obtain the position of the hands in detected skeleton. We cut out a rectangle region around hands. We try to detect dangerous objects in cut out images.

B. *VGG-16* [3] [4]

VGG-16(VGG) is a model of CNN(Convolutional neural network) composed of 13 convolution layers and 3 fully connected layers. VGG performs image recognition and image classification. VGG is relatively simple, but the model has deep layers and extracts features well. We detect dangerous objects by VGG in a cut out image.

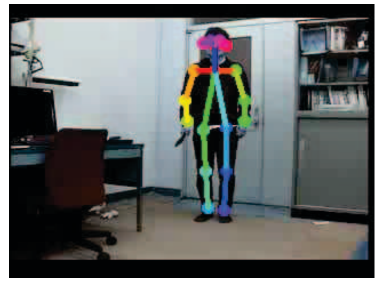


Fig. 1. An example of skeleton detection by OpenPose.

Fig. 2. An example of cut out image by pan-tilt camera. This image is very blocky.

When we detect dangerous objects in the image (Fig. 2), the situation is dangerous. Keras, which is a neural network library, has VGG models that have weights trained by ImageNet [5]. These models can classify objects according to 1,000 classes. There is a knife label ("cleaver") in the 1,000 classes.

III. Transfer learning for knife detection

A. *Pan-tilt camera*

The camera used for the experiment is a pan-tilt camera. The pan-tilt camera can move the direction of camera lens. We can obtain a wide view by pan and tilt functions.

In this paper we propose a countermeasure of too much data and cameras for security. This proposed method may reduce increasing cameras by getting wide views of this camera. Furthermore, this camera has a zoom function. When objects

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

will be magnified, it will be easy to detect the objects that humans hold with their hands.

We detected a knife by VGG of Keras. However the target objects were very small in size and blocky (Fig. 2). Therefore, we constructed a VGG model that could detect a small object. We created a new model of the VGG by transfer learning. We explain about training of a new model in the next subsection.

B. VGG transfer learning

We created a VGG model. Relearning labels were knives and cellphone that human have with hands. A learning dataset consists of 349 images obtained from imageNet dataset and 78 images by a pan-tilt camera. Pan-tilt camera's images were increased by transferring the values of RGB, transferring the value of gamma, and flipping horizontal. Total training data of knives were 427 images. Negative data were "cellphone" images by imageNet dataset. The transfer VGG model was trained after the 15th layer. We thought that a model of VGG could be constructed by this transfer learning.

C. VGG model evaluation

In this subsection, we evaluated the transfer learned VGG model. We prepared test images that were other training data. The test images were constructed by 50 images with a knife and 50 images without a knife.

Precision was 95.2% and recall was 80% as a result for those data. Next, we experimented this VGG model with images by a pan-tilt camera.

D. Experimental setup for knife detection

We used pan-tilt camera's images. The images were converted from a movie. A 20s male who had a knife was in the movie. The total images were 446 images that were in around his hand with a knife and without a knife. We detected a knife with the above VGG model.

IV. RESULT AND CONSIDERATION

A. Result

Precision was 58.5% and recall was 54.3% for the pan-tilt camera images. We used a movie by pan-tilt cameras. The reason of low precision is due to high FP (False Positive) and the reason of low recall is due to high FN (False Negative). We considered why the values of FP and FN were high.

B. Consideration

Table II is the result about Precision and Recall. Firstly, we consider the values of TP (True Positive) and FN. According to visual contact, the VGG model could not recognize knives that were very black or very white. It is apparent that the color of a knife was shifting by metallic luster. Therefore, the color transferred images were added to the training data. For example, HSV conversion and changing the values of RGB were carried out. For the reason, other color conversion will be carried out to the test images too. Because we defined a dangerous object as a knife, it is necessary that we think color change by metallic luster.

Secondly, we consider the values of TN (True Negative) and FP. We used skeleton imges obtained by Openpose. Fig. 3 is an example of the sletlton images. Detection results of human's other parts were also inputs to the VGG model as a part of the image. We thought the value to FP increased owing to circles or lines in skeleton images. 62 images out of all FP results were skeleton images.

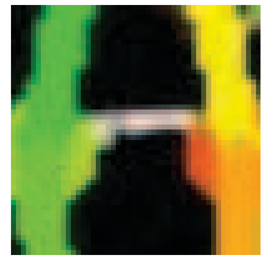


Fig. 3. An example of the FN images. Green ,orange and yellow lines are skeleton detection results.

TABLE I
RESULT OF DETECT KNIFE BY A VGG MODEL.

TP	121	FP	86
TN	137	FN	102
Total			446

TABLE II
PRECISION AND RECALL

Precision	58.5%(121/(121+86))
Recall	54.3%(121/(121+102))

V. CONCLUSION

We proposed a method of knife detection for security improvement with images of security cameras. It was shown that the relearned VGG model could recognize small and blocky objects in this paper. We used a pan-tilt camera for getting data, because we thought to be able to solve managing too much data by pan and tilt functions. We recgnized knives in images with VGG. However, we couldn't use a VGG model of keras.For the reason, we carried out transfer learning. We trained a VGG model with a dataset that was increased the number of images by changing the values of RGB. As a result, Precision was 58.5% and Recall was 54.3% for the pan-tilt camera images.

REFERENCES

[1] Metoropolitan Police Department. Security camera system in public facilities, 2017, from http://www.keishicho.metro.tokyo.jp/kurashi/anzen/anshin/gaitocamera.html

[2] Gines Hidalgo, Zhe Cao, Tomas Simon, Shih-En Wei, Hanbyul Joo, and Yaser Sheikh, "Realtime Multi-Person 2D Human Pose Estimation using Part Affinity Fields", April2017

[3] Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Visual Recognition", Octobar 2014

[4] Keras Documentation - Applications, from https://keras.io/ja/applications/#vgg16, July 2019

[5] imageNet, from http://www.image-net.org/

Zero-Phase Impulsive Noise Suppression with Iterative Phase Reconstruction

1st Aki Ishibashi
Faculty of Information Science and Engineering
Kyoto Sangyo University
Kyoto, Japan
i1986016@cc.kyoto-su.ac.jp

2nd Arata Kawamura
Faculty of Information Science and Engineering
Kyoto Sangyo University
Kyoto, Japan
kawamura@cc.kyoto-su.ac.jp

***Abstract*—This paper proposes a zero-phase impulsive noise suppression method with an iterative phase reconstruction. A zero-phase signal is obtained as IFFT of FFT spectral amplitude, and the impulsive noise suppression is achieved on the zero-phase domain. So, the zero-phase impulsive noise suppression is a process of the FFT spectral amplitude, and the spectral phase is unprocessed. To improve the noise suppression capability, we introduce an iterative phase reconstruction method to the zero-phase impulsive noise suppressor. The iterative phase reconstruction generates spectral phase with no contradiction on FFT overlap-add method. Simulation results showed that the introduced iterative phase reconstruction improves PESQ value in comparison to the conventional zero-phase method.**

***Index Terms*—zero-phase signal, impulsive noise, noise suppression, phase reconstruction**

I. Introduction

In many single channel noise suppression methods, FFT is performed for each analysis frame, and noise spectral amplitude is suppressed where the spectral phase is usually not processed [1]–[4]. Noise suppression processing is performed independently in each frame, and the noise suppressed spectrum is converted to a time domain signal based on an overlap-add method. Unfortunately, since the spectal amplitude is changed and the spectral phase is not changed, there is no guarantee that the frames are properly connected.

To connect the frames with no contradiction, Gliffin and Lim proposed an iterative phase reconstruction method (we called it as 'GL') [5]. The GL method is a batch processing, and cannot implemented as an online system. Based on GL, online phase reconstruction methods are proposed [6], [7]. They iteratively reconstruct spectral phase by frame-by-frame.

Phase reconstruction technique is important in the area of impulsive noise suppression especially in low SNR environment, because impulsive noise makes a linear phase. The linear phase works to concentrate energy of signal in a short duration in time domain. Hence, we have to process not only spectral amplitude but also spectral phase for impulsive noise removal.

In this paper, we introduce an iterative phase reconstruction method into a zero-phase impulsive noise suppressor [1], [2]. The zero-phase signal is obtained as IFFT of FFT spectral amplitude. In the zero-phase domain, the energy of impulsive noise concentrates around the origin. Impulsive noise suppression is acheved by replacing zero-phase signals around the origin [1]. Since zero-phase signal is obtained from spectral amplitude, the impulsive noise suppression is effective only on the spectral amplitude.

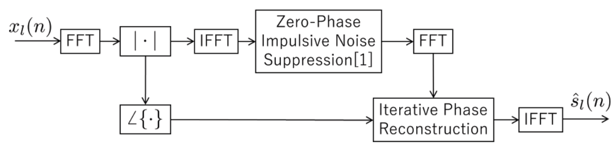


Fig. 1. Block diagram of proposed method.

As phase reconstruction medhods to be introduced to the zero-phase impulsive noise suppressor, we use RTISI (Real-Time Iterative Spectrogram Inversion) [6] and RTISI-LA (RTISI with Look Ahead) [7]. Since these methods are online algorithm, we can easily introduce into the zero-phase impulsive noise suppressor. RTISI is a method of reconstructing the spectral phase by repeating FFT and IFFT for the current frame and the past frame. Also, RTISI-LA, which is an extended version of RTISI, reconstructs spectral phase referring to future frames as well as the current frame and the past frames. Hence, in RTISI-LA, an output delay occurs depending on the number of future frames to be referenced.

Simulation results showed that the introduction of the iterative phase reconstruction to the zero-phase impulsive noise suppressor improves PESQ value, which is an objective test recommended by ITU-T [8].

II. Zero-Phase Impulsve Noise Suppression with Iterative Phase Reconstruction

The block diagram of the proposed method is shown in Fig.1. Let an observed signal at frame l be $x_l(n)$, where $0 \le n < N$ and N is the frame length. Similarly, $s_l(n)$ and $d_l(n)$ be speech signal and noise, respectively, and we assume $x_l(n) = s_l(n) + d_l(n)$.

$X_l(k)$, $S_l(k)$, $D_l(k)$ denote FFTs of $x_l(n)$, $s_l(n)$, $d_l(n)$, respectively. We have $X_l(k) = S_l(k) + D_l(k)$, where $0 \le k < N$. $X_l(k)$ is also written as $X_l(k) = |X_l(k)| \exp(j\angle X_l(k))$, where $|\cdot|$ and $\angle\{\cdot\}$ denote spectral amplitude and spectral phase, respectively, and $j = \sqrt{-1}$.

The zero-phase signal $x_l^{(0)}(n)$ is obtained by taking IFFT of $|X_l(k)|$, where $x_l^{(0)}(n) = s_l^{(0)}(n) + d_l^{(0)}(n)$ under the strict assumption $|X_l(k)| = |S_l(k)| + |D_l(k)|$. Here $s_l^{(0)}(n)$ and $d_l^{(0)}(n)$ are the zero-phase signal of $s_l(n)$ and $d_l(n)$, respectively.

An impulsive noise exists only around the origin in zero-phase domain, i.e., $d_l^{(0)}(n) = 0$ $(n \geq L)$ where L denotes an index around the origin. On the other hand, voiced speech signal maintains a periodicity in zero-phase domain. In [1], samples around the second peak are copied to around the origin. When $d_l^{(0)}(n) = 0$ around the second peak, the impulsive noise is suppressed. Taking IFFT of the processed zero-phase signal gives a noise suppressed spectral amplitude $|\hat{S}_l(k)|$.

Next, we apply RTISI [6] to $|\hat{S}_l(k)|$. We assume that the current frame index is l and the past noise suppressed speech signal in time domain, $\hat{s}_m(n)$ $(m < l)$, is already fixed. Using $\angle X_l(k)$ as an initial phase, IFFT of $|\hat{S}_l(k)| \exp(\angle X_l(k))$ gives a time domain signal. It is overlap and added to $\hat{s}_m(n)$. After that, we apply FFT to it and get the complex spectrum $A_1(k) \exp(j\theta_1(k))$. Here, $|\hat{S}_l(k)|$ and $A_1(k)$ are different, and $\angle X_l(k)$ and $\theta_1(k)$ are also different. We make $|\hat{S}_l(k)| \exp(j\theta_1(k))$. Taking IFFT of it, and overlap and added to $\hat{s}_m(n)$. Again, taking FFT of it, we have $A_2(k) \exp(j\theta_2(k))$. We make $|\hat{S}_l(k)| \exp(j\theta_2(k))$, and repeat the procedure. When the i-th iteration, we have $\theta_i(k)$ which is more appropriate to connect the frames in comparison to the initial phase [6]. RTISI refers only single past frame for the iteration. Its extended version, RTISI-LA refers some past and future frames for the iteration.

III. Simulation Results

We carried out an impulsive noise suppression for a male speech signal sampled at 16kHz, where we artificially added 40 impulse noises for generating an observed signal. We put $N = 512$ and overlap is $7/8$. The impulse noise suppression result is shown in Fig.2, where (a) shows observed signal, (b) shows the output signal of the conventional zero-phase impulsive noise suppression (ZPS) [1], (c) shows the output of ZPS with RTISI (RTISI) where the iteration was 10 times. We see from (b) that the impulse noises remain, where these noises are caused from the linear phase of the impulse signal. On the other hand, from (c), the residual noises are reduced as an effect of the iterative phase reconstruction.

For objective evaluation of noise suppression capability, we used PESQ [8]. In the objective evaluation, 50 male and 50 female speech signals were used where they were respective added 40 impulse noises. The evaluated noise suppression methods are ZPS, ZPS with RTISI (RTISI), ZPS with RTISI-LA (RTISI-LA), ZPS with Griffin and Lim method (GL). 10 iterations applied to RTISI and RTISI-LA, and 100 iterations applied to GL. We note that only GL cannot be implemented as an online system.

Averaged values of PESQ for 100 speech signals are shown in Fig.3, where the observed signal was 0.8. ZPS acheved PESQ of 2.46. Other phase reconstruction methods improved about 0.1$\sim$0.2 of PESQ in comparison to ZPS.

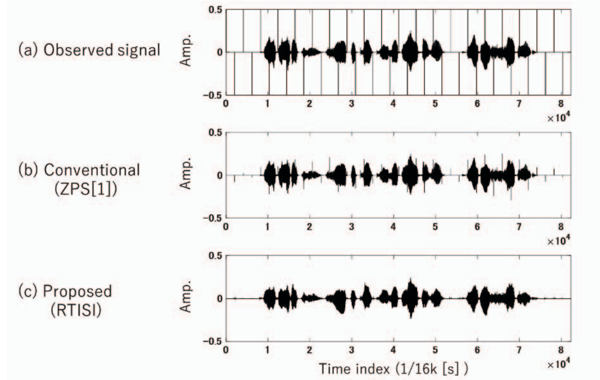


Fig. 2. Waveforms of noise suppression results. (a) observed signal, (b) conventional method (ZPS [1]), (c) Proposed method (RTISI).

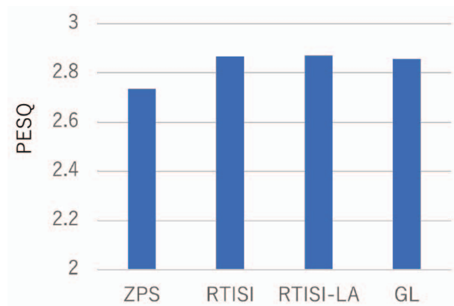


Fig. 3. Averaged PESQ values for ZPS [1], RTISI [6], RTISI-LA [7], GL [5]. Averaged PESQ for observed signal was 0.8.

IV. Conclusion

We proposed an impulsive noise suppressor which is based on a zero-phase signal replacement and an iterative spectral phase reconstruction method. The introduction of the phase reconstruction method RTISI or RTISI-LA to the zero-phase impulsive noise suppressor was useful to improve PESQ value.

References

[1] W. Thanhikam, Y. Kamamori, A. Kawamura, and Y. Iiguni, "Stationary and non-stationary wide-band noise reduction using zero phase signal," IEICE Trans. Fundamentals, vol.E95-A, no.5, pp.843–852, May 2012.

[2] S. Kohmura, A. Kawamura, and Y. Iiguni, "A zero phase noise reduction method with damped oscillation estimator," IEICE Trans. Fundamentals, vol.E97-A, no.10, pp.2033-2042, Oct. 2014.

[3] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. ASSP., vol.27, no.2, pp.113–120, April 1979.

[4] A. Kawamura, W. Thanhikam, and Y. Iiguni, "A speech spectral estimator using adaptive speech probability densityfunction," Proc. of 18th European Signal Processing Conference (EUSIPCO-2010), pp.1549-1552, Aug. 2010.

[5] D.W. Griffin and J.S. Lim, "Signal estimation from modified short-time Fourier transform," IEEE Trans. on ASSP, vol.32, no.2, pp.236–243, April 1984.

[6] G.T. Beauregard, X. Zhu, and L. Wyse, "An efficient algorithm for real-time spectrogram inversion," in Proc. 8th Int. Conf. Digital Audio Effects (DAFX-05), pp. 116?121, Sep. 2005.

[7] X. Zhu, G. Beauregard, and L. Wyse, "Real-time signal estimation from modified short-time Fourier transform magnitude spectra," IEEE Trans. ASSP, vol.15, no.5, pp.1645–1653, July 2007.

[8] ITU-T Recommendation P.862, Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, 2001.

Target Speech Signal Extraction Using Variance of Phase Difference

1st Eiji Saito
Faculty of Information Science and Engineering
Kyoto Sangyo University
Kyoto, Japan
i1986070@cc.kyoto-su.ac.jp

2nd Arata Kawamura
Faculty of Information Science and Engineering
Kyoto Sangyo University
Kyoto, Japan
kawamura@cc.kyoto-su.ac.jp

***Abstract*—It has been attracted attention to a hearable device which is one of wearable devices on human ears for communications, entertainment, health coaching, and so on. For human-to-human speech communication, an important issue of the hearable device is to enhance a target speech signal in an environment where multiple speech signals exist. In this paper, we propose a method to extract the target speech signal under the assumption that the target speaker be the nearest and the front to the hearable device user. The proposed method separates an observed speech mixture to each speech signals and estimates the target speech signal based on a variance of a phase difference between both ears. The proposed method can be implemented as a batch algorithm or an online algorithm. We evaluate the capability of the proposed method on both algorithms, and show that the online algorithm is more effective to a hearable device.**

***Index Terms*—source extraction, distance estimation, online**

I. Introduction

Recently, hearable devices have attracted attention. Hearable devices are small and ear-mounted devices to assist or enhance human hearing in a variety of acoustic environments [1]. One of the important feature of hearable devices is improving speech communication in an environment where multiple speech signals exist. There is ILRMA [2] as a blind source separation method for such environment. This method shows high separation performance when $N \leq M$, where N is the number of sources and M is the number of microphones. However, it is difficult to apply ILRMA when $N > M$. In recent years, source separation method based Deep Neural Networks (DNN) also have been studied [3], and it has been reported that high performance can be obtained. However, DNN-based methods require training data in advance. On the other hand, as a method which does not require prior learning without depending on the number of sources, there is a source separation method by Time-Frequency masking(T-F masking) [4] [5]. These methods are equally separated without considering the importance of source. When we obtain not only the separation but also the distance and direction of each source, we can extract only the target voice based on such information.

In this paper, target speech signal extraction is derived by introducing distance estimation and direction estimation based on variance of phase difference [6] to the conventional method [4]. We assumed that only the nearest source to the front is the target source. The proposed method can be implemented as a batch algorithm or an online algorithm.

II. Proposed method

The flow of the proposed method is shown in the Fig.1. Here, $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^T$ is the observed signal at time t, and the subscript indicates microphone index.

First, $\boldsymbol{x}(t)$ is transformed into the time-frequency domain by short-time Fourier transform (STFT). The observed spectrum is expressed as

$$\boldsymbol{X}(\tau, f) = \boldsymbol{H}_k(f) S_k(\tau, f) + \boldsymbol{N}(\tau, f), \tag{1}$$

where $\boldsymbol{X}(\tau, f) = [X_1(\tau, f), X_2(\tau, f)]^T$, and $X_j(\tau, f)$ and $S_k(\tau, f)$ represent the STFT of $x_j(t)$ and the k-th source signal, respectively. Also, τ and f indicate the frame index and frequency, respectively. $\boldsymbol{H}_k(f) = [H_{1k}(f), H_{2k}(f)]^T$, and $H_{jk}(f)$ is the STFT of the impulse response from source k to microphone j. $\boldsymbol{N}(\tau, f)$ is a noise term. Next, single voice activity (SVA) detection is applied to the observation signals. Then, Leader-Follower clustering (L-F clustering) is performed on the SVA frames. The direction and distance estimation are performed by using the phase difference of each cluster obtained by L-F clustering.

When only direct sound exists, the phase difference $\phi(f)$ is expressed as

$$\phi(f) = af + 2n\pi, \tag{2}$$

where a is slope of phase difference, n is an arbitrary integer. Here, assuming that the sound velocity v and the distance D between the microphones are constants, the direction θ of the sound source can be obtained as follows.

$$\theta = \sin^{-1}(av/2\pi D) \tag{3}$$

In other words, the direction θ of the sound source can be obtained from the slope a of the phase difference. From Eq. (2), phase difference ϕ is proportional to the frequency when only direct sound exists. However, reverberation actually exists. Due to the effect of the reverberation, the linear relationship between ϕ and f is disturbed and variance occurs.

According to reference [6], it is reported that the variance of the phase difference increases as the distance between the

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

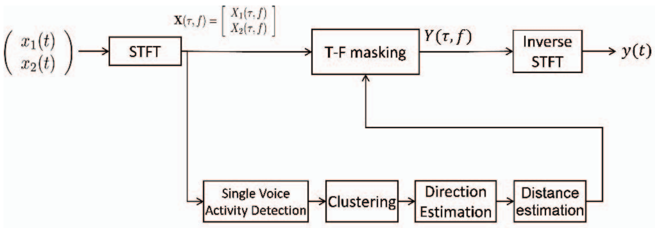


Fig. 1. Flow of the proposed method

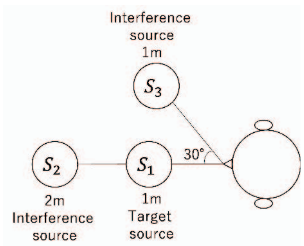


Fig. 2. Arrangement of source and microphone.

microphone and the sound source increases. Using this result, the nearest source is identified.

The specific way of obtaining the target source is as follows. The slope a of the phase difference is can be obtained as

$$a = \arg\min_k \varepsilon(k), \tag{4}$$

$$\varepsilon(k) = \sum_{i=0}^{M} (\phi_i - k f_i + 2n\pi)^2, \tag{5}$$

where f_i is frequency, ϕ_i is phase difference value at f_i obtained from the cluster, and M is the number of frequencies. Also, variance of phase difference is defined as

$$V = \varepsilon(k)/M. \tag{6}$$

It is determined that the distance is closer as the variance is smaller.

Finally, the cluster of the target source is extracted from the results of variance of phase difference and direction estimation of each obtained cluster. Only clusters in front are selected from the result of direction estimation. Here, the range that is judged to be the front is -5 degrees to 5 degrees. When there are multiple clusters in the front direction, the cluster with the smallest variance is selected as the target cluster. A separated spectrum is obtained by creating a T-F mask of this cluster and multiplying it with the observed spectrum. Separated signal in the time domain is obtained by inverse STFT of the separated spectrum. This system can be implemented as batch or online algorithm.

III. Evaluation

In order to confirm the effectiveness of the proposed method, we carried out speech extraction simulation. To create the observed signals, the speech signals were convoluted the impulse response which was measured in the arrangement shown in Fig.2. In the simulation, 5 male and 5 female Japanese speech signals in the ATR speech database [8] were used. The experimental conditions are shown in Table I. The observed signals were prepared for a total of 720 combinations (6 combinations of the placement of each speaker and 120 combinations of speakers). Source separation performance is evaluated by the SDR [7]. The results are shown in Table II. We see that the online algorithm has higher performance. The processing time of the online algorithm was measured, it was 8 seconds for speech of about 40 seconds. From this, it is prospected that the proposed method can work in a real time.

TABLE I
EXPERIMENT CONDITIONS

number of sources	$N = 3$
number of microphones	$M = 2$
Sampling frequency	$16kHz$
STFT frame size	4096
STFT frame shift size	256

TABLE II
EXPERIMENT RESULT

	Batch Algorithm	***Online Algorithm***
SDR[dB]	5.35	6.20

IV. Summary

In this paper, we proposed a method to extract the nearest source in the front direction by distance estimation and direction estimation using phase difference variance. The effectiveness of the proposed method is clarified by experiments. We confirmed that the online algorithm improves 0.85 over the batch algorithm in SDR.

References

[1] F. Satoshi, K. Takafumi, and O. Kouji, "Human-Oriented IoT Solutions Using Hearable Technology from NEC,"NEC Technical Journal, vol.12, no.1, pp.50–54, Oct. 2017.
[2] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined Blind Source Separation Unifying Independent Vector Analysis and Nonnegative Matrix Factorization," IEEE/ACM Trans. ASLP, vol. 24, no. 9, pp. 1626–1641, Sep. 2016.
[3] S. Mogami, H. Sumino, D. Kitamura, N. Takamune, S. Takamichi, H. Saruwatari, and N. Ono, "Independent deeply learned matrix analysis for multichannel audio source separation," EUSIPCO, 2018., pp. 1557-1561.
[4] A. Matsuda, A. Kawamura, Y. Iiguni, "Low computational two-channel blind source separation using single voice activity segment for unknown number of sources," Acoustical Society of Japan, volume 72, issue 3, pp.115-122, Mar. 2016.
[5] H. Sawada, S. Araki and S. Makino, "Underdetermined convolutiveblind source separation via frequency bin-wise clustering and permutation alignment," IEEE Trans. ASLP, 19, 516-527, 2010.
[6] K. Murakami, A. Kawamura, Y. Fujisaka, N. Hiruma, and Y. Iiguni, "Distant Sound Suppression Using Spectral Phase Variance for Two Channel Blind Source Separation," APSIPA, 2018, pp. 272–278.
[7] E. Vincent, H. Sawada, P. Bofill, S. Makino, J. Rosca, "First stereo audio source separation evaluation campaign: data, algorithms and results," Proc, ICA 2007 , pp552 − 559, 2007.
[8] A. Kurematsu, K. Takeda, Y. Sagisaka, S. Katagiri, H. Kuwabara, and K. Shikano, "ATR Japanese speech database as a tool of speech recognition and synthesis," Speech Commun., vol.9, pp.357–363, 1990.

Machine Learning-Based Rate Control Scheme for High Efficiency Video Coding

Jui-Hung Hsieh[a], *Member*, *IEEE*, Jing-Cheng Syu, and Hui-Lan Zhang

Department of Computer and Communication Engineering,
National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan
E-mail: jhhsieh@nkfust.edu.tw[a]

Abstract—**The up-and-coming high-efficiency video coding (HEVC) offers superior compression performance compared to the previous-generation video compression standard H.264/AVC. The rate control is the most important coding tool for compressed data transmission because of the time-varying transmission bandwidth. Based on a wide range of offline data analyses and statistics, this paper presents a rapid rate control scheme that has low computational complexity for HEVC applications within electronic devices. Experimental results demonstrate that the presented scheme can perform HEVC under time-varying bit-rate conditions.**

Keywords—machine learning, motion estimation, rate control, HEVC

I. Introduction

The latest video coding standard HEVC [1] provides many coding tools for realizing massive video compression, which includes coding-sized blocks of *coding tree units* (CTUs), *transform units* (TUs), and symmetric/asymmetric *prediction units* (PUs), along with up to 34 intra-prediction modes and a useful coding scheme, the advanced motion vector prediction (AMVP), for inter-frame prediction or intra-frame prediction. For awareness of the time-varying transmission bandwidth, the HEVC reference software (i.e., HM 16.6 [2]) provides a remarkable rate-control scheme for precisely controlling the coding bit rate of video coding design. The introduced rate-control scheme can perform HEVC with a time-varying coding bit rate, and the experimental results under diverse video-sized test sequences demonstrate excellent coding performance in terms of coding quality and coding bit rate. However, the introduced rate-control technique is difficult to implement in power-limited electronic devices because of the complicated algorithmic modeling and substantial computational complexity.

To overcome these issues, a wide variety of solutions [3-9] have been proposed in the literature. The work in [3] introduced a group of picture-level rate control methods with low bit fluctuations and superior rate-distortion performance. A λ-domain frame-level rate control scheme for low-delay video coding was presented in [4] by using an improved parameter updating method and an adaptive frame-level bit allocation. The design in [5] proposed a new λ-QP relationship for the conventional λ-domain rate control algorithm in high-dynamic-range video coding. In addition, the study in [6] introduced an intra-frame rate control scheme that is based on a reinforcement learning method and combines the CTU-layer texture complexity and the bit balance. Meanwhile, a rate-control algorithm with a piecewise-linear approximation of the rate-distortion curve was presented in [7] for intra-coding. The design in [8] proposed a CTU-layer rate-control scheme that is based on spatial and temporal visual masking effects for coding performance improvement. An efficacious frame-level HEVC rate-control scheme using an enhanced rate-distortion model and a group-of-picture (GOP) level bit allocation approach with the recursive Taylor expansion were introduced in [9]. The aforementioned studies indeed presented exceptional rate-control designs for solving bit allocation problems. Nonetheless, this paper introduced a machine learning (ML)-based HEVC rate-control design scheme that achieves low-hardware-complexity and low-coding-time design.

The remainder of this paper is organized as follows. Our proposed rate-control method and the corresponding simulation results are presented and shown in Section II. Section III summarizes and concludes this paper.

II. Proposed Framework

We intend to establish a rapid HEVC rate-control framework for mobile video coding applications. The scheme is comprised of two procedures and the detailed process is elaborated as follows.

A. Proposed ML-based HEVC Rate-Control Scheme

Initially, we will analyze the mathematical relationship between the coding bit rate and the quantization parameters (QPs) because the coding bit rate is a function of the quantization parameter (QP) [10] from our offline simulation with diverse video-sized test sequences [11]. At this stage, denoted as data training stage, we can acquire a curve via the linear regression method [12] by using MATLAB. Then, we develop and establish the data classification and data prediction strategy for the proposed framework's use. The classification strategy that is used in this paper is a low-hardware-complexity method instead of the conventional classification method, such as a support vector machine (SVM) [13], which is difficult to apply in VLSI. Meanwhile, the prediction strategy that is used here simultaneously considers the coding and hardware complexity by using the temporal and spatial correlations [14] within the video coded frames. Then, we integrate these mathematical formulations into the constrained optimization model; accordingly, we will receive the optimized QP from the proposed constrained optimization framework. After that, the received optimum QPs are fed into the HEVC for use by the function block of motion

TABLE I. PERFORMANCE COMPARISON WITH CLASS-D-SIZED (416×240-SIZED) FOR THE LOW DELAY MAIN P SETTINGS

Video Sequence	Frame rate (frames/sec)	CBR_T (kilobits/sec)	∆PSNR (dB)	∆BR (%)
BasketballPass	30	2285	-0.19	-3.79
		1143	-0.67	-6.53
		114	-1.48	-24.56
BlowingBubbles	30	2285	2.58	55
		1143	1.42	39.67
		114	-0.99	-27.76
BQSquare	30	2285	3.19	55.23
		1143	2.25	48.15
		114	-0.49	-18.16
RaceHorses	30	2285	2.13	37.78
		1143	1.52	30.07
		114	0.51	2.3
Average			0.82	15.62

estimation (ME). Finally, real-time coding-bit-rate-aware HEVC is performed.

B. *Simulation Results*

To evaluate our proposed HEVC rate-control scheme, we implement the proposed rate-control algorithm in HEVC reference software [2] for a full-search ME scheme and simulate it on a workstation with a 2.60 GHz Intel® Xeon™ E5-2620 v3 24-core CPU. The coding performances are compared in terms of the coding quality difference (i.e., $\triangle$ PSNR = $PSNR_{PROS}$ - $PSNR_{[2]}$) and the coding bit rate difference (i.e., ∆BR= BR_{PROS} - $BR_{[2]}$ / $BR_{[2]}$) under target coding-bit-rate limitations (i.e., CBR_T) for a class-D test sequence (i.e., 416 ×240). In which, $PSNR_{PROS}$ and BR_{PROS} respectively denote the PSNR and BR of our proposed algorithm.

The simulation results are listed in Table 1, according to which our proposed rate-control scheme performs adequately for HEVC due to coding quality improvement and low coding bit rate degradation (i.e., average PSNR increase of 0.82 dB; average BR increase of 15.62%).

III. CONCLUSION

A rapid HEVC rate-control design based on the machine learning method was presented in this paper for accelerating the coding time and the coding complexity is lower due to our offline training model development instead of run-time implementation. The experimental results demonstrate that the presented ML-based rate-control design is suitable for time-varying HEVC with acceptable coding performance. Specifically, the low-coding-complexity and low-coding-time design scheme is adequate for mobile video applications.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology of the Republic of China under Grant MOST 107-2221-E-992-082-MY2.

REFERENCES

[1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 22, pp. 1649-1668, 2012.

[2] *Jonint Collaborative Team on Video Coding (JCT-VC), HM 16.6, Reference Software*. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6/

[3] F. Song, C. Zhu, Y. Liu, and Y. Zhou, "A new GOP level bit allocation method for HEVC rate control," in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2017, pp. 1-4.

[4] H. Guo, C. Zhu, Y. Gao, and S. Song, "A frame-level rate control scheme for low delay video coding in HEVC," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, 2017, pp. 1-6.

[5] J. Mir, D. S. Talagala, and A. Fernando, "Optimization of HEVC λ-domain rate control algorithm for HDR video," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, 2018, pp. 1-4.

[6] J. H. Hu, W. H. Peng, and C. H. Chung, "Reinforcement learning for HEVC/H. 265 intra-frame rate control," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1-5.

[7] V. Sanchez, "Rate control for HEVC intra-coding based on piecewise linear approximations," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 1782-1786.

[8] H. Wang, L. Song, R. Xie, Z. Luo, and X. Wang, "Masking effects based rate control scheme for high efficiency video coding," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1-5.

[9] H. Guo, C. Zhu, S. Li, and Y. Gao, "Optimal bit allocation at frame level for rate control in HEVC," *IEEE Transactions on Broadcasting,* vol. 65, pp. 270-281, 2018.

[10] L. Li, B. Li, H. Li, and C. W. Chen, "λ-domain optimal bit allocation algorithm for high efficiency video coding," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 28, pp. 130-142, 2018.

[11] F. Bossen, "Common Test Conditions and Software Reference Configurations, Document JCTVC-K1100 " Oct, 2013.

[12] W. W. Hines, D. C. Montgomery, D. M. Goldman, and C. M. Borror, *Probability and statistics in Engineering*, 4th ed. New York: Wiley, 2003.

[13] X. Wu, W. Zuo, L. Lin, W. Jia, and D. Zhang, "F-SVM: combination of feature transformation and SVM learning via convex relaxation," *IEEE transactions on neural networks and learning systems,* vol. 29, pp. 5185-5199, 2018.

[14] K. I. Lee, A. C. Tsai, J. F. Wang, and J. F. Yang, "Efficient inter mode prediction based on model selection and rate feedback for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 21, pp. 708-716, 2011.

The regression model of NOx emission in a real driving automobile

1st Hung-Ta Wen
Mechanical Engineering
Chung Hsing University
Taichung, Taiwan
wenhungta@gmail.com

2nd Ming-An Li
Mechanical Engineering
Chung Hsing University
Taichung, Taiwan
a901691@gmail.com

3rd Jau-Huai Lu
Mechanical Engineering
Chung Hsing University
Taichung, Taiwan
jhlu@dragon.nchu.edu.tw

***Abstract*—Although the Portable Emission Measurement System (PEMS) has become the official certification procedure for vehicles, but it is very expensive and cannot be used as a monitoring device for massive use. The purpose of this study is to establish a low-cost vehicle pollution monitoring system (NGK NOx sensor, Arduino with CanBus) and uses this system to measure the actual vehicle road pollution emissions. In addition, proposes an ANN nonlinear autoregressive exogenous model (NARX) to predict NOx emissions.**

Keywords—PEMS, NARX, NGK, Arduino, CanBus

I. Introduction

According to the World Health Organization (WHO) report, there are 7 million people passed away due to air pollution in 2018[1]. And the air pollution source is mainly from transportation emission.

Establish emissions monitoring is an effective way to control pollution. However, the actual road measurement time cost is high, so need to develop a new way to reduce the cost [2].

A solution proposes a system consists of an IOT device, a smartphone and web-based simulation model to monitor air quality exposure and pollutant emissions [3]. In addition, using three machine learning techniques (SVM, RF, ANN) can predict fuel consumption successfully [4].

This study will use NGK NOx sensor with pitot tube flowmeter to build a set of vehicle pollution measurement equipment, and use Arduino with CanBus expansion version as the core control, with NB-IOT extended version to establish a vehicle pollution monitoring system. Finally, use ANN NARX model to predict NOx emissions.

II. Methon

A. Data

The experimental vehicle adopts the TOYOTA Zace 1.5L gasoline engine version. It was produced in 1991. As shown in the Fig. 1, it can be seen that the exhaust pipe has been installed with a special exhaust pipe for sampling. And the experimental route is shown in the Fig. 2.

Fig. 1. Vehicle

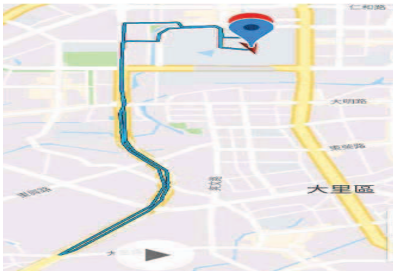


Fig. 2. Route

The NGK NOx sensor (Fig. 3) is a component of an SCR system developed by NGK. This study uses Arduino with the CanBus expansion version (Fig. 4) to enable the Arduino microcontroller to utilize the read/write CanBus data link and control the NGK NOx sensor.

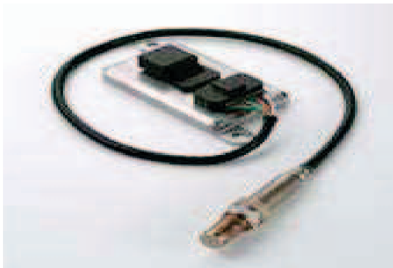


Fig. 3. NGK NOx sensor

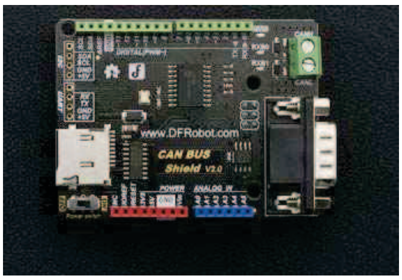


Fig. 4. Arduino CanBus

The total driving time is 25 mins and 44 seconds, the total driving distance is 7.35 km, and the average speed is 18 km/hr. The NOx emissions is 0.213 g/km.

B. Neural network

NARX models are widely used in nonlinear dynamic behavior analysis applications[5]. This study uses MATLAB NARX model to predict NOx emissions (Fig. 5). The whole dataset was splitted into three parts : train dataset (70% of whole dataset), test dataset (15% of whole dataset) and validation dataset (15% of whole dataset). The input features are vihicle speed, accelerate and flow rate. And the target value is NOx value.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

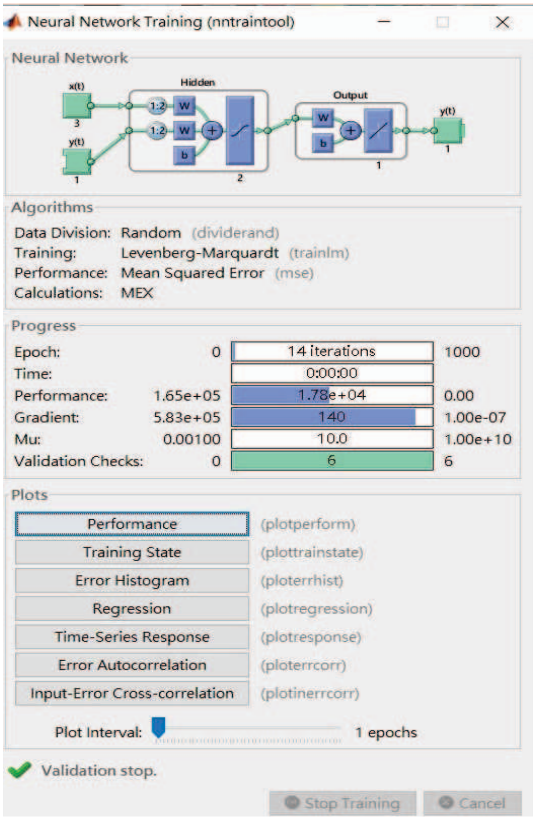


Fig. 5. MATLAB NARX model

III. Results and Discussion

The Neural Network training regression results are shown in Fig 6. And the predicted value against the actual value are shown in Fig 7 and Fig 8. The performance of the trained model are evaluated and shown in Table Ⅰ. After obtaining the NOx prediction values, then can calculate the emissions per kilometer and compared to NOx emissions calculated from dataset shown in Table Ⅱ.

Road traffic conditions such as heavy traffic, traffic lights, etc., affect driving speed when driving. And these will also determine the NARX model prediction performance.

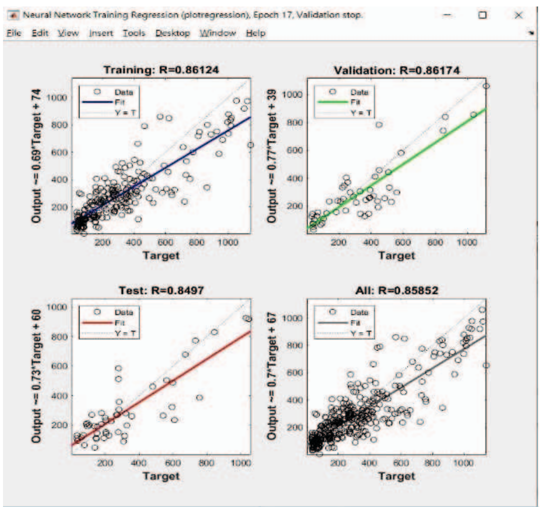


Fig. 6. Neural Network Training Regression

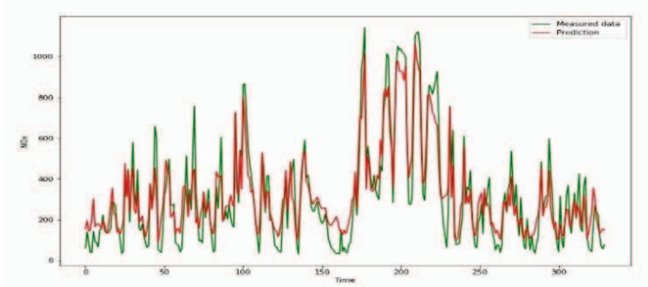


Fig. 7. Predicted value against the actual value

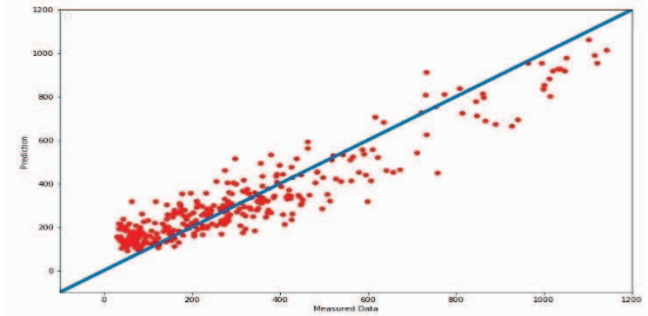


Fig. 8. Predicted value against the actual value

TABLE I. PERFORMANCE OF THE TRAINED MODEL

Error	Performance		
	RMSE	***MAE***	***R^2***
statistical errors	96.785	77.487	0.856

TABLE II. NOX EMISSIONS ERROR %

Emissions	Error % calculation		
	From dataset calculation	***NARX model prediction***	***Error %***
NOx	0.213g/km	0.222g/km	4.2

IV. Conclusion

This study establish a low-cost vehicle pollution monitoring system (NGK NOx sensor, Arduino with CanBus). And propose an ANN nonlinear autoregressive exogenous model (NARX) to predict NOx emissions. The NOx emission value is compared with the actual measured value, and the error is 4.2%.

References

[1] World Health Organization and others, “WHO Global Urban Ambient Air Pollution Database,” 2018.

[2] Deng-SiangJhang, Real road emission measurement and analysis with PEMS for light diesel vehicles:NCHU, 2018.

[3] Thibault, Laurent, et al. "A new GPS-based method to estimate real driving emissions." 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2016.

[4] Perrotta, Federico, Tony Parry, and Luis C. Neves. "Application of machine learning for fuel consumption modelling of trucks." 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017..

[5] Pisoni E., Farina M., Carnevale C., Piroddi L., "Forcasting peak air pollution levels using NARX models." Engineering Applications of Artificial Intelligence (Elsevier) Vol. 22 (2009) p. 593-602

Dictionary learning on l_1-norm fidelity for non-key frames in distributed compressed video sensing

Tsugumi Oishi
National Institute of Technology,
Kurume College
1-1-1 Komorino, Kurume-shi,
Fukuoka, 830-0001, Japan
Email: a3503to@kurume.kosen-ac.jp

Yoshimitsu Kuroki
National Institute of Technology,
Kurume College
1-1-1 Komorino, Kurume-shi,
Fukuoka, 830-0001, Japan
Email: kuroki@kurume-nct.ac.jp

***Abstract*—Distributed Compressed Video Sensing (DCVS), which consists of Compressed Sensing (CS) and Distributed Video Coding (DVC), is an encoding scheme transferring computational burden from encoder to decoder. By assuming that given signals are sparse, the CS enables accurate decoding only referring low dimensional observations which are obtained by low-rank random projection of original signals. The DVC divides image sequences into key and non-key frames and regards the decoding of the non-key frames as error correction using the key frames. The quality of the non-key frames depends on the design of the dictionaries. Then, many studies optimize dictionaries with convex optimization solvers for functions consisting of the weighted sum of two terms: l_2-norm error estimation term and l_1-norm regularization term. This paper proposes to use l_1-norm error instead of l_2-norm to increase the robustness against outliers. We apply ADMM (Alternating Direction Method of Multipliers), which is a convex optimization solver, to minimization the cost function. Simulation results show the proposed method generates better quality images than the conventional method.**

***Index Terms*—Distributed Compressed Video Sensing, ADMM, Dictionary learning**

I. INTRODUCTION

Distributed Compressed Video Sensing (DCVS) [1] is an encoding scheme. It consists of Compressed Sensing (CS) and Distributed Video Coding (DVC) and transfers computational burden from the encoder side to the decoder side. The CS enables accurate decoding only referring low-dimensional observations which are obtained by low-rank random projection of original signals. The sparseness of coefficient vectors is measured with l_0-norm, and this problem is classified into an NP-hard problem. Therefore many studies relax l_0-norm problems to l_1. The DVC divides image sequences into key and non-key frames and regards the decoding of the non-key frames as error correction using key frames. The quality of non-key frames depends on the design of the dictionaries. Then, many studies designed dictionaries with convex optimization problems that consist of the weighted sum of two terms: l_2-norm error estimation term and l_1-norm regularization term [1]. In this paper, we use l_1-norm error instead of l_2-norm to increase the robustness against outliers [3].

A conventional method applies Alternating Direction Method of Multipliers (ADMM) [4] to solve minimization problem and obtain dictionaries [1]. ADMM is a convex optimization solver divided into three steps: an error minimization, a coefficient-norm minimization, and a dual variable update of an augmented Lagrangian. This paper proposes ADMM applies to the problems, and adopts the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [5] for the minimization step.

II. DISTRIBUTE COMPRESSED VIDEO SENSING

A given video sequence is divided GOP (Group of Picture) which consists of a key frame $\boldsymbol{f}_K$ and non-key frames $\boldsymbol{f}_{NK}$. The former is reconstructed by the CS, and the latter is reconstructed by $\boldsymbol{f}_{SI}$, which is predicted by reconstructed key frame's motion compensated interpolation (MCI). This paper assumes GOP is key, non-key, key, non-key, key...; then, $\boldsymbol{f}_{SI}$ of a $\boldsymbol{f}_{NK}$ is predicted by using a half motion vectors of the motion vectors generated by the previous and the following key frames of the $\boldsymbol{f}_{NK}$. Both $\boldsymbol{f}_K$ and $\boldsymbol{f}_{NK}$ are divided into non-overlapped blocks and a block of $\boldsymbol{f}_K$ is reconstructed by the following convex optimization problem:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y}_K - \boldsymbol{\Phi}_K\boldsymbol{\Psi}\boldsymbol{x}\|_2^2 + \mu\|\boldsymbol{x}\|_1, \tag{1}$$

where $\boldsymbol{\Psi} \in \mathbb{R}^{N\times N}$ is the DCT matrix, $\boldsymbol{x} \in \mathbb{R}^N$ is a sparse coefficient vector using $\boldsymbol{\Psi}$ for the block, μ is a weight parameter to balance the two terms, $\boldsymbol{\Phi}_K \in \mathbb{R}^{M\times N}$ is an observation matrix, and $\boldsymbol{y}_K \in \mathbb{R}^M$ $(M < N)$ is an observation vector. The observation matrix is shared with both encoder and decoder sides, and only $\boldsymbol{y}_K$ is sent to the decoder. Decoders estimate the solution $\hat{\boldsymbol{x}}$ of (1), and generate $\hat{\boldsymbol{f}}_K = \boldsymbol{\Psi}\hat{\boldsymbol{x}}$. A $\boldsymbol{f}_{NK}$ block is reconstructed as follows:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y}_{NK} - \boldsymbol{D}\boldsymbol{x}\|_2^2 + \mu\|\boldsymbol{x}\|_1, \tag{2}$$

where $\boldsymbol{y}_{NK} \in \mathbb{R}^M$ is the observation vector which is multiplied a block of $\boldsymbol{f}_{SI} \in \mathbb{R}^N$ by $\boldsymbol{\Phi}_{NK}$, and $\boldsymbol{D}$ is a learned dictionary.

III. CONVENTIONAL METHOD

In the conventional method, the dictionaries are designed by K-SVD algorithm [2]. However, [6] shows that the dictionary obtained by using ADMM is faster than that designed with K-SVD whereas PSNR values using ADMM are slightly lower than those of K-SVD. Therefore, we design dictionaries using ADMM as the same manner of [6], and the problem is given by

$$\min_{\boldsymbol{x},\boldsymbol{D}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_2^2+\mu\|\boldsymbol{x}\|_1, \tag{3}$$

where λ_1 and μ are parameters. To apply ADMM, firstly we divide (3) into the $\boldsymbol{x}$ minimization and the $\boldsymbol{D}$ minimization problems:

$$\min_{\boldsymbol{x}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_2^2+\mu_1\|\boldsymbol{x}\|_1 \tag{4}$$

$$\min_{\boldsymbol{D}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_2^2+i_s(\boldsymbol{D}), \tag{5}$$

where i_s is the indicator function which projects each column vector of $\boldsymbol{D}$ onto unit sphere. The solutions $\boldsymbol{x}$ and $\boldsymbol{D}$ are computed alternately until their convergence.

IV. Proposed method

The proposed method replaces the l_2 fidelity term with l_1-norm and considers following minimization problem:

$$\min_{\boldsymbol{x},\boldsymbol{D}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_1+\mu\|\boldsymbol{x}\|_1. \tag{6}$$

The optimization scheme is same as the problems using l_2 fidelity term, and (6) is divided into the following two sub-solutions:

$$\min_{\boldsymbol{x}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_1+\mu_1\|\boldsymbol{x}\|_1 \tag{7}$$

$$\min_{\boldsymbol{D}} \frac{\lambda_1}{2}\|\boldsymbol{f}_{SI}-\boldsymbol{D}\boldsymbol{x}\|_1+i_s(\boldsymbol{D}). \tag{8}$$

To solve the two problems alternatively, we adopt the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [5] for the minimization step. Please see details in [7].

V. Experiments

The performance of our method is evaluated on the video sequence ”foreman” with 256×256 pixels. The parameters are as follows: $\mu = 1$ in (1), $\mu = 0.1$ in (2), $\lambda_1 = 1$ in (4)(5)(7)(8). We changed the value of the parameter μ_1 in (4)(5)(7)(8) from 0.01 to 1 in 0.01 increments. Let $\boldsymbol{\Phi}_K$ and $\boldsymbol{\Phi}_{NK}$ be random matrices generated by the standard normal distribution. We definined measurement rate (MR) of measurement matrix, which is the ratio between its height and width as follows:

$$MR = \frac{M}{N}. \tag{9}$$

The MR to the key frames and the MR of non-key frames are fixed to 0.5. Figure 1 shows that average PSNR of conventional method and proposed method when the value of the parameter μ_1 are varied. Figure 2 compares reconstructed images of both method at MR=0.5, and also shows high quality of our method. Figure 3 shows generated dictionaries of these images in Fig.2.

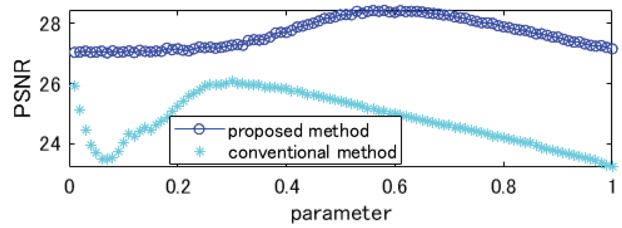


Fig. 1. Average PSNR when the value of the parameter μ_1 are varied

(a) conventional method
PSNR=31.30 ［dB］
SSIM=0.853

(b) proposed method
PSNR=34.64 ［dB］
SSIM=0.928

Fig. 2. Reconstructed images for MR=0.5

(a) DCT (b) conventional method (c) proposed method

Fig. 3. Generated dictionaries

VI. Conclusion

This study has proposed that dictionary learning of non-key frames in distributed compressed video sensing replacing l_2-norm fidelity term with l_1-norm in minimization problem. Empirical experiments show that our method reconstructed better quality images than the conventional method. In future, we will find better parameters to reconstruct images.

References

[1] F. Tian, J. Guo, B. Song, H. Liu, and H. Qin, ”Distributed compressed video sensing with joint optimization of dictionary learning and l_1-analysis based reconstruction”, IEICE Trans. Inf. & Syst, volE99-D, no.4, pp.1202-1211, 2016. DOI:10.1587/transinf.2015EDP7373.

[2] Chen, Hung-Wei, Li-Wei Kang, and Chun-Shien Lu. ”Dictionary learning-based distributed compressive video sensing.” 28th Picture Coding Symposium. IEEE, 2010. DOI:10.1109/PCS.2010.5702466

[3] J. Nocedal and S.J. Wright, “Numerical optimization,” Springer-Verlag, 1999.

[4] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation”, Computers & Mathematics with Applications, vol.2, pp.17-40, Jan. 1976. DOI:10.1016/08981221(76)90003-1.

[5] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. Imag. Sci., Vol. 2, No. 1, pp. 183–202, 2009. DOI. 10.1137/080716542

[6] Y. Akiyoshi, T. Sumi, and Y. Kuroki, “Dictionary design and disparity interpolation on distributed compressed sensing for light field image,” Proc. IEEE, APSIPA, 2017.

[7] T. Oishi and Y. Kuroki, ”An l_1-l_1-norm minimization solution using ADMM with FISTA.” 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). Vol. 3. IEEE, 2018. DOI:10.1109/ICIIBMS.2018.8549934

Design of SAR ADC to Light Charging for Optical Sensors Applications

Wen-Cheng Lai, Chung-Ling Wang, Guo-Fan Wang, Chia-Yu Tsai, Chien-Wen Chen and Yu-Wen Lo
Dept. of Electrical Engineering, National Penghu University of Science and Technology, Taiwan R.O.C.
Email: wenlai@gms.npu.edu.tw

***Abstract*—The proposed successive approximation register (SAR) analog-to-digital converter (ADC) for integrates optical sensor. The heart rate sensor including SAR ADC and OPVs are presented. To reduce DAC switching energy, a hybrid resistor-capacitor DAC is applied. The optimal design used charging near infrared laser driven (NIRLD) with thin Fresnel and OPVs light with VCSEL. The charging near infrared laser driven (NIRLD) with thin Fresnel might be a promising OPVs light power supplier and low consumption, asynchronous control logic to drive the ADC is used. A pre-amplifier-based comparator circuit is built to reduce the kickback noise from the dynamic latch designs for optical sensor applications.**

Keywords—SAR ADC, near infrared charging, optical sensors

I. INTRODUCTION

The near-infrared laser driven (NIRLD) with thin fresnel adopts wireless and optical charging power energy that organic photovoltaic devices (OPVs), which can convert optical about 980 nm directly into wireless and power charging. The designed solution redounds the NIR photovoltaic answer to the long wavelength assimilate of charging transfer (CT) states. Direct drive through high effectiveness CT condition might break open new channels for harvesting the far wavelength spectrum with thin Fresnel of OPVs irradiation. The proposed OPVs with thin Fresnel can enhance received performance. From dissertations use an active rectifier or a phase shift inverter to replace the solutions of the DC – to - DC converters from fresh issues as hard exchange [1] and maximum/ peak output [2] [3]. SAR ADC has the advantages of power efficiency compared with other types of ADC architectures [4]-[7]. Authors proposed a monotonic capacitor switching procedure to lower power consumption [4]. A literature presents a SAR ADC uses a hybrid resistive-capacitive DAC [5]. Using common-mode based switching, originally proposed for capacitive DACs, can be utilized [6].

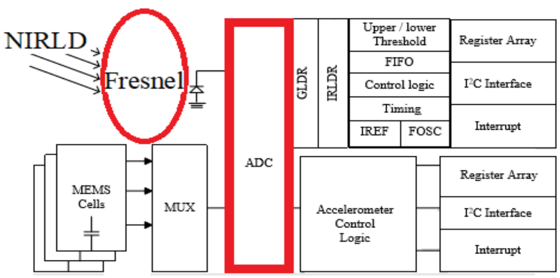


Fig. 1. The proposed SAR ADC architecture in optical sensor.

II. CIRCUITS DESIGN

Recently authors present a SAR ADC with capacitance value that uses a resistive ladder to arrange a switching circuit which reduces the input capacitance for 10-bit resolution, but also displays the predictive capacitor switching sequence to further reduce the power consumption [8]. The proposed SAR and switching logic solutions are shown the single-sided switching method and store the comparator output result [9]. The provided SAR ADC architecture composes of a T/H circuit, a comparator, a SAR logic block diagram and a resistor-capacitor hybrid DAC network as shown in Fig. 2. The proposed T/H circuit is shown in Fig.3. Fig. 4 shows comparator circuit with a preamplifier and a latch. The proposed DC-DC converter circuit is shown in Fig. 5, which is composed of two parts: oscillator and charge-pump circuits. DC-DC converter circuit can support NIRLD in heart rate monitor (HRM) and pulse oximeter (OX) module to sense photoplethysmogram (PPG) signal as show in Fig. 6. The proposed design used OPVs with vertical-cavity surface-emitting laser (VCSEL) for charging transfer (CT) and signal input. The OPVs with vertical-cavity surface-emitting laser (VCSEL) cover can enhance OPVs CT performance more than EEL and LED as show in Fig. 7.

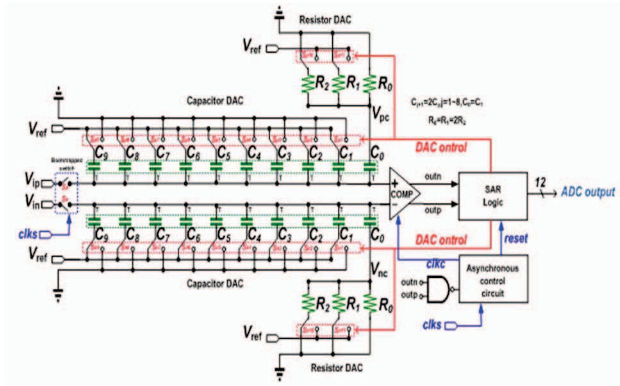


Fig. 2. The proposed 12-bit 5MS/s SAR ADC architecture.

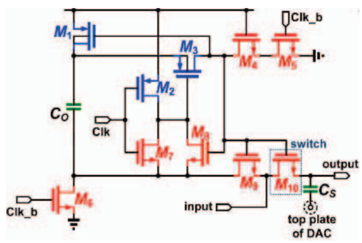


Fig. 3. Schematic of the T/H circuit.

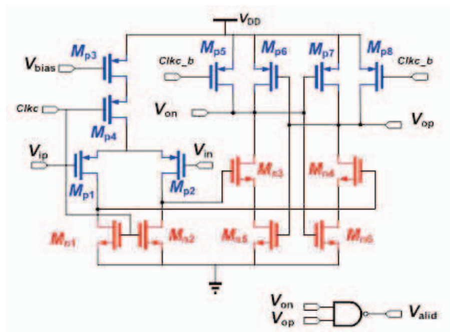
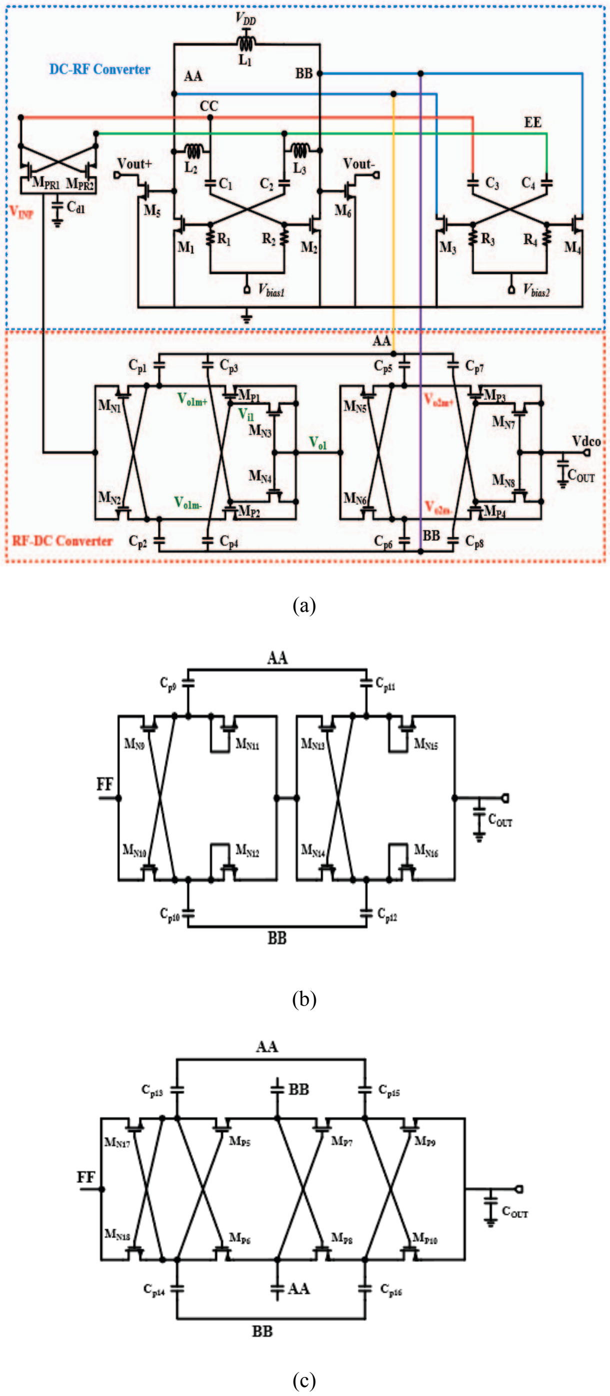


Fig. 4. Comparator circuit with a preamplifier and a latch.



(a)

(b)

(c)

Fig. 5. Schematic of the proposed DC-DC converter. (b) and (c) other charge pump circuits.

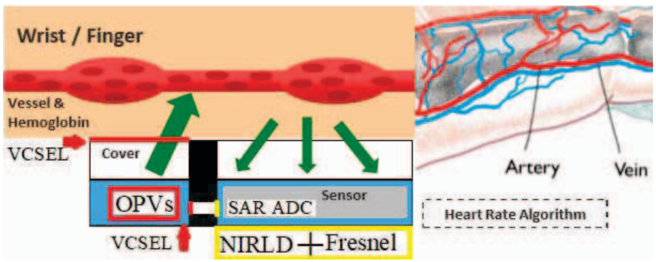


Fig. 6. DC-DC converter circuit can support NIRLD in heart rate monitor (HRM) and pulse oximeter (OX) module to sense photoplethysmogram (PPG) signal for optical sensors applications

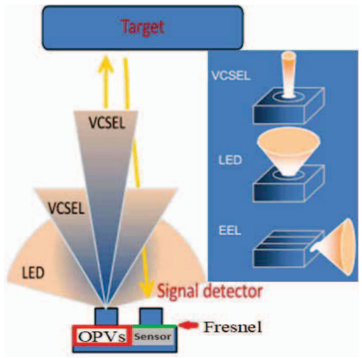


Fig. 7. The performance comparison of OPVs with vertical-cavity surface-emitting laser (VCSEL), EEL and LED.

III. Conclusions

The proposed heart rate sensor measured results is shown in Fig.8. The integrated design of SAR ADC with DAC to near infrared charging for low power consumption and high optical and bio-signal performance.

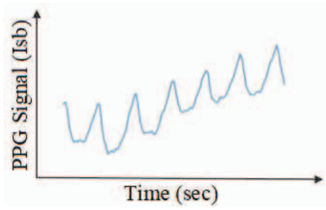


Fig. 8. The proposed heart rate sensor measured results.

References

[1] T. Diekhans and R. W. De Doncker, "A Dual-Side Controlled Inductive Power Transfer System Optimized for Large Coupling Factor Variations and Partial Load," *IEEE Trans. on Power Elect.*, vol. 30, pp. 6320-6328, Nov. 2015.

[2] K. Hata, T. Imura, and Y. Hori, "Dynamic wireless power transfer system for electric vehicles to simplify ground facilities - power control and efficiency maximization on the secondary side," *IEEE Applied Power Elect. Conf. and Exposition (APEC)*, pp. 1731-1736, 2016.

[3] G. Lovison, M. Sato, T. Imura, and Y. Hori, "Secondary-side-only simultaneous power and efficiency control for two converters in wireless power transfer system," in *Industrial Electronics Society, IECON - 41st Annual Conf. of the IEEE*, pp. 004824-004829, 2015.

[4] C.-C. Liu, S.-J. Chang, G.-Y. Huang and Y.-Z. Lin, “A 10-bit 50-MS/s SAR ADC with a Monotonic Capacitor Switching Procedure,” *IEEE J. Solid-State Circuits*, vol 45, pp. 731-740, Apr. 2010.

[5] B. Sedighi, A. T. Huynh, E. Skafidas, D. Micusik, “Design of hybrid resistive-capacitive DAC for SAR A/D converters Electronics,” *IEEE International Conference on Circuits and Systems (ICECS),* pp.508 – 511, Dec. 2012.

[6] Y. Zhu, C.-H. Chan, U.-F. Chio, S.-W. Sin, S.-P. U. R. P. Martins, and F. Maloberti, “A 10-bit 100-MS/s reference-free SAR ADC in 90 nm CMOS,” *IEEE Solid-State Circuits*, vol. 45, pp. 1111–1121, June 2010.

[7] T. Z. Chen, S. J. Chang and G. Y. Huang, “A successive approximation ADC with resistor-capacitor hybrid structure,” *IEEE VLSI Design, Automation, and Test (VLSI-DAT),* pp. 1-4, Apl. 2013.

[8] J.-F. Huang, W.-C. Lai and C.-G. Hsieh, “A 10-bit 100 MS/s Successive Approximation Register Analog-To-Digital Converter Design,” *IEICE Trans. on Electron,* vol. E97-C, no. 08, pp. 833-836, Aug. 2014.

[9] J. F. Huang, W. C. Lai and T. Ye, An 8-bit 2 MS/s Low Voltage Successive Approximation Register Analog-To-Digital Converter, *International Conference on Control Engineering and Information Technology (ICCEIT 2013)*, pp. 1-4, Nanning China, 09-11 Aug 2013.

Multi Band Antenna Design for Mobile Devices

Wen-Cheng Lai[1*] and Chang-Chun Zhang[2]
[1]*Dept. of Electrical Engineering, National Penghu University of Science and Technology, Taiwan R.O.C.*
[2]*College of electronic science and engineering, Nanjing University of Posts and Telecommunications, China.*
*Email: wenlai@gms.npu.edu.tw

***Abstract*— A laser direct structuring (LDS) antenna for multiband operation covering the GSM-850/900, GPS, DCS-1800, PCS-1900, UMTS-2100, WLAN-2400, LTE-2500 bands in the 5G communication devices is presented. The proposed antenna achieves this by exciting the antenna's wide radiating plate. The proposed antenna uses asymmetric T-type feed line structure to enchase impedance matching and achieve wide bandwidth with voltage standing wave ratio (VSWR) of 3 : 1. Due to the compact volume, only occupied 50x113mm², the proposed antenna design is fitting to be applied in the mobile devices.**

Keywords—coupling feed, mobile device application, T-type feed line, multiband.

I. INTRODUCTION

With the rapid growth of the personal portable communications, the demand for antennas with small size, light weight, high performance, low cost and ease manufacture has been considerably increased; hence the compact antenna designs have been received much attention. Several antenna designs for reducing the antenna size for mobile phone applications have been proposed, Promising on-laser direct structuring (LDS) antennas with a small printed area on the mobile device cover for GSM-850/900, GPS, DCS-1800, PCS-1900, UMTS-2100, WLAN-2400, LTE-2500 bands in the mobile device have also been reported very recently. These LDS antennas support to achieve reduced antenna size and wide operating bands. In the near future, it is expected that the LTEA (long term evolution advance) service which can provide better mobile broadband and multimedia services than the existing GSM and UMTS mobile networks will become very attractive for the mobile users. The 5G communication can also support hand-over to the existing mobile networks. For this application, how to use simply pattern and easy way to design broadband antennas, which bands is become very important. In this article, it present a promising simply pattern to cover the eight-band operation. The proposed antenna is suitable to be directly printed on the system circuit board of the mobile phone, making it easy to fabricate at and attractive for slim mobile device applications.

II. FDTD PRINCIPLE

The FDTD method, a time-domain approach was first treated by Yee [1] to deal with electromagnetic field distributions. Through the work of Yee and others, FDTD has become a powerful tool in predicting scattering problems. Workers have found applications of FDTD in transient electromagnetics [2]. One calculates the distant scattered field in back direction and one is interested in the distant refracted field in an unlimited space. All of them, the field propagates will not account for much about surrounded obstacle shape. Therefore we apply this method to analyze the designed antenna in a limited and irregular space, especially in a modern mobile transceiver which antenna must satisfy Maxwell's curl equations.

$$\nabla \times \overline{H} = \varepsilon \frac{\partial \overline{E}}{\partial t}, \tag{1}$$

where ε and μ are medium permittivity and permeability, respectively.

$$\nabla \times \overline{E} = -\mu \frac{\partial \overline{H}}{\partial t}, \tag{2}$$

Follow Yee's suggestion and denote any function of space and time as:

$$F^n(i,j,k) = F(i\Delta x, j\Delta y, k\Delta z, n\Delta t), \tag{3}$$

where Δt is the time increment and n denotes the nth-time step.

The FDTD approach proceeds to solve for the electric and magnetic fields resulting from an radiating wave interacting by dividing both time and space into a numerical grids. Fig. 1 illustrates the grid points of the electric and magnetic field components about a unit cell of the FDTD lattice in Cartesian coordinates. It should be noted that each electric field component is surrounded by four circulating magnetic components and vice versa. With each cell, six components of the electric and magnetic fields are computed successively and alternatively at different points in time steps. It is also noted that electric and magnetic fields are calculated by the finite-difference data of the magnetic field about each electric field point and vice versa. The FDTD technique begins the solution of the fields with the assumption that the system is starting from some known condition, must often from an assumption of zero initial fields. Thus it is assumed that the electric (or magnetic) fields are known at each grid point at *t=0*, and the magnetic (or electric) fields are also known at each lattice point at previous half-step $t = -\Delta t/2$. Under this assumption, at $t = \Delta t/2$, the magnetic fields at grid points will be obtained successively. Once the values of the magnetic fields at $t = \Delta t/2$ have been found in each cell of FDTD lattice, then an application of Maxwell's curl equations in central difference form will yield the electric fields at time $t = \Delta t$. These procedures will continue until the desired time-step.

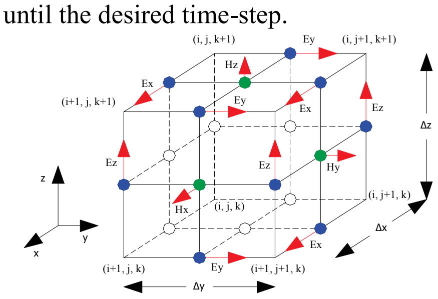


Fig. 1. A FDTD lattice cell.

From the above discussions, by using the numerical theory of finite difference, Maxwell's equations in the curl form of (2) can be written in the following difference forms [3].

$$E_x^n\left(i+\frac{1}{2},j,k\right)=E_x^{n-1}\left(i+\frac{1}{2},j,k\right)+\frac{\Delta t}{\varepsilon}\left[\frac{H_z^{n-\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)-H_z^{n-\frac{1}{2}}\left(i+\frac{1}{2},j-\frac{1}{2},k\right)}{\Delta y}+\frac{H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k-\frac{1}{2}\right)-H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)}{\Delta z}\right] \quad (4a)$$

$$E_y^n\left(i,j+\frac{1}{2},k\right)=E_y^{n-1}\left(i,j+\frac{1}{2},k\right)+\frac{\Delta t}{\varepsilon}\left[\frac{H_x^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)-H_x^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k-\frac{1}{2}\right)}{\Delta z}+\frac{H_z^{n-\frac{1}{2}}\left(i-\frac{1}{2},j+\frac{1}{2},k\right)-H_z^{n-\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)}{\Delta x}\right] \quad (4b)$$

$$E_z^n\left(i,j,k+\frac{1}{2}\right)=E_z^{n-1}\left(i,j,k+\frac{1}{2}\right)+\frac{\Delta t}{\varepsilon}\left[\frac{H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)-H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k-\frac{1}{2}\right)}{\Delta x}+\frac{H_x^{n-\frac{1}{2}}\left(i,j-\frac{1}{2},k+\frac{1}{2}\right)-H_x^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)}{\Delta y}\right] \quad (4c)$$

$$H_x^{n+\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)=H_x^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right)+\frac{\Delta t}{\mu}\left[\frac{E_y^n\left(i,j+\frac{1}{2},k+1\right)-E_y^n\left(i,j+\frac{1}{2},k\right)}{\Delta z}+\frac{E_z^n\left(i,j,k+\frac{1}{2}\right)-E_z^n\left(i,j+1,k+\frac{1}{2}\right)}{\Delta y}\right] \quad (4d)$$

$$H_y^{n+\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)=H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)+\frac{\Delta t}{\mu}\left[\frac{E_z^n\left(i+1,j,k+\frac{1}{2}\right)-E_z^n\left(i,j,k+\frac{1}{2}\right)}{\Delta x}+\frac{E_x^n\left(i+\frac{1}{2},j,k\right)-E_x^n\left(i+\frac{1}{2},j,k+1\right)}{\Delta z}\right] \quad (4e)$$

$$H_z^{n+\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)=H_z^{n-\frac{1}{2}}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)+\frac{\Delta t}{\mu}\left[\frac{E_x^n\left(i+\frac{1}{2},j+1,k\right)-E_x^n\left(i+\frac{1}{2},j,k\right)}{\Delta y}+\frac{E_y^n\left(i,j+\frac{1}{2},k\right)-E_y^n\left(i+1,j+\frac{1}{2},k\right)}{\Delta x}\right] \quad (4f)$$

Equation (4a) is the final FDTD expressions for evaluating the x-axis electric fields at time $t = n\Delta t$ in each cell. From (4) the fields at time $t = n\Delta t$ are obtained entirely in terms of fields earlier times $(n-1/2)\Delta t$ and $(n-1)\Delta t$. Therefore, once the values of the electric fields at $t = n\Delta t$ have been calculated in each cell, the others can be calculated by the similar finite difference equations derived from Maxwell's equations. Space increments Δx, Δy, Δz and time increment Δt, in the previous discussion, are related by accuracy and stability. The increments of Δx, Δy, Δz and Δt must be small enough to ensure the calculating accuracy. However, whose increments could not be too small to guarantee the calculating stability. It is well-known that a stability criterion is given as follows. In order to simulate the computation zone to infinity, before we mentioned that one has to add absorbing boundary surface which can absorb the propagating wave, so that the wave will not be reflected when it propagates to the boundary surface. The absorbing boundary surface is only a virtual surface in order to reach the above purpose. Here, we choose the surface having the following properties [4]:

$$v_{\max}\Delta t \le \left(1/\Delta x^2 + 1/\Delta y^2 + 1/\Delta z^2\right)^{-1/2}, \quad (5)$$

where $v_{\max}$ is the maximum phase velocity expected within the FDTD cells.

$$E_y^n(M) = E_y^{n-1}(M-1) \quad (6)$$

Equation (6) simulates the free-space propagation of the magnitude of E_y from the point M-1 to the truncation point M in one step. From the computed results, the wave is almost absorbed. The errors caused by the finite boundary surface can be neglected.

III. Conclusions

The proposed LDS multiband antenna achieves gain for GSM-850/900 to be about 2.48-3.08 dBi, GPS to be about 1.47-1.8 dBi, DCS-1800 to be about 0.83-3.18 dBi, PCS-1900 to be about 0.83-3.18 dBi, UMTS-2100 to be about 0.83-3.18 dBi, WLAN-2400 to be about 1.42-1.66 dBi, LTE-2500 to be about 1.42-1.66 dBi bandwidth. The proposed 2D pattern is shown in Fig. 3, Fig. 4, Fig. 5, Fig. 6. The proposed multi band LDS antenna for VSWR ≤ 2 can be achieved by means of tuned traps, each trap consisting of an equivalent parallel-tuned LC circuit [5]. Both simulation and measured results are highly matched. The LDS antenna which is economic cost has been proposed and successfully implemented [6].

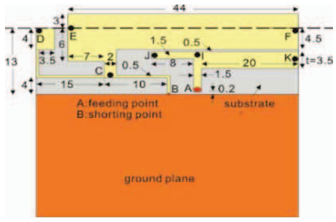


Fig. 2. The proposed multiband antenna architecture.

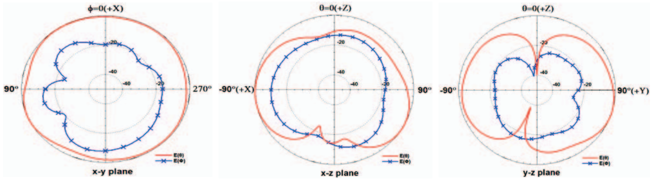


Fig. 3. 925 MHz

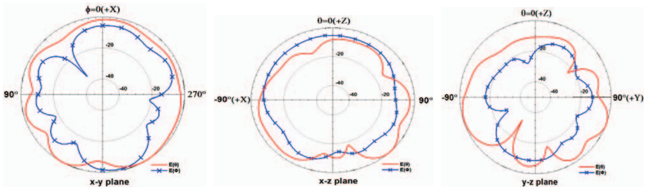


Fig. 4. 1575 MHz

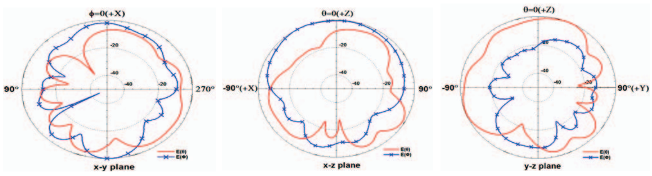


Fig. 5. 1940 MHz

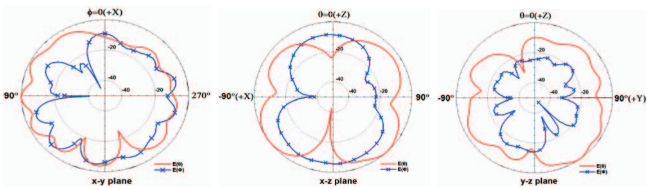


Fig. 6. 2450 MHz

References

[1] K.-S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. on Antennas and Propag.*, p.302-307, May 1966.

[2] Y.-W. Su and J.-F. Huang, FDTD analysis of EMP scattering on a small square conductor coated with a thin dielectric material, *Progress in Electromagnetic Research Symp.*, p. 605. 9, Jan. 1997.

[3] J.-F. Huang, J.-Y. Wen, and S.-C. Lin, "Multiband notebook computer antenna design by FDTD method for GSM WCDMA system applications," *Microwave Opt Technol Lett* 51, 2212–2216, 2009.

[4] A. Taflove and K.-R. Umashankar, "Radar cross section of general three-dimensional scatterers," *IEEE Trans Electromagn Compat* AP33, 662–666, 1985.

[5] W.-C. Lai and J.-F. Huang, "Integrated RFIC On-Chip and GPS Antenna with Human Body for Wrist and Wearable Communication Applications," *Applied Computational Electromagnetics Society Journal* , vol. 31, no. 9, Sep. 2016.

[6] J.-F. Huang, J.-Y. Wen and W.-C. Lai, "Design of a Printed Dipole Array Antenna with Wideband Power Divider and RF Switches," *Microwave Opt Technol Lett.*, vol.55, no. 10, pp. 2410–2413, Jul. 2013.

Frontend with Antenna for EM Body Analysis

Wen-Cheng Lai[1]* and Chang-Chun Zhang[2]
[1]*Dept. of Electrical Engineering, National Penghu University of Science and Technology, Taiwan R.O.C.*
[2]*College of electronic science and engineering, Nanjing University of Posts and Telecommunications, China.*
*Email: wenlai@gms.npu.edu.tw

***Abstract*—The proposed LDS antenna with TR switch and LNA with WuRx circuit design for multiband operation covering the GSM-850/900, GPS, DCS-1800, PCS-1900, UMTS-2100, WLAN-2400, LTE-2500 bands in the 5G communication devices is presented. The proposed multiband antenna is fitting to be applied in the mobile devices with human body for EM analysis.**

Keywords*—*TR Switch, EM analysis, LNA, multiband antenna, WuRx

I. INTRODUCTION

The proposed laser direct structuring (LDS) antennas with a small printed area on the mobile device cover for GSM-850/900, GPS, DCS-1800, PCS-1900, UMTS-2100, WLAN-2400, LTE-2500 bands in the mobile device have also been reported very recently. The proposed multiband antenna merger T/R switch, power amplifier and low noise amplifier design for 5G application as shown Fig. 1. The low power transceiver develop insertion loss and FET based T/R switch is more importance [1][2].

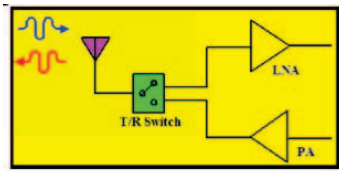


Fig. 1. The proposed transceiver front end design with antenna

II. FDTD PRINCIPLE

Fig. 2 shows an arbitrarily oriented superquadric antenna and a proximate human-body model. The electric parameters of the human-body tissue are expressed as (σ, ε, μ_0). The total EM field in space is the sum of the incident field maintained by multiband antenna and the scattered field from the body. By using the dyadic Green's function technique, the total electric field inside the body and multiband antenna current distribution can be formulated as,

$$\frac{-1}{j\omega\varepsilon_o}\int_{ant}\left[\hat{s}\cdot\hat{s}'k_o^2 I(s')+\frac{\partial I(s')}{\partial s'}\cdot\frac{\partial}{\partial s}\right]G(s,s')ds'-\int_{V_b}\hat{s}\cdot\bar{\bar{G}}(s,\bar{r}')\cdot\tau(\bar{r}')\bar{E}(\bar{r}')dv' = \hat{s}\cdot\bar{E}^i(s) \quad (1)$$

$$\int_{ant} I(s')\hat{s}'\cdot\bar{\bar{G}}(\bar{r},s')ds' + PV\int_{V_b}\tau(\bar{r}')\bar{E}(\bar{r}')\cdot\bar{\bar{G}}(\bar{r},\bar{r}')dv' - \left[1+\frac{\tau(\bar{r})}{3j\omega\varepsilon_o}\right]\bar{E}(\bar{r}) = 0 \quad (2)$$

Where

$$\bar{\bar{G}}(\bar{r},\bar{r}') = -j\omega\mu_o\left[\bar{\bar{I}}+\frac{1}{k_o^2}\nabla\nabla\right]G(\bar{r},\bar{r}'); \quad G(\bar{r},\bar{r}') = \frac{e^{-jk_o|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|}$$

$$\bar{J}_{eq}(\bar{r}) = \{\sigma(\bar{r})+j\omega[\varepsilon(\bar{r})-\varepsilon_o]\}\bar{E}(\bar{r}) = \tau(\bar{r})\bar{E}(\bar{r})$$

Equations (1) and (2) are solved numerically by the MoM. The proposed multiband antenna is partitioned into N_a segments, and the body is subdivided into N_b cubic cells. Pulse-function expansions are used to approximate the unknown multiband antenna current and induced electric field inside the body. $(\Delta\theta)_l$ and $(\Delta v)_n$ represent the lth segment of the multiband antenna and the nth cell in the body, respectively.

$$I(\theta)=\sum_{l=1}^{N_a} I_l p_l(\theta), \quad \bar{E}(\bar{r})=\sum_{n=1}^{N_b}\sum_{p=1}^{3}\hat{x}_p E^n_{x_p} p_n(\bar{r}) \quad x_1=x,\ x_2=y,\ x_3=z \quad (3)$$

I_l = antenna current at the lth partition

$E^n_{x_p}$ = x_p component of induced electric field in the nth subvolume

$$p_l(\theta)=\begin{cases}1, & \theta \text{ in } (\Delta\theta)_l\\ 0, & \text{otherwise}\end{cases} \qquad p_n(\bar{r})=\begin{cases}1, & \bar{r} \text{ in } (\Delta v)_n\\ 0, & \text{otherwise}\end{cases}$$

By applying pulse-function expansion for the unknowns and point matching at the central points of each cell in the body and each segment on the multiband antenna, (1) and (2) can be transformed as

$$\left[\begin{array}{c|ccc} [AA] & [AB^x] & [AB^y] & [AB^z] \\ \hline [BA^x] & [BB^{xx}] & [BB^{xy}] & [BB^{xz}] \\ [BA^y] & [BB^{yx}] & [BB^{yy}] & [BB^{yz}] \\ [BA^z] & [BB^{zx}] & [BB^{zy}] & [BB^{zz}] \end{array}\right]\left[\begin{array}{c} [I_a] \\ \hline [E_b^x] \\ [E_b^y] \\ [E_b^z] \end{array}\right]=\left[\begin{array}{c} [E_\theta^i] \\ \hline [0] \\ [0] \\ [0] \end{array}\right] \quad (4)$$

where [AA] is a Na × Na sub-matrix, [AB] is Na × 3Nb, [BA] is 3Nb × Na, [BB] is 3Nb × 3Nb, and $[I_a]$ and $[E_b]$ are corresponding column vectors of the antenna current and a component of the induced field in the body as Fig. 2.

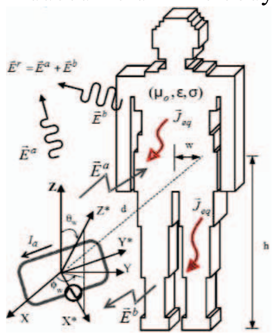


Fig. 2. Illustration of EM coupling between an arbitrarily oriented superquadric loop antenna and a proximate human-body model..

III. CONCLUSIONS

The proposed LDS multiband antenna in Fig. 3 (a) with low noise finger and high linear amplifier for low power consumption receiver. The proposed architecture of single-stage low noise, broadband and series peaking amplifier with Cascode and Chebyshev filter is depicted in Fig. 3(b). Fig. 4 shows the impedance matching design for amplifier circuit. Common-source amplifier circuit input impedance is (5)

$$Z_{in} \approx sL + \frac{1}{sC_{gs}} + \omega_T L \quad (5)$$

$$Z_{in} = sL_b // \left[s\left(L_g+L_s\right)+\frac{1}{sC_{gs}}+\omega_T L_s\right] = \frac{sL_b\left[s^2(L_g+L_s)C_{gs}+sL_sC_{gs}+1\right]}{s^2(L_g+L_s+L_b)C_{gs}+sg_mL_s+1} \quad (6)$$

and target design Z_{in} is 50Ω ; The impedance matching design for wideband amplifier circuit as shown in Fig. 5 (a). Regarding function (6), the Z_{in} parameter of center frequency is leading L_b parallel front-end inductor L and capacitor C series matching. The series-peaking architecture can enhance high-frequency gain to increase bandwidth (7) as shown in Fig. 5 (b). Assume as L_1 increases and wider frequency range, the reported gain will be upgrade path from the RC series feedback structure.

$$v_o(s) = \frac{1}{sC} // (sL_1 + sL_2) \cdot \frac{sL_2}{sL_1 + sL_2} \cdot I_{in} = \frac{sL_2}{s^2 C (L_1 + L_2) + 1} \cdot I_{in} \quad (7)$$

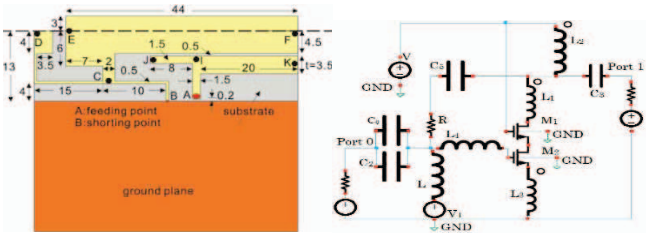
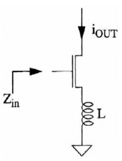


Fig. 3. (a) the proposed multiband antenna architecture (b) the proposed LNA (low noise amplifier)

The proposed T/R-switch circuit as shown in Fig. 6. Due to ultra-low-power WuRx's have been demonstrated with low consumption. The block diagram presents single-stage low noise amplifier in Fig. 7. The proposed WuRx (wake up receiver) circuit design for the proposed RF detector embedded with input matching network, depicted in Fig 8, directly converts the amplitude modulated RF signal from a 50 ohm antenna to baseband. To address this solution proposed a unique protocol, which adopts a duty cycle scheduling and radio wake-up identification circuit in Fig. 9 [3][4][5].



$$Z_{in} = sL + \frac{1}{sC_{gs}} + \frac{g_m}{C_{gs}} L \approx sL + \frac{1}{sC_{gs}} + \omega_T L$$

Fig. 4. The impedance matching design for MOS amplifier circuit

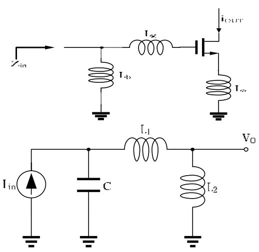


Fig. 5. (a) Wideband amplifier circuit impedance matching (b) Series-Peaking amplifiers

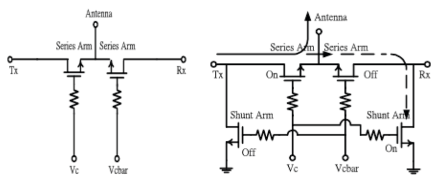


Fig. 6. T/R Switch architecture diagram and schematic

T/R Switch characteristics table

T/R Switch: Characteristics	
Control Voltage	1.8V
Insertion loss	<1.0dB
Input return loss (Tx and Rx modes)	>12dB
Input P_{1dB} (Tx and Rx modes)	>21dBm
Isolation (Tx-Rx in Tx mode)	>24dB
Isolation (ANT-Tx in Rx mode)	>18dB

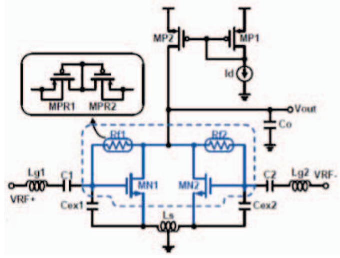


Fig. 7. Circuit of RF detector

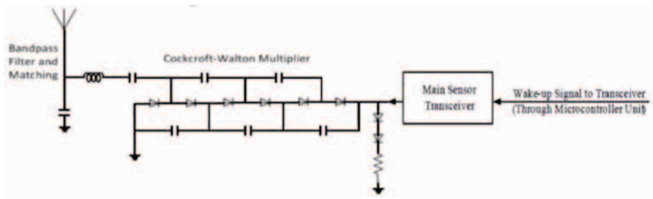


Fig. 8. Block diagram of wake-up receiver front end

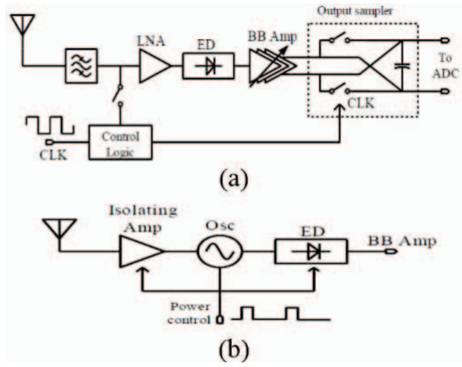


Fig. 9. (a) A WuRx with double sampling (b) super-regenerative receiver block diagram.

References

[1] W.-C. Lai and J.-F. Huang, "Integrated RFIC On-Chip and GPS Antenna with Human Body for Wrist and Wearable Communication Applications," *Applied Computational Electromagnetics Society Journal* , vol. 31, no. 9, Sep. 2016.

[2] J.-F. Huang, J.-Y. Wen and W.-C. Lai, "Design of a Printed Dipole Array Antenna with Wideband Power Divider and RF Switches," *Microwave Opt Technol Lett.*, vol.55, no. 10, pp. 2410–2413, Jul. 2013.

[3] X. Huang, S. Rampu, X. Wang et al., "A 2.4GHz/915MHz 51 ʃW wake-up receiver with offset and noise suppression," in IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers, 2010, pp. 222-223.

[4] J. Choi, K. Lee, S.-O. Yun et al., "An interference-aware 5.8GHz wake-up radio for ETCS," *in IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers,* 2012, pp. 446-448.

[5] C. Hambeck, S. Mahlknecht, and T. Herndl, "A 2.4uW Wake-up Receiver for wireless sensor nodes with -71dBm sensitivity," *in IEEE Int. Symp. on Circuits and Systems*, 2011, pp. 534-537.

Preliminary Design of a Waveguide-Fed Milimeter Wave Metasurface Antenna with LCD Controlled Array Factor for 5G User Equipment

Shen Shou Max Chung
Department of Electrical Engineering
National Penghu University of Science and Technology
Penghu, Taiwan, R.O.C.
maxchung@ms3.hinet.net

Shih-Chung Tuan
Department of Communication Engineering
Oriental Institute of Technology
Taipei, Taiwan, R.O.C.
fo012@mail.oit.edu.tw

***Abstract*—The millimeter wave (mmW) band in 5G communication promises wide bandwidth for high data rate throughput, however the available area on cellular user equipment (UE) is quite limited, and 2 dimensional phased array occupies a large area and is fairly complicated. In this article we explore the possibility a waveguide-fed millimeter wave metasurface antenna for 28 GHz band. The element is a generic T-metasurface, and LCD control of array factor is suggested with a glass lens to focus the beam. A fan shape beam is generated from this antenna with maximum gain around 16.7 dB. This design may serve the 5G mmW communication with simpler structure to nearby small cell station.**

I. Introduction

5G communication [1] promises 10 Gbps data rate with 100 MHz bandwidth carrier aggregation, and the main new feature that enable this is the mmW communication [2]. Due to the nature of high path loss of mmW, current concept for the assigned 28 GHz band may be a small cell (< 200 m) backhaul cellular network [3] for outdoor mobile communication in which traffic light or lamp post play important role as base station, and 802.11ax (Wi-Fi 6) with its 160 MHz channel shall take care of the indoor mobile communication [4] with router connected to fiber optic or cable backhaul. In terms of antenna technology, massive MIMO technology [5] with hybrid beam forming complies naturally with base station, and 28 GHz transceiver driven phased array is proposed for base station (BS) and user equipment (UE) as well [6].

However, is it really necessary to use two-dimensional phased array on UE? A Two-dimensional phased array at 28 GHz requires substantial antenna area ($\lambda/2 \sim 5$ mm, 8x8 means 25 mm^2) in order to form a pencil beam, and array control is quite complicated even for hybrid beam forming. Phased array usually requires Wilkinson power dividers and phase shifter networks, the former occupy some die area and the latter tends to be quite complicated and requires real time calibration at the output end. Since UE in street are usually surrounded by multiple of traffic lights and lamp posts where BS can be installed, the communication between UE and BS can be accomplished in a rectangular box space, a simple fan beam can locate the BS and commence download and upload. In this article we investigate the feasibility of using a waveguide-fed milimeter wave metasurface antenna with LCD controlled array factor for 5G UE. Waveguide-fed antenna is widely used in military applications; it has high gain and has only one feed point, which is simple. Meta-surface antennas [7,8] are basically combinations of inductors and capacitors of various shapes, their diversity add freedom to antenna design, and the metasurface stores energy which leads to a smaller waveguide. LCD can control the amplitude of each element to perform beam forming [9] and it is a much simpler mechanism than the phase shifter controlled phased array, plus there is already LCD or OLED control chip inside UE.

II. The Waveguide-fed Milimeter Wave Metasurface Antenna Concept

Shown in Fig. 1 is the possible scenario for outdoor 5G mmW communication where massive MIMO phased arrays are installed on lamp posts and traffic lights to serve as base station. For the UE on a person or a vehicle, it only needs to produce a fan shape beam in order to communicate with the BSs which are within a box of rectangular space with limited height. Fig. 2 shows the explosive view of the proposed waveguide-fed metasurface LCD controlled antenna. On the top is a cylindrical glass lens to focus and narrow down the beam. Below it is the LCD array that is used to control each element's amplitude to perform beam forming. Below the LCD is the bottom glass, where the metasurface is deposited on its lower surface. Next are the metasurface elements, there are total 17 of them and each is separated at $\lambda/2$. Below them is the waveguide, which is fed from the right. Fig. 3 is the 3D gain radiation pattern of the metasurface antenna with zero and even array factors, the maximum gain is quite high (~ 16.7 dBi). Fig. 4 is the Phi=0° and Phi=90° cut gain pattern. Table I lists all the relevant external sizes of the antenna.

III. Summary

In this article we present the preliminary design of a 28 GHz waveguide-fed LCD controlled metasurface antenna that we think is smaller than an 8x8 phased array antenna and simpler to make and control but still capable of beam forming for 5G communication. The metasurface element size is 4x4 mm^2, and the minimum feature is 0.5 mm. The size of the antenna can be reduced if microstrip is used instead of waveguide; however waveguide transmission loss is lower. A typical metasurface pattern is used here, but other patterns are also possible. LCD amplitude control of electromagnetic wave is a fairly mature technology.

TABLE I: Relevant Sizes FOR the Metasurface LCD Controlled Antenna

Frequency	Total Length	Height	Width	Waveguide
28 GHz	96.36 mm	10.19 mm	8.37mm	WR34

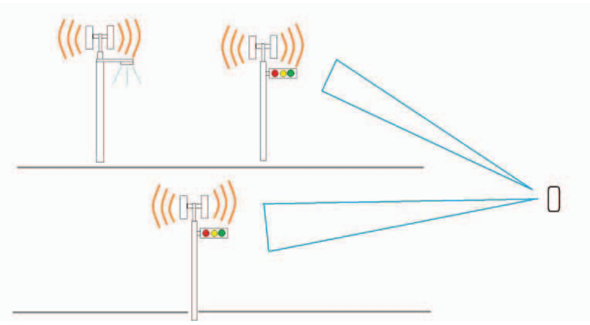


Fig. 1. Possible scenario for outdoor 5G mmW communication where massive MIMO phased arrays are installed on lamp posts and traffic lights to serve as base station.

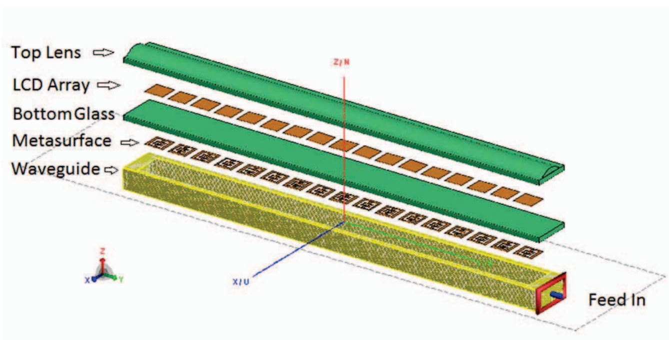


Fig. 2. Explosive view of the proposed waveguide-fed metasurface LCD controlled antenna. On the top is a cylindrical glass lens to focus and narrow the beam. Below it is the LCD array that is used to control each element's amplitude to perform beam forming. Below the LCD is the bottom glass, where the metasurface is deposited on its lower surface. Next are the metasurface elements, total 17 of them. Below them is the waveguide, which is fed from the right.

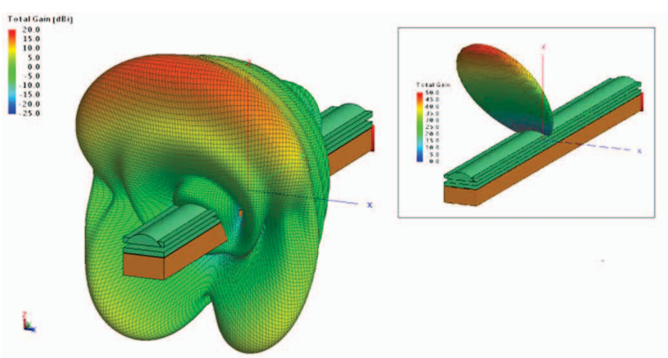


Fig. 3. The 3D beam shape of the metasurface antenna with zero array factors, the maximum gain is quite high (~ 16.7 dBi). Inset is the gain in linear scale.

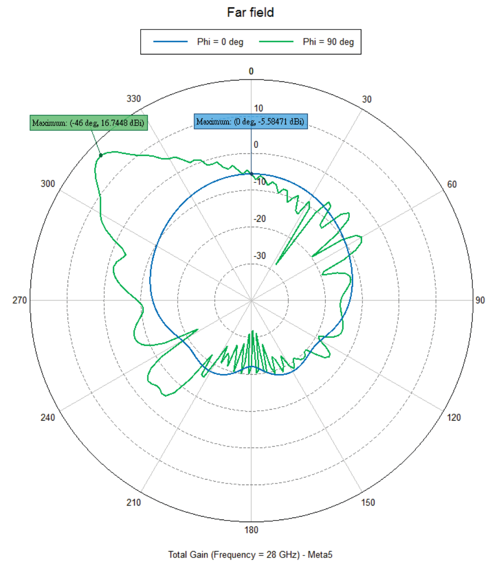


Fig. 4. The Phi=0° and Phi=90° cut gain pattern. The maximum gain in YZ plane (green line) is 16.74 dB, maximum gain in XZ plane is -5.58 dB.

ACKNOWLEDGEMENT

The authors would like to express gratitude towards Altair for the academic license of FEKO. This work was supported in part by the Ministry of Science and Technology of Taiwan, R.O.C., under project MOST 107-2221-E-161-001.

REFERENCES

[1] J. G. Andrews *et al*., "What will 5G be?", IEEE J. Sel. Areas Commun., vol. 32, no.6, pp. 1065-1082, June 2014.

[2] T. S. Rappaport *et al.,* "Millimeter-wave mobile communication for 5G celluar: It will work!," IEEE Access, pp. 335-349, 2013.

[3] Wei Feng, Yong Li, Depeng Jin, Li Su and Sheng Chen, "Millimetre-wave backhaul for 5G networks: Challenges and solutions," Sensors, vol. 16, pp. 892, 2016.

[4] Evgeny Khorov, Anton Kiryanov, Andrey Lyakhov, and Giuseppe Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," IEEE Comm. Survey & Tutorials, vol. 21, no. 1, pp. 197-216, 2019.

[5] Xiang Gao, Ove Edfors, Fredrik Rusek, and Fredrik Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," IEEE Trans. on Wireless Comm., vol. 14, no. 7,pp. 3899-3911, July 2015.

[6] Jeremy Dunworth *et al., "*28GHz Phased Array Transceiver in 28nm Bulk CMOS for 5G Prototype User Equipment and Base Stations," 2018 IEEE/MTT-S Int. Microwave Symp.

[7] David R. Smith *et al*., "Analysis of a waveguide-fed metasurface antenna," arXiv:1711.01448v1 [physics.app-ph] 4 Nov 2017.

[8] Aobo Li, Shreya Singh and Dan Sievenpiper, "Metasurfaces and their applications," Nanophotonics, https://doi.org/10.1515/nanoph-2017-0120.

[9] Felix Gölden, "Liquid crystal based microwave components with fast response times: material, technology, power handling capability," PhD dissertation, Elektrotechnik und Informationstechnik der Technischen Universität Darmstadt, 2009.

Duty Ratio and Capacitance Analysis of AC/DC Converter without Current Control Circuit

Yuki Tanaka
Department of Electrical Engineering
Graduate School of Science and Technology
Tokyo University of Science
Noda-shi, Chiba, Japan

Tatsuji Matsuura, Ryo Kishida, Akira Hyogo
Department of Electrical Engineering
Faculty of Science and Technology
Tokyo University of Science
Noda-shi, Chiba, Japan

***Abstract*—This paper analyzes duty ratio and capacitance of AC/DC converter without Power Factor Correction (PFC) circuit. This converter adds a capacitor in order to flow current through a diode of the boost converter every switching cycle. The input current flows every switching cycle without current control and follows input voltage. We calculate circuit equations of the AC/DC converter and simulate using proper duty ratio and capacitance. As simulation results, this circuit convert input voltage and current to constant 12 V and 125 A. We demonstrate the equation and calculated parameters are useful in the AC/DC converter.**

Keywords—AC/DC Converter, PFC converter, power factor

I. Introduction

AC/DC converter converts alternating current (AC) to direct current (DC) voltage. The general AC/DC converter mainly consists of three circuits that are a rectifier, Power Factor Correction (PFC), and an isolated DC/DC (LLC) converter. The AC/DC converter requires cost reduction and small circuit area. To meet these demands, The AC/DC converter using a capacitor instead of current control is proposed[1]. This circuit can reduce the number of components and circuit area by removing current control circuits.

We analyze circuit equations and find proper duty ratio and capacitance in the AC/DC converter. Section 2 describes the AC/DC converter using a capacitor instead of current control. Section 3 describes the analysis of the circuit equations. Section 4 shows the simulation results. In the simulation, simulation condition is that from 90 to 264 Vac voltage, 50 Hz frequency,12 Vdc output voltage and 125 A output current assuming the server in data center. Finally, section 5 shows our conclusion.

II. AC/DC Converter Using a Capacitor Instead of Current Control

AC/DC converter proposed in [1] is shown in Fig. 1, however the circuit topology is changed for high power. AC/DC converter mainly consists of three circuits that are a rectifier, Power Factor Correction (PFC), and an isolated DC/DC converter. Point A is the intersection of the SW_1 and inductor L. Point B is the intersection of the SW_2 and SW_3. This circuit connects these points using a capacitor C_a. V_{cin} is output voltage of the boost circuit, and V_F is the voltage across the diode. The key timing is when SW_1 turns off and SW_3, SW_4 turns on. At this state, the inductor current initially flows to C_a because the diode dose not operate. Next, the voltage across to C_a (V_{Ca}) increases. When V_{Ca} is larger than $V_{cin}+V_F$, the diode operates and the inductor current I_L flows through diode current I_D. The I_L flows every switching cycle of the boost converter. This current is determined by the voltage across the inductor. The output voltage of the rectifier is applied to the inductor. By flowing the inductor current as the input current every switching cycle, the same frequency component of input voltage in input current increases. Therefore, it is possible to improve the power factor.

III. Analysis of Circuit Equations

This section shows design procedures of the AC/DC converter without current control circuit. First, we design commonly used boost converter and LLC converter respectively[2][3]. This design determines component values, D_b and D_l, where D_b and D_l are the duty ratio of the boost converter and LLC converter respectively. Second, we design C_a in two cases ($D_b < D_l$, $D_b > D_l$). In both cases, we consider inductor current I_L, and charge and discharge of capacitor in boost circuit Q in steady state.

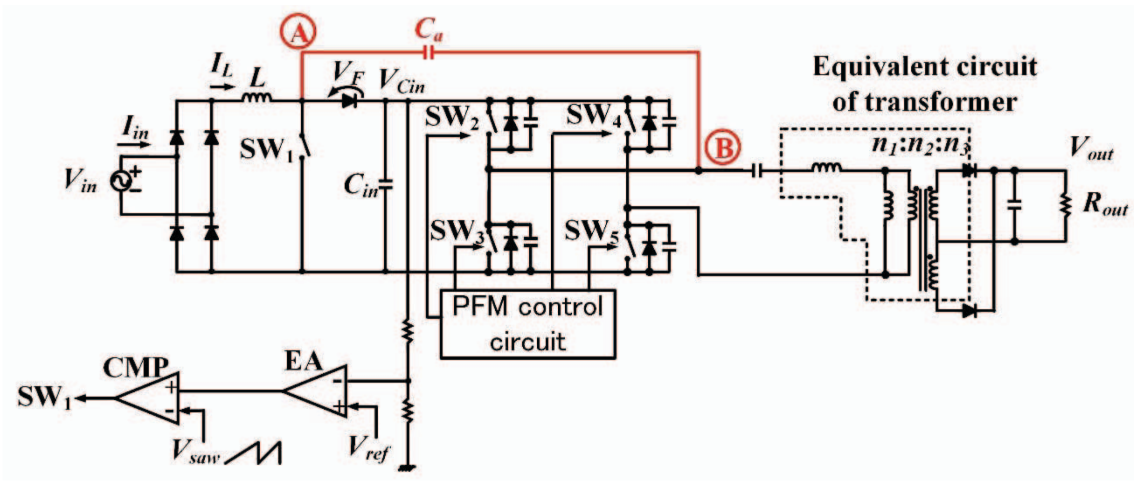


Fig. 1. AC/DC converter using a capacitor instead of current control [1]. C_a is connected between point A and B. C_a is charged by the inductor current I_L. When the voltage across C_a exceeds $V_{cin}+V_F$, the diode of the boost circuit turns on. The input current flows and improves power factor.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

In steady state, inductor current and charge / discharge on C_{in} are established following (1) and (2).

$$I_L(T) = I_L(0) \quad (1)$$

$$Q(+) = Q(-) \quad (2)$$

Where T is switching period of boost circuit and $Q(+)$ and $Q(-)$ are charge and discharge of the capacitor respectively. It is important to calculate at $V_{Ca}=V_{cin}+V_F$ because I_L flows through diode when the charge of the C_a is filled.

A. $D_b>D_l$

V_{Ca} becomes $V_{cin}+V_F$ at time t_1 when SW_3 and SW_4 turn on, and SW_1 turns off. V_{Ca} is followed by (3),

$$V_{Ca}(t) = V_{in}\left\{1-\cos\left(\frac{t-D_bT}{\sqrt{LC_a}}\right)\right\} + \sqrt{\frac{L}{C_a}}\left(\frac{V_{in}}{L}D_bT+I_L(0)\right)\sin\left(\frac{t-D_bT}{\sqrt{LC_a}}\right) \quad (3)$$

where L is an inductor of the boost converter. We can design C_a to use (1)-(3).

B. $D_b<D_l$

V_{Ca} becomes $V_{cin}+V_F$ at time t_1 and also t_2. t_1 is a time when SW_2 and SW_5 turn on, and SW_1 turns off, and t_2 is a time when SW_3 and SW_4 turn on, and SW_1 turns off.

We are required two times, and V_{Ca} is followed by (4) and (5),

(for t_1) $$V_{Ca}(t) = -V_{in}\left\{\cos\left(\frac{t-D_bT}{\sqrt{LC_a}}\right)-1\right\} + \frac{1}{C_a}\sqrt{\frac{L}{C_a}}\left(\frac{V_{in}}{L}D_bT+I_L(0)\right)\sin\left(\frac{t-D_bT}{\sqrt{LC_a}}\right) \quad (4)$$

(for t_2) $$V_{Ca}(t) = -\{V_{in}+V_F(1-C_a)\}\left\{\cos\left(\frac{t-D_lT}{\sqrt{LC_a}}\right)-1\right\} + \{(V_{in}-V_{cin}-V_F)(D_lT-t_1)+LI_L(t_1)\}\sqrt{LC_a}\sin\left(\frac{t-D_lT}{\sqrt{LC_a}}\right) \quad (5)$$

Finally, we need to calculate D_b using C_a and (1)-(5) again.

IV. SIMULATION

The validity of the equation and operation of the circuit is confirmed by circuit simulation using the high-speed circuit simulator SCALE. Assuming the AC/DC converter for server in data center, the input is 50 Hz 100 Vac voltage and the output voltage and current are 12 Vdc and 125 A, respectively. By calculation, C_a is determined 60 nF and D_b is 0.67.

The simulation results of the input and output voltage and the input current are shown in Fig. 2. The output voltage are constant 12 Vdc, and input current follows input voltage.

Fig. 3 (a)-(c) show voltage and current waveforms. When SW_1 turns off and SW_3 and SW_4 turn on, V_{Ca} increases. Then, the diode turns on and current flows. The AC/DC converter operates correctly and analyzed C_a and D_b are appropriate.

V. CONCLUSION

The AC/DC converter adding a capacitor becomes smaller area because current control circuit and components are removed. We analyze circuit equations of the AC/DC converter and calculate the value of C_a. As simulation results, AC/DC converter operates properly, and it demonstrates the calculated parameters are appropriate.

REFERENCES

[1] Y. Tanaka, T. Matsuura, R. Kishida, and A. Hyogo, "A Worldwide Input Voltage AC/DC Converter Using a Capacitor Instead of Current Control,"Analog VLSI Circuits (AVIC), Oct, 2019.

[2] M K. kazimierczuk, "Pulse-width Modulated DC-DC Power Converters," WILEYH, pp.85-137, 2008

[3] S. Abdel-Rahman, "Resonant LLC Converter: Operation and Design," AN 2012-09, Infineon, Application note, 2012.

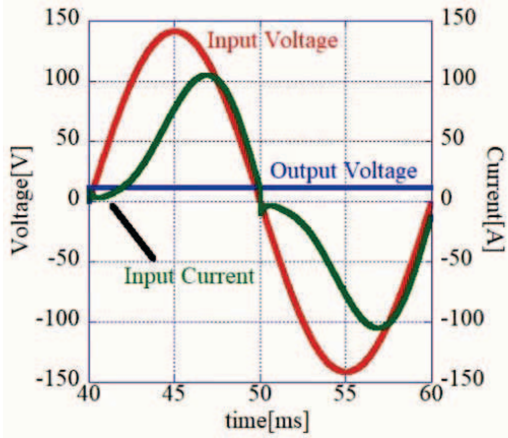
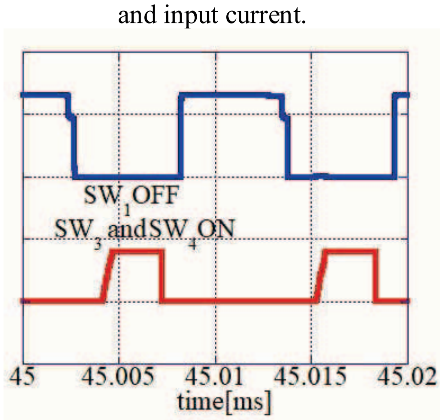
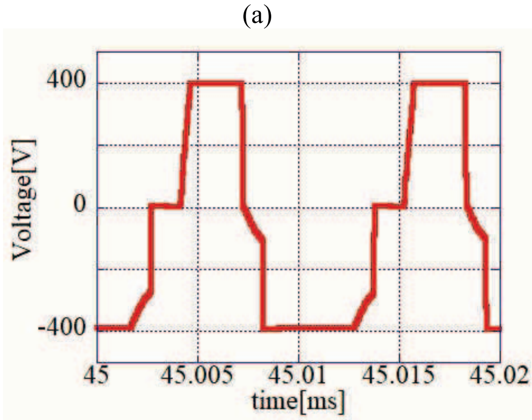


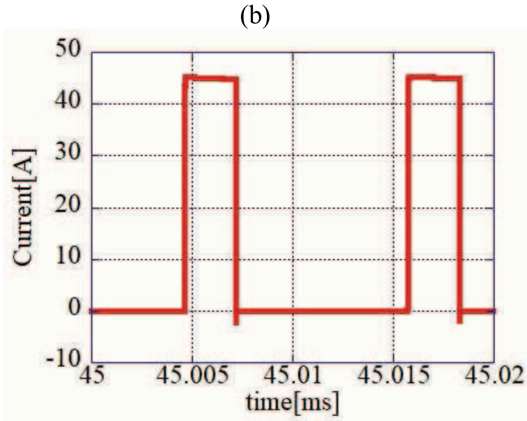
Fig.2. Simulation results of input and output voltage, and input current.



(a)



(b)



(c)

Fig. 3. Simulation results of the circuit operation. (a) Control voltage of SW_1 and SW_3, SW_4. (b) Voltage across C_a increases when SW_1 turns off and SW_3 and SW_4 turns on. (c) I_D flows when voltage across C_a exceeds $V_{cin}+V_F$.

Frontend Design for FMCW MIMO Radar Sensor

Wen-Cheng Lai*, Wen-Hao Chen, Ting-Jung Shih and Ze-Sheng Hong
Dept. of Electrical Engineering, National Penghu University of Science and Technology, Taiwan R.O.C.
*Email: wenlai@gms.npu.edu.tw

***Abstract*—The proposed article presents an integrated frontend design for FMCW (Frequency-modulated continuous-wave) MIMO radar sensor that can be used for 3D imaging purposes applications is presented. By providing transceiver bands the MIMO principle for data reconstruction has been realized. In this publication, a 3-phase class-C current reused VCO can generate signal with 3 output phases and it uses three identical single-ended current reused VCOs in a ring configuration. A novel Gm–C loop filter instead of a conventional passive loop filter used in PLL with divider by 8 ILFD. The innovative advantage of the proposed architecture is tunable loop filter bandwidth.**

Keywords—VCO, PLL, ILFD, Gm–C loop filter

I. Introduction

The transmitter circuit in the radar system frontend is shown in Fig 1. Besides a fundamental VCO (voltage controlled oscillator), this integrated circuit additionally contains a divide-by-8 frequency divider so it can be stabilized with a commercial off the-shelf PLL design. The proposed transmitter circuits are typically measured over large distances in case of a FMCW MIMO (multiple-input and multiple-output) antenna topology. The performance measured 3D sensing by FMCW MIMO are better than optical sensing for depth camera.

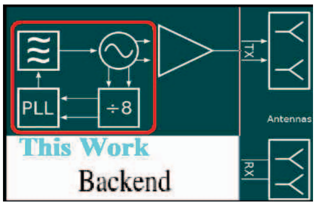


Fig. 1. The proposed transmitter circuit in the radar system frontend.

II. Circuits Design

A combination of active coupling and triple-push techniques is used in the presented VCOs [1]. Alternatively, a 3P VCO can be formed by using three single-ended oscillators coupled by the triple-push signal [2] or coupled by a closed loop composite transmission-line LC network [3]. Fig. 2(a) shows a current reused sub-VCO used in the proposed VCO. In conjunction with a coupling network, three sub-VCOs as shown in Fig. 2(a) are modified to form a 3P VCO. The proposed 3-phase current-reused VCO is shown in Fig. 2(b) consisted of three sub-VCOs connected in a ring architecture. The active ring coupling approach introduces noise and power consumption, however it leads to robust phase accuracy. The sub-VCOs are the same as Fig. 2(a), except the bodies of pMOSFETs are used for coupling [4]. Fig. 3 shows a phase-locked loop (PLL) with a divide-by-4 and a divide-by-2 ILFD. The two ILFDs (injection locked frequency divider) perform the divide-by-8 function, which also can be fulfilled by using three ÷2 LC ILFDs [5]. Three ÷2 current mode logic (CML) flip-flop-based dividers can be used due to their wide locking range [6], however they consume considerable power to operate at higher frequency. Three-step mixing is used in the ÷8 ring-oscillator ILFD [7], this is done through a 3step divide-by-2 down conversion.

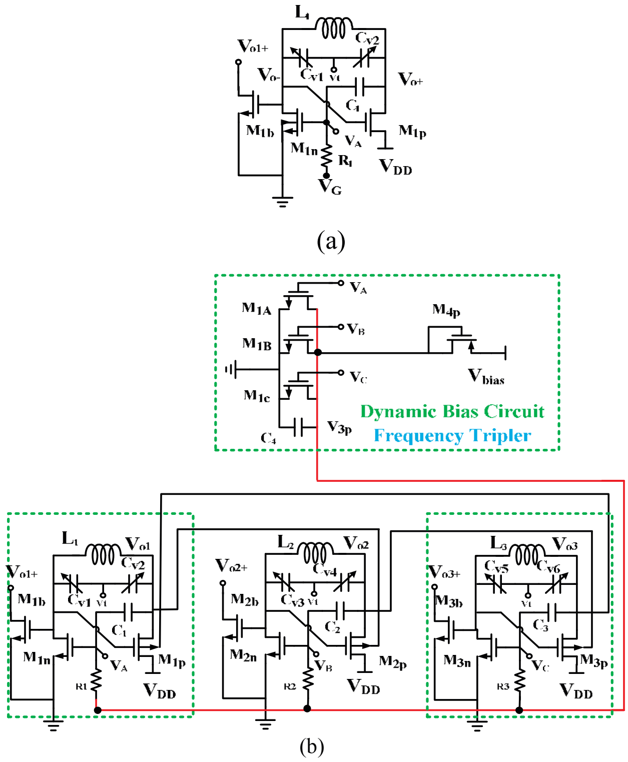


Fig. 2. Schematic of a current reused VCO adapted from [9]. (b) Schematic of the proposed 3-phase VCO.

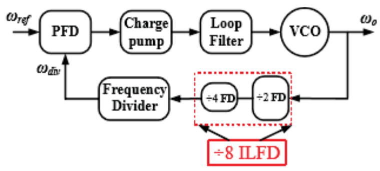


Fig. 3. Block diagram of PLL with a ÷8 ILFD.

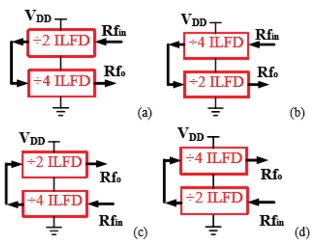


Fig. 4. Block diagram of current –reused divide-by-8 ILFDs. Rfin: input signal. Rfo output signal.

The regenerative ÷8 frequency divider (FD) [8] has wide locking range but suffers from large power consumption because it uses three CML ÷2 FDs, one LC mixer and 4 quadruplers. Two ÷4 LC ILFDs have been developed so far, the harmonic mixer ILFD has narrow locking range, while linear mixer ÷4 ILFD [9] has wide locking range. There are many ways to design a ÷8 current-reused ILFD, for example, the circuit as shown in Fig. 4(a) uses a high frequency ÷2

ILFD followed by a low frequency ÷4 ILFD. Fig. 4(b) shows a ÷8 ILFD with a high-frequency ÷4 ILFD followed by a low frequency ÷2 ILFD. This proposed ÷8 ILFD uses the topology shown in Fig. 4(c). Fig. 4(d) is counterpart of Fig. 4(b), it uses the ILFD connected to the ground as the input. The ÷8 ILFDs use the current reused architecture to save power. For the design shown in Fig. 4(c), there are still many options, for example the ÷2 ILFD can be implemented either with n-core cross-coupled oscillator or with p-core oscillator [11]-[13]. Fig. 6 illustrates in the proposed implementation. the souring and sinking devices M_{12} 、 M_{23} are independent current sources, and the switch M_6、M_7、M_{18}、M_{19} are used to enable UP/DN charge to the LPF. To reduce current mismatch and long channel modulation effect, replica mirror has been employed. The gain-boosted technique is consists of M_4、M_5、M_{10}、M_{16}、M_{17}、M_{22}.

$$i_d = (-1-A) \times gmv_s \tag{1}$$

where A in Fig.5(b), which is composed of M_4 and M_{21}, is an amplifier with gain that doesn't proposed any extra power consumption, then the CP can get more gain of (1+A) under the same power consumption. Fig. 6 shows the current mismatch of the gain boosted charge pump.

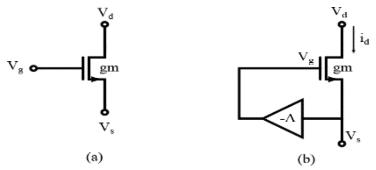


Fig. 5. Schematic of proposed Gm-Boosting technique (a) conventional amplifier circuit (b) Gm-Boosting amplifier circuit.

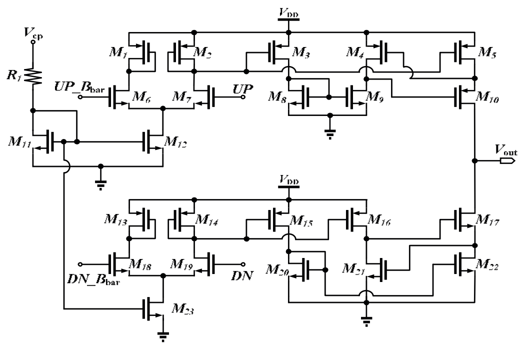


Fig. 6. Proposed gain boosted charge pump

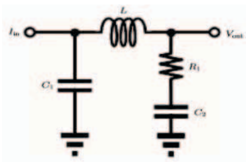


Fig. 7. Passive 3rd order loop filter circuit

It is well known that loop filter is one of the most important building blocks in PLL and the circuit noise performance and locking behavior are a • ected by loop bandwidth. The general 3rd order passive loop filter is chosen and illustrated in Fig. 7. The inductor can be emulated by Gm cells and a capacitor.

III. Conclusions

The proposed article can serve as an integrated frontend design for FMCW MIMO radar sensor that can be used for 3D imaging purposes as shown in Fig. 8 for depth sensing applications.

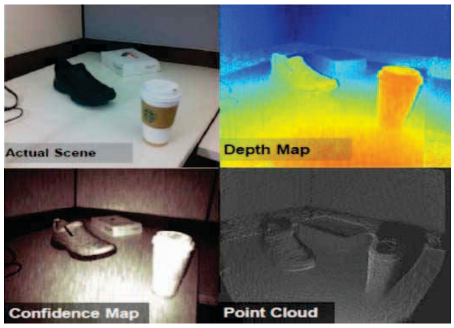


Fig. 8. 3D imaging for depth sensing applications

References

[1] S.-L. Jang," A three-phase PMOS voltage-controlled oscillator using ring and triple-push coupling,"*Microw. Opt. Technol. Lett.*vol. 57, no. 11, pp.2529-2532, 2015.

[2] S.-L. Jang, Y.-S. Lin, C.-W. Chang, and M.-H. Juang, " A three-phase voltage-controlled oscillator using a composite LC transmission-line resonator," *Progress In Electromagnetics Research Letters*, Vol. 27, pp.151–160, 2011.

[3] S.-L. Jang, T.-Y. Cheng, C.-W. Chang, and C.-W. Hsue," A three-phase complementary Colpitts VCO implemented with a LC-ring resonator," *Microw. Opt. Technol. Lett.*vol. 53, issue 10, pp.2308-2310, Oct. 2011.

[4] W.-C. Lai, S.-L. Jang, Y.-Y. Liu and M.-H. Juang, "A Triple-Band Voltage-Controlled Oscillator Using Two Shunt Right-Handed 4th - Order Resonators," *Journal of Semiconductor Technology and Science*, vol. 16, no. 4, pp.1961-1964, Aug. 2016.

[5] S. Ann, J. Yu, J. Park, Y. Kim, and N. Kim ," Low power CMOS 8:1 injection-locked frequency divider with LC cross-coupled oscillator ," *in IEEE European Modelling Symposium (EMS)*, 2015.

[6] R. Shu, V. Subramanian, and G. Boeck, "A 8:1 static frequency divider operating up to 34 GHz in 0.13- um CMOS technology," *in IEEE MTT-S Int. Microw. Symp. Dig.*, pp. 17–20, Sep. 2011.

[7] A. Musa, K. Okada and A. Matsuzawa, "Progressive mixing technique to widen the locking range of high division-ratio injection-locked frequency dividers," *IEEE Trans. Microw. Theory Techn,* vol. 61, no. 3, pp. 1161-1173, Mar. 2013.

[8] Y.-S. Lin, W.-H. Huang, C.-L. Lu, and Y.-H. Wang, "Widelocking-range multi-phase-outputs regenerative frequency dividers using even-harmonic mixers and CML loop dividers," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 12, pp. 3065-3075, Aug. 2014.

[9] S.-L. Jang, T.-C. Kung and C.-W. Hsue," Wide-locking range divide-by-4 injection-locked frequency divider using linear mixer approach," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 4, pp. 398–400, April 2017.

[10] S.-L. Jang, W.-C. Lai, G.-Y. Lin, and C.-Y. Huang, "Injection-Locked Frequency Divider With a Resistively Distributed Resonator for Wide Locking Range Performance," *IEEE Trans. Microw. Theory Techn.*, vol. 67, issue: 2, p.p.505-517, Feb. 2019.

[11] S.-L. Jang, C.-T. Hung, Y.-R. Huang, W.-C. Lai and M.-H. Juang, "Wide-Band Varactorless Dual-Resonance Divide-by-4 Injection-Locked Frequency Divider," *Microw. and Opt. Techn. Lett.*, vol. 59, issue 7, pp.1503-1507, Jul. 2017.

[12] S.-L. Jang, Y.-Y. Liu, C.-H. Fang, W.-C. Lai and M.-H. Juang, "Divide-by-2 LC Injection-Locked Frequency Divider with Wide Locking Range at Low and High Injection Powers," *International Journal of Circuit Theory and Application,* vol. 44, no.12, Apr. 2016.

[13] S.-L. Jang, Y.-J. Chen, C.-H. Fang and W.-C. Lai, "Enhanced Locking Range Technique for Frequency Divider Using Dual-Resonance RLC Resonator," *Electronics Lett.*, pp.1-2, vol. 51, no. 23, 15 Nov. 2015.

Deep Wavelets for Heart Sound Classification

Kun Qian*, Zhao Ren†, Fengquan Dong‡, Wen-Hsing Lai§, Björn W. Schuller†¶, and Yoshiharu Yamamoto*
*Educational Physiology Laboratory, Graduate School of Education, The University of Tokyo, Japan
†ZD.B Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Germany
‡Department of Cardiology, Shenzhen University General Hospital, P. R. China
§Department of Computer and Communication Engineering,
National Kaohsiung University of Science and Technology, Taiwan
¶GLAM – Group on Language, Audio & Music, Department of Computing, Imperial College London, UK
Email: qian@p.u-tokyo.ac.jp, zhao.ren@informatik.uni-augsburg.de

***Abstract*—Cardiovascular diseases have a high morbidity, and remain the leading cause of mortality. In the past two decades, developing an intelligent auscultation system has attracted tremendous efforts from the field of signal processing and machine learning. We propose a novel framework based on wavelet representations and deep recurrent neural networks for recognising three heart sounds, i. e., *normal*, *mild*, and *severe*. The Heart Sounds Shenzhen corpus ($n = 170$) is used to validate the proposed method. The experimental results demonstrate the efficacy of the proposed method in a rigorous subject independent scenario, which can reach an unweighted average recall at 43.0 % (chance level: 33.3 %).**

***Index Terms*—Heart Sound, Healthcare, Cardiology, Wavelets, Deep Learning**

I. Introduction

Cardiovascular diseases (CVD), cause 45 % of all deaths in Europe annually [1]. On one hand, as a simple, convenient, and less-expensive method, auscultation is widely used in clinical and medical practice. On the other hand, physicians need tremendous and extensive training for gaining experiences and skills in auscultation [2]. In addition, it was reported that, only approximately 20 % of the medical interns on average can make an efficient use of the stethoscope to measure a subject's heart status [3]. In the past two decades, plenty of endeavours were made in the field of developing an intelligent auscultation system, which can facilitate an automatic analysis of the heart sound for monitoring the health status of the subject [4]. From a recent literature survey [4] on heart sound classification we can see that, the previous work has achieved encouraging and promising results, which demonstrated the feasibility of using state-of-the-art techniques from signal processing and machine learning to automatically monitor a subject's health status from the heart sound, or the respective Phonocardiogram (PCG). However, there are still some limitations among the existing work: Firstly, publicly accessible heart sound databases are extremely limited. The PhysioNet Cinc Challenge database [5] is the biggest one currently. Nevertheless, this database was combined by multiple medical centres, which share inconsistent data acquisition, annotation, and pre-processing methods. Secondly, most of the previous work ignored subject independence, which renders results overoptimistic. Thirdly, advanced signal processing methods, e. g., wavelets combined with deep learning methods were not comprehensively studied. Therefore, we propose a novel framework based on wavelet representations and deep recurrent neural networks (DRNNs) for the task of heart sound classification. To the best of our knowledge, it is the first time to investigate the capacity of combining multi-resolution analysis and deep sequential learning for heart sound classification. In addition, the experiments were implemented in a publicly accessible database, i. e., the heart sounds Shenzhen (HSS) corpus, which makes this study reproducible and comparable. The remainder of this paper will be organised as follows: First, a relation to prior work will be given in Section II. Then, we will introduce the database and methods in Section III. Subsequently, the experimental results are given in Section IV. Finally, we conclude the study in Section V.

This work was partially supported by the Zhejiang Lab's International Talent Fund for Young Professionals (Project HANAMI), P. R. China, the JSPS Postdoctoral Fellowship for Research in Japan (ID No. P19081) from the Japan Society for the Promotion of Science (JSPS), Japan, the Grants-in-Aid for Scientific Research (17H00878) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, the Horizon H2020 Marie Skłodowska-Curie Actions Initial Training Network European Training Network (MSCA-ITN-ETN) project under grant agreement No. 766287 (TAPAS), Germany, and the Natural Science Foundation of Shenzhen University General Hospital (No. SUGH2018QD013), P. R. China. The first author would like to thank Ms. Xiao (Lexie) Li for her love and support.

II. Relation to Prior Work

In our previous studies, wavelets had been found to be efficient for extracting robust representations from body acoustical signals, e. g., snore sounds [6]. In this study, we first introduce wavelet energy features (WEF) into the field of automatically recognising heart sounds. In addition, we investigate the performance of sequential learning by DRNNs, which has been ignored in the aforementioned studies.

III. Materials and Methods

A. HSS Database

The HSS corpus was first released in the INTERSPEECH 2018 ComParE challenge heart beats sub-challenge [7]. This study was approved by the ethics committee of the Shenzhen University General Hospital. There were 170 participants (female: 55, male: 115, 65.4 ± 13.2 years, 21 to 88 years) involved in the data collection (the heart sounds were annotated as *normal*, *mild*, and *severe*). To make a subject independent

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

TABLE I
THE DATA SET PARTITIONING OF THE HSS CORPUS.

#	Train	Dev	Test	Σ
Normal	84	32	28	144
Mild	276	98	91	465
Severe	142	50	44	236
Σ	502	180	163	845

evaluation, and taking the gender, age, and class distributions into account, we split the whole database into train, development (dev), and test sets (see Table I). All the heart sound recordings were collected with an electronic stethoscope (Eko CORE, USA) set up via a Bluetooth 4.0 at a 4 kHz sampling rate. The average length of the recordings is 30 s (from 29.8 s to 30.2 s), which results in a whole length of approximately 423 min with 845 recordings.

B. Features and Classifier

We use WEF [6] as the low-level descriptors (LLDs) for representing the characteristics of the heart sounds. Compared to the traditional Fourier transformation, wavelet transformation can provide a multi-resolution analysis of the signal [8]. We further chose gated recurrent unit (GRU) cells [9] to overcome the *vanishing gradient* issue when training a DRNN [10]. The WEF based LLDs are extracted from 1 s length frames (with 50 % overlap) segmented from one recording, and stacked as a sequence when feeding into the DRNN model.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

We selected 'coif3' as the wavelet type for extracting WEF (dimension: 287) from 7 decomposition levels (refer to [11]) by experiments on the development set. The DRNN model was optimised to have three hidden layers (512-256-128) with the 'Adam' optimiser. The learning rate, batch size, and iteration number was set to .001, 64, and 500, respectively. To improve the reproducibility of this work, we set the random seed as 12 in the experiments. To evaluate the performance of the proposed model, the unweighted average recall (UAR), i. e., the averaged recall for each class of heart sounds, was used due to the imbalanced distribution of HSS.

B. Results

The confusion matrix of the proposed model on the test set is shown in Fig. 1. Generally speaking, the wavelet based DRNN works for the three-class heart sound recognition task (UAR: 43.0 %, Chance Level UAR: 33.3 %). We can see that, the model has an excellent performance for recognising the *mild* class. However, its capacity in classifying *normal*, and *severe* needs to be improved. In particular, it is difficult for the current model to distinguish *normal* and *mild*, or *severe* and *mild*.

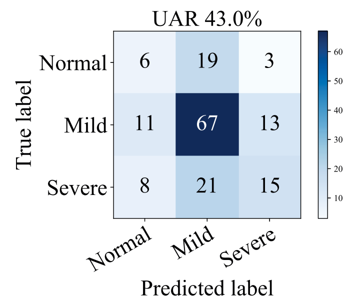


Fig. 1. Confusion Matrix of the Proposed Model on the Test Set.

V. CONCLUSION

In this study, we proposed a novel framework based on wavelet representations and deep recurrent neural networks for classifying heart sounds into *normal*, *mild*, and *severe*. The experiments showed the encouraging and promising performance of the model. In future work, we will involve more sophisticated methods like attention based models [12].

REFERENCES

[1] E. Wilkins, L. Wilson, K. Wickramasinghe, P. Bhatnagar, J. Leal, R. Luengo-Fernandez, R. Burns, M. Rayner, and N. Townsend, *European Cardiovascular Disease Statistics 2017*. Brussels, Belgium: European Heart Network, 2017.

[2] D. Roy, J. Sargeant, J. Gray, B. Hoyt, M. Allen, and M. Fleming, "Helping family physicians improve their cardiac auscultation skills with an interactive cd-rom," *Journal of Continuing Education in the Health Professions*, vol. 22, no. 3, pp. 152–159, 2002.

[3] S. Mangione, "Cardiac auscultatory skills of physicians-in-training: a comparison of three english-speaking countries," *The American Journal of Medicine*, vol. 110, no. 3, pp. 210–216, 2001.

[4] S. Ismail, I. Siddiqi, and U. Akram, "Localization and classification of heart beats in phonocardiography signals: a comprehensive review," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 26, 2018.

[5] C. Liu, D. Springer, Q. Li, B. Moody *et al.*, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, p. 2181, 2016.

[6] K. Qian, M. Schmitt, C. Janott, Z. Zhang, C. Heiser, W. Hohenhorst, M. Herzog, W. Hemmert, and B. Schuller, "A bag of wavelet features for snore sound classification," *Annals of Biomedical Engineering*, vol. 47, no. 4, pp. 1000–1011, 2019.

[7] B. Schuller, S. Steidl, A. Batliner, P. B. Marschik *et al.*, "The INTERSPEECH 2018 Computational Paralinguistics Challenge: Atypical & self-assessed affect, crying & heart beats," in *Proc INTERSPEECH*, Hyderabad, India, 2018, pp. 122–126.

[8] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[9] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. NIPS Deep Learning and Representation Learning Workshop*, Montreal, Canada, 2014, pp. 1–9.

[10] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber *et al.*, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies," in *A Field Guide to Dynamical Recurrent Neural Networks*. Piscataway, NJ, USA: IEEE Press, 2001, pp. 237–244, J. F. Kolen, and S. C. Kremer, Ed.

[11] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121–131, 2011.

[12] Z. Ren, Q. Kong, J. Han, M. D. Plumbley, and B. Schuller, "Attention-based atrous convolutional neural networks: Visualisation and understanding perspectives of acoustic scenes," in *Proc. ICASSP*, Brighton, UK, 2019, pp. 56–60.

Combine Facebook Prophet and LSTM with BPNN Forecasting financial markets : the Morgan Taiwan Index

Wen-Xiang Fang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
ypps920054@gmail.com

Po-Chao Lan
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
s981724@gmail.com

Wan-Rung Lin
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
linyumiko@gmail.com

Hsiao-Chen Chang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
jean91525@gmail.com

Hai-Yen Chang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
irischang1014@gmail.com

Yi-Hsien Wang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
holland@mail2000.com.tw

***Abstract*—The recurrent neural network (RNN) used by many people in machine learning often faces the situation where the gradient disappears, in order to solve this problem. Modern scholars often use Long Short-Term Memory (LSTM) proposed in 1997 to predict time series samples. However, Facebook believes that most of the past time series models have missing adjustment parameters. Therefore, it developed a set of predictive tools Prophet for periodic parameters and trend parameters in 2017. In this paper, LSTM and Prophet are used to predict the trend of time series data, and the prediction trend is combined with the inverse neural network model (BPNN) for prediction. The empirical results show that this method can indeed achieve accurate forecasting trends and reduce errors. This research promises to contribute to this research literature in the future, thereby enhancing the ability of investors to target the long-term layout.**

Keywords—Deep learning、Time series、Forecast、MSCI Taiwan Index Futures、Financial market

I. Introduction

Long et al. (2019) stated that time series prediction has been the most difficult problem faced by researchers and speculators. Time series predictions can be applied to any data that changes over time, especially for stock prices. Moreover, it can be used to track changes in stock prices over time, and the results of the forecast can inform investors of the timing of investment in and out of the market.

In recent years, more and more studies have used various algorithms and machine learning to predict financial markets(Sisodia, Jadhav, 2018; Shah, Isah & Zulkernine, 2019 ; Picasso, Merello, Ma & Oneto, 2019). Therefore, in order to achieve more effective predictions, this study uses the additivity model Prophet developed by Facebook and Long Short-Term Memory (LSTM) time series model in deep learning. The MSCI Taiwan Index Futures (hereinafter referred to as MSF), which is operated by the Singapore Stock Exchange, forecasts trends and compares their trend forecasting capabilities. Based on the predicted trends of both, the reverse transfer neural network model (BPNN) was further introduced for short-term closing price prediction.

The MSF is included in the 105 stocks listed on the Taiwan Stock Exchange as the target, and the index set by capital weighting is selected. The selected companies are all enterprises with high market share in the industry. This index accounts for the total market value of the Taiwan stock market 70% what is very valuable reference. Nowadays, the main participants of MSF are foreign capital, the main purpose is to hedge and arbitrage, so this article uses this as the research target to predict, in the hope that it will bring wealth maximization effect for long-term foreign investment or short-term arbitrage.

This study hopes to improve the ability of foreign capital to hedge in MSF and the possibility of arbitrage opportunities through MSF trends and short-term price forecasts. This paper is divided into four paragraphs. The first part is the introduction, which explains the research background, motivation and research purposes. The second part is the research method; the third part is the empirical result; the last is the research conclusion of this paper.

II. Introduction of Research Method

A. Facebook Prophet

Prophet is an additivity (GAM) model that Facebook has open sourced, mainly for time series data. Prophet has excellent processing ability for predicting highly seasonal data with long-term non-stationary rends or for missing data (Upadhayula, Varadarajan, Ravilla,&Viswanathan,2019). Facebook has adopted Prophet as its application, and has applied changes to its parameters in R software and Python open users. Currently, users have used Prophet to predict sales and repurchase rates.

B. KERAS Construction LSTM KERAS with LSTM model

LSTM is a branch of the Recurrent Neural Network (RNN). Because LSTM sets up Dropout, it can solve the problem of gradient disappearance and excessive gradient during the operation. Moreover, because LSTM has more control units, it can deal with long-term time series data. With extremely good trend prediction ability.

LSTM is based on Python's deep learning database KERAS for model construction. In the process, Sickit-Learn's MinMaxScaler is used to scale the data between 0-1 in 3D matrix mode, and the Adam optimizer is used as the compilation model. Nelson (2017) uses LSTM to use the past price data of the stock as a training group to predict the future trend of the stock price. To evaluate whether LSTM is improved compared to other machine learning methods, empirical results show that LSTM has excellent predictive power.

C. Neural network

The back-transfer neural network trains the multilayer feedforward neural network using the error back propagation algorithm. It consists of an input layer, a hidden layer, and an output layer. The hidden layer is used for one or more layers, and each layer of neurons is called a node.

Zhou, Zhou, Yan., & Yang (2019) proposed the use of inverted-transverse neural networks to predict stock prices from nonlinear data, that is, through the selection of activation functions in inverse neural networks. Learning rate and hidden neuron number layer adjustment. It demonstrates the effectiveness of the inverse neural network model in predicting stock returns. The advantage of a class-like neural network lies in its ability to predict nonlinear models, which are widely used in time series prediction.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

D. Data Source

MSF's share price time series data from January 1, 2014 to December 31, 2018, during the sample period, is based on the data published by Taiwan Economics News Database TEJ.

III. EMPIRICAL ANALYSIS

A. facebook prophet

Figure 1 is the stock price trend chart predicted by prophet. The black point in the figure is the actual value of MSF and the blue line is the prediction curve of MSF. In this paper, the data of the first five years is used as the training group to predict the price trend of the last 60 days. The predicted curve and the predicted point are roughly the same in the trend. The empirical results show that prophet has certain advantages in predicting future price trends.

Fig1. Prophet 60-day prediction interval

Figure 2 and Figure 3 show the "seasonal" of the data using prophet, and see the monthly and daily fluctuations of the data. From Figure 2, the MSF usually shows an upward trend on Monday.In Figure 3, the MSF showed an upward trend in the first half of the year (January to June), however, the second half (July to December) showed a downward trend.

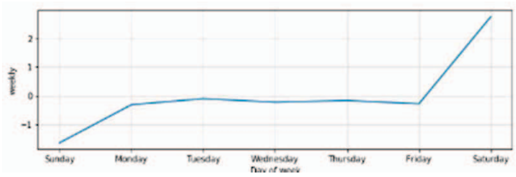


Fig2 Weekly seasonality of Prophet

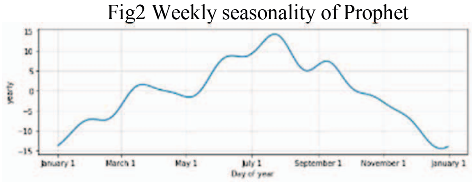


Fig3 Annual seasonality of Prophet

Figure 4 shows the use of prophet to predict the yearly stock price trend. Investors can use this chart to target the long-term layout of MSF as a reference for investors' investment decisions.

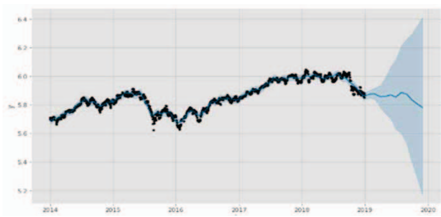


Fig4 Prophet 60-day Predictive value

B. LSTM

Figure 5 shows the stock price trend chart predicted by LSTM. The red line is the actual value curve of MSF and the blue line is the predicted value curve of MSF. In Figure 5, the LSTM predictions are consistent with the actual values, and the empirical results confirm the validity of the price trends predicted by LSTM.

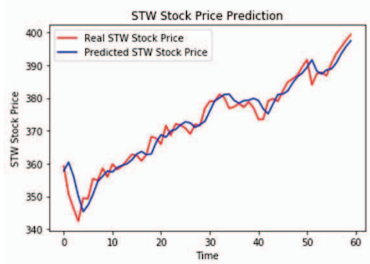


Fig5. the stock price trend chart predicted by LSTM.

C. Neural network

Figure 6 integrates the prediction results of LSTM into the Backpropagation Neural Network, the red line is the actual value curve of MSF, and the blue line is the predicted value curve of MSF.In figure 6 that the actual value and the predicted value are almost coincident, and the RMSE has reached 0.04464.The empirical results show that the use of integrated LSTM reverse-transfer-like nerves has an accurate predictive predictive power in short-term prediction.

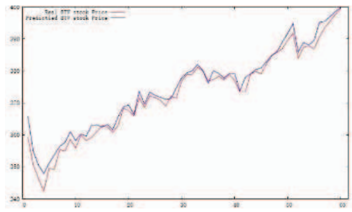


Fig6. the stock price trend chart predicted by BPNN.

IV. CONCLUSIONS

This study uses Prophet to predict the 60-day stock price trend after MSF. The empirical results show that Prophr has been able to predict the stock price trend of MSF in predicting the 60-day MSF stock price trend, and the MSF shows an upward trend on Monday.

It is concluded that the MSF showed an upward trend in the first half of the year (January to June), however, the second half (July to December) showed a downward trend.

The LSTM forecast and trend trends tend to be consistent, and the empirical results confirm the validity of the price trend predicted by LSTM.

Finally, the LSTM stock price prediction value is brought into the inverse transfer type nerve as the input variable. The results show that the predicted value of the integrated LSTM inverted transfer type nerve almost coincides with the actual value two curve, indicating that the LSTM reverse transfer type nerve can accurately predict the MSF. The stock price, RMSE has reached 0.04464.

REFERENCES

[1] Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, *164*, 163-173.

[2] Weytjens, H., Lohmann, E., & Kleinsteuber, M. (2019). Cash flow prediction: MLP and LSTM compared to ARIMA and Prophet. *Electronic Commerce Research*, 1-21.

[3] Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 1419-1426). IEEE.

An analysis of combining correlation screening with artificial neural network for FITX futures prediction

Fu-Ming Lai
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
lfm0619@gmail.com

Wen-Xiang Fang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
ypps920054@gmail.com

Chong-Kai Fu
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
fuchongkai@gmail.com

Fu-Ju Yang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
fuju@faculty.pccu.edu.tw

Kuang-Hsun Shih
CTBC Business School
Tainan, Taiwan
president@ctbc.edu.tw

Yi-Hsien Wang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
holland@mail2000.com.tw

***Abstract*—FITX futures provide investors with tools to evade risks and serve as an investor commodity; in other words, speculators and arbitragers make use of the high-level feature of futures to earn enormous returns with a small sum of money. FITX price is easily influenced by various factors, including finance, economy, politics, society, and investors' mind. In this study, some features of the FITX futures market and the back propagation network (BPN) are used to predict FITX futures. The BPN is adopted to predict the next-day closing prices of FITX futures. In this paper, the variables of the original data regarding price, quantity, and time, as well as the fluctuation estimated by GARCH, were taken as the input variables. The R software was used to establish the BPN. The next-day closing indices of FITX futures were predicted after training, as based on the back propagation model. In addition, the coefficient of correlation between the input variables and the closing prices is used to select the optimal input variable to establish the BPN model. According to the empirical research results, the proposed methods are more accurate in prediction in comparison with the original BPN model.**

Keywords—Artificial Neural Network (ANN), FITX futures, predict

I. INTRODUCTION

A robust prediction tool can assist investors in the improvement of return on investment and the effectiveness of hedging strategies. [2] indicated Stock investment has become an important investment activity in Taiwan. However, investors usually get loss because of unclear investment objective and blind investment. Therefore, to create a good investment decision support system to assist investors in making good decisions has become an important research problem. Artificial Neural Networks (ANN) can provide relatively good performances in forecasting stock price but it cannot explain the forecasting rules clearly.

The futures market has witnessed vibrant growth over recent years, as margin trading allows investors to achieve generous returns with financial leverage. Meanwhile, taking futures as an instrument serves hedging functions for domestic and foreign investors trading at large volumes, which renders the futures market healthier and more robust. Reference [9] suggested that index futures serve three purposes, an instrument for market trading, a hedging tool for market risks assistance in price discovery of the spot market. The first two functions are closely aligned with the portfolio theory. There are many index futures available in the market. There are many index futures available in the market. This paper seeks to establish an accurate prediction model for investors trading TAIEX futures.

The back propagation neural network (BPNN), as proposed by The input layers receive a large amount of variable data and simulate decision patterns via learning, while the hidden layers generate evaluations and outputs via the output layers. While neural networks are highly complex and difficult to explain, they are very suitable for the analysis and prediction of stocks, futures, and bond ratings, which is possibly due to their high-speed learning capability. Reference [12] posited The investment behavior also varied for different kinds of investors (traders, arbitrageurs and investors). If and only if the information obtained, relating to the stock prices, were pre-processed efficiently, using the machine learning method like the artificial neural network, the forecasting would become more accurate and the investors could ensure earning capital appreciation, for their stock investments, which would ensure maximization of wealth in the long run.

This paper consists of four parts. Part 1 is the introduction of the research background, motivation, and objectives. Part 2 explains the research methodology. Part 3 presents the empirical findings and analysis. Part 4 summarizes the research conclusions and suggestions.

II. RESEARCH METHOD

A.Back propagation network (BPN)

BPNNs obtain training data sets and predictive data sets from the question domain. The training set is the input variable of the network, and the weighted coefficient is adjusted time and again by using the gradient steepest descent method, in order to minimize the error and bring the output value as close as possible to the target/actual value. BPNNs are known for their capability in parallel processing of massive data, non-linear outputs, and predictive ability by utilizing the multiple-layer structure. Therefore, BPNNs are suitable for prediction and analysis. A back propagation neural network is widely used well known multi-layer supervised feed forward neural network algorithm since its simplicity and high problem-solving ability. In the traditional back propagation neural network, weight updating done by gradient decent based learning algorithm which is falling into local minima and learning rate is slow. Hence, keep away from above mentioned drawbacks; the opposition based learning (OBL) algorithm is used for weight adjustment in a back propagation neural network[10]. Reference [3] suggested the use of non-structured data in BPNNs to predict share prices, which includes the selection of functions, learning rates, and the application of hidden neural layers. Zhou also proposed a validation method to prove the effectiveness of BPNNs in the prediction of return on equity investments.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

III. Empirical Results

A. Establishment of the BPN model

This study adopted R software to establish the BPN model, where price, quantity, time, and GARCH were taken as the input variables to train the BPN model. Then, the R software was used to obtain the parameter combination of the number of hidden ANN layers and neuron nodes that have the least RMSE.

1) Data Sources

In this study, the FITX futures of the Taiwan Futures Exchange from January 1, 2008 to December 31, 2018 were chosen. The data source was the TEJ Futures Database of the FITX Daily Exchange List, as released by the Taiwan Futures Exchange.

2) Input Variables

The BPN was taken as the model of predicting the next-day closing prices of FITX. In the selection of main influencing factors, the variables of the original data regarding price, quantity, and time were adopted. Regarding price, target stock price and basis were selected; regarding quantity, open quantity and turnover were chosen; regarding time, the due time of futures was selected. In addition, GARCH fluctuation was chosen.

3) Data Processing

The collected variable data were standardized in this paper. The difference between Variable X and Variable Mean X was obtained through standardization, and it was "σ" times of the standard deviation; in other words, the standard deviation "σ" was taken as the unit. The standardization of the location of the two numbers in their own data pool could prevent excessive difference among the variables, which would otherwise affect the prediction ability of ANN.

B. Performance Assessment Method

The Root Mean of Squared Error (RMSE) was used for the performance assessment in this study.

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{1}^{n}(x_i-\widehat{x_\imath})^2} \tag{1}$$

In this equation, " x_i" is the actual value; "$\widehat{x_t}$" is the predicted value; "n" is the quantity of samples.

C. Optimal Prediction Model Neuron Combination in Different Years

the RMSE of the artificial neural model was around 0.03 from 2008 to 2018. This shows that the model has the performance of accurate prediction.

D. Selection of Optimal Input Variable Combination

In this paper, the Pearson correlation coefficient was used to select the correlation coefficient, the significant ones for the next-day closing price were chosen as the input variables, The input variables were used to determine whether the correlation coefficient could improve the prediction performance.."TX" (weighted share price index price of Taiwan), "Basis" (basis), "Volume" (turnover), "OI" (open interest), and "GARCH" (GARCH fluctuation) were significant for the next-day closing price (P). Therefore, "TX" (weighted share price index price of Taiwan), "Basis" (basis), "Volume" (turnover), "OI" (open interest), and "GARCH" (GARCH fluctuation) were taken as the input variables.

The RMSE became significantly lower before and after variable selection in the correlation coefficient-based selection of the optimal variable. This demonstrates that the correlation coefficient-based selection can improve prediction.

IV. Conclusion and Suggestions

This paper adopted Pearson correlation coefficient to choose the optimal variable combination, and the results is "TX" (weighted share price index price of Taiwan), "Basis" (basis), "Volume" (turnover), "OI" (open interest), and "GARCH" (GARCH fluctuation) were significant for the next-day closing price (P); therefore, they were taken as the input variables. In addition, the empirical research shows that RMSE became significantly lower (as shown in Table 4) before and after the variable selection in the correlation coefficient-based selection of the optimal variable. This demonstrates that the correlation coefficient-based selection can improve prediction.

FITX futures were taken as the research subject in this study. The ANN prediction tool can serve as a risk evasion strategy for investors when making decisions about investments. Future studies can also delve into other futures, as it is believed that they will bring a higher investment return rate to investors.

References

[1] A. Vejendla, and D. Enke, "Evaluation of GARCH, RNN and FNN Models for Forecasting Volatility in the Financial Markets," IUP Journal of Financial Risk management, vol.10, 2013, pp.41-49.

[2] F. Zhou, H. M. Zhou, Z. Yang, and L. Yang, "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction," Expert Systems with Applications, vol.115, 2019, pp.136-151.

[3] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[4] G. H. Kim, and S. H. Kim, "Variable Selection for Artificial Neural Networks with Applications for Stock Price Prediction," Applied Artificial Intelligence, vol.33, 2019, pp.54-67.

[5] H. Guan, Z. Dai, A.Zhao, and J. He, "A novel stock forecasting model based on High-order-fuzzy-fluctuation Trends and Back Propagation Neural Network," PloS one, vol.13, 2018, e0192366.

[6] H. Karabiyik, P. K. Narayan, D. H. B. Phan, and J. "Westerlund, Islamic spot and index futures markets: Where is the price discovery?" Pacific-Basin Finance Journal, vol.52, 2018, pp.123-133.

[7] K. Zhou, K. Velusamy, and C. Gomathi, "Financial prediction using back propagation neural networks with opposition based learning," In Journal of Physics: Conference Series (Vol. 1142, No. 1, p. 012008), 2018. IOP Publishing.

[8] Li Weiping, Li Guocheng, Lai Jinhui, Liao Yirou, Chen Junxi, & Zhong Wei, "Using the neural network to construct the forecast model of Taiwan stock index futures," Journal of Advanced Engineering, vol.13, 2018, pp.73-77.

[9] M. O. Sigo, M. Selvam, B. Maniam, D. Kannaiah, C.Kathiravan, and T. Vadivel, "Big data analytics-application of artificial neural network in forecasting stock price trends in india," Academy of Accounting and Financial Studies Journal, vol.22, 2018, pp.1-13.

[10] Y. Tsai, "Discussion on the Asymmetric Interpretation Factors of Taiwan's Stock Volatility——Taking the Panel VAR Model as an Example," Dissertation of the Institute of Financial Management, Sun Yat-Sen University,2000, pp.1-56.

Comparison of Forcasting Ability between Back-propagation Network and ARIMA in the Prediction of Bitcoin Price

Chung-Chieh Chen
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
corey0206cc@gmail.com

Jung-Hsin Chang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
ejo3vm@gmail.com

Fang-Cih Lin
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
c87610@gmail.com

Jui-Cheng Hung
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
hung660804@gmail.com

Cheng-Shian Lin
Department of Artificial Intelligence
CTBC Business School
Tainan, Taiwan
charcheslin@ctbc.edu.tw

Yi-Hsien Wang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
holland@mail2000.com.tw

***Abstract*—Bitcoin is a peer-to-peer (P2P) electronic currency that allows online payments around the world without the management of a third party. Many studies have been conducted on the performance prediction of the stock market; in particular, the Autoregressive Integrated Moving Average model (ARIMA) is one of the linear regressive models widely used in the time series. Nevertheless, as artificial intelligence has become a heated research topic in modern days, A number of studies have also shown that the Back-propagation Neural Network (BPNN) is very effective in prediction. Hence, this paper compares the ARIMA model with the BPNN model in the prediction of Bitcoin price.**

Keywords—neural network (NN), Bitcoin, time series, prediction

I. Introduction

Of the encrypted currencies, Bitcoin, as created in 2008, enjoys the highest visibility. In his academic article entitled "Bitcoin: A P2P Electronic Cash System", said that most previous transactions were conducted through electronic payments by a third party, which resulted in a loophole in the trust mechanism. Thus, proposed combining the Internet with computer operation, and used the decentralization of the block chain to achieve P2P online payment, which enhanced the security of online transactions. As Bitcoin is free from the supervision of any interested party and is an independent financial product in the market, the Bitcoin price has soared over the years, and even reached a new height in 2017. daily fluctuation may be as high as 4.61%; therefore, it is very important to predict Bitcoin price to increase the benefits for investors[1].

the time series model is not effective in analyzing or predicting non-linear relations[2]. To solve the problem of inadequate analysis and prediction of non-linear samples, an increasing number of studies have employed both the time series model and the NN model to predict prices and compare the two models, in order to reduce deviation and achieve effective investment prediction[3][4].

According to previous academic papers, this study adopted the ARIMA model and BPNN to predict Bitcoin price, respectively. The empirical results show that the non-linear relations of BPNN make it more effective in predicting Bitcoin price. This paper intends that this research will enrich the studies of the machine learning-based prediction of Bitcoin, and that the prediction model established in this research will help investors predict prices in a more accurate manner and make effective investment. This paper is comprised of four parts.

II. Research Method

A. Back-propagation Neural Network (BPNN)

BPNN is a multi-layered feedforward network with the self-learning ability, which uses the gradient steepest descent method to obtain an optimal weight vector among non-linear vectors. Only by reducing the deviation to an acceptable range through the re-definition of the weight can the optimal calculation of partial weight be outputted through the Output layer. This will improve the estimated value. The calculation of the gradient steepest descent method is shown in Eq. (1):

$$\Delta W_{ij} = c \cdot \ddot{a}_j^n \cdot A_i^{n-1} \tag{1}$$

It's necessary to evaluate the effectiveness of convergence in the machine learning of ANN. Usually, the root mean square is adopted for the measurement, as shown in Eq. (2):

$$\sqrt{\frac{\sum_p^M \sum_j^N \left(\hat{O}_j^p - \tilde{O}_j^p\right)^2}{M \cdot N}} \tag{2}$$

The Root Mean of Squared Error (RMSE) is used in all performance assessment methods, where "x_i" is the actual value; "x^_i", the predicted value; "n", the quantity of samples, as shown in Eq. (3):

$$\text{RMSE}= \sqrt{\frac{1}{n}\sum_1^n (x_i - \hat{x}_i)^2} \tag{3}$$

B. Autoregressive Integrated Moving Average (ARIMA) model

The ARIMA model is a linear regressive model (p, d, q). "AR" refers to autoregressive; "MA" refers to moving average; "p" denotes the quantity of the delayed observation in the model; "d" denotes the number of differences of the original observed value; "q" refers to the size of the moving average window. The ARIMA model is mainly used to analyze and predict the statistical model of smooth time series data, and its equations are shown in Eqs. (4) and (5):

$$\hat{X}_t = C + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{4}$$

$$= C + \sum_{k=1}^{p} \emptyset_k X_{t-k} - \sum_{l=1}^{q} \theta_l \varepsilon_{t-1} \tag{5}$$

C. Sample Processing and Model Establishment

1) Data Sources

The research samples are the Bitcoin prices during the period from January 1, 2014 to December 31, 2018. The data sources are the "Bitcoin in US Dollars", as released by Yahoo Finance and the exchange rates from Investing.com.

2) Data Summary

For the first step, the variable data from Yahoo Finance and Investing.com were statistically standardized. Frequently used in research, standardization uses the mean (μ) and the standard deviation (σ) to obtain the consistency of the variable (x) in the given range. The standardized variable contributes to a more accurate prediction by BPNN.

3) Establishment of the BPNN

In this study, the free statistical programming software R was adopted to establish the BPNN, where the moving average of the 20th day (MA20), the closing price, the highest price, and the US dollar-to CNY exchange rate (USACNY) were taken as the input variables. seek an optimal combination and strengthen the prediction ability.

The established model is shown in Eq. (6):

$$P = F\ (SKD,\ MA20,\ Close,\ USDCNY) \tag{6}$$

D. Variables Influencing Bitcoin Price

1) Moving average

The moving average is one of the indicators of technical analysis, it is also an essential indicator that is most widely adopted by investors. For that reason, this study included it as an input variable, as shown in Eq. (7):

$$MA_t = (P\_t + P\_(t-1) + P\_(t-2) + \cdots + P\_(t - N + 1))\ /\ N \tag{7}$$

2) Price

Market expectation and the entry and exit of speculators are two factors that exert direct impact on Bitcoin price; in particular price holds the greatest attention of investors, this study includes closing price and the highest price as two output variables.

3) Exchange rate

Bitcoin is closely related to the US dollar, and China is one of the main participants. Some academic papers show that encrypted currency would have direct impact on China's exchange rate[5]. Hence, this study takes USACNY as an input variable to discuss the change to deviation of different years when the exchange rate is included as a variable.

III. Empirical Results

year	**Table1. Comparison of the best RMSE for each year**	
	ARIMA	***BPNN***
2018	0.17905	0.19267
2017	0.09803	0.05688
2016	0.13747	0.13192
2015	0.19390	0.18411
2014	0.23400	0.23758

Note: * indicates lower RMSE

Table 1 shows the BPNN-based prediction of the optimal combinations of different years and the ARIMA (1, 0, 1)-based prediction. The research findings show that the predication deviation of BPNN is less than that of the ARIMA model, which demonstrates that ANN has some prospect of development.

IV. Conclusion and Suggestions

In this study, Bitcoin in US dollars was taken as the research subject, and ARIMA (1, 0, 1) and BPNN were used for prediction, respectively. In this way, this study aims to compare the prediction abilities between linear and non-linear relation models. According to the empirical results, BPNN was effective in predicting the samples over the past five years, especially in 2017, when the Bitcoin price rose from less than USD 1,000 to the unprecedented USD 20,000. According to the prediction results in 2017, ARIMA has remarkable prediction deviation over RMSE. However, the BPNN model can reduce the deviation to lower RMSE.

In 2014 and 2018, when the price fluctuation was relatively stable, the ARIMA model had low deviation. Therefore, this study suggests that investors should observe the properties of recent historical data when establishing a prediction model for Bitcoin prices. Meanwhile, they should adopt the ARIMA model and BPNN for prediction, in order to increase investment return and establish comprehensive risk evasion strategies.

References

[1] A. Radityo, Q. Munajat, and I.Budi, "Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods," In *2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2017, pp. 433-438. IEEE.

[2] D. Ö. Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Engineering Applications of Artificial Intelligence*, vol.*23*, 2010, pp.586-594.

[3] E. S. Karakoyun, and A. O. Cibikdiken, "Comparison of ARIMA Time Series Model and LSTM Deep Learning Algorithm for Bitcoin Price Forecasting," In *The 13th Multidisciplinary Academic Conference in Prague 2018*

[4] S. Karasu, A. Altan, Z. Saraç, and R. Hacioğlu, "Prediction of Bitcoin prices with machine learning methods using time series data," In *2018 26th Signal Processing and Communications Applications Conference (SIU)*, 2018, pp. 1-4. IEEE.

[5] W. Fang, S. Tian, and J. Wang, "Multiscale fluctuations and complexity synchronization of Bitcoin in China and US markets," *Physica A: Statistical Mechanics and its Applications*, vol.*512*, 2018, pp.109-120.

Single-Input-Single-Output System with Interference-Unaware Time-Based Receive Transformation under Cochannel Interference and Intersymbol Interference

Jui Teng Wang
Department of Electrical Engineering
National Chi Nan University
Puli, Nantou, Taiwan
jtwang@ncnu.edu.tw

Tsun-Chia Wang
wtsunchia@yahoo.com

***Abstract*—In this paper, we propose the interference-unaware time-based receive transformation (IUTRT), which can employ the time diversity provided by cochannel interferences (CCI) and intersymbol interference (ISI) in the single-input-single-output (SISO) system. We show that IUTRT can completely eliminate the effect of CCI and the inter-frame ISI (between different data frames) so that the time diversity provided by the intra-frame ISI (within the same data frame) can maximize the achievable rate. We also show that the achievable rate of the SISO system with IUTRT can be larger than the CCI-free and ISI-free SISO channel capacity and this advantage increases with the symbol rate and/or the number of symbols per frame.**

Keywords—CCI , ISI, SISO.

I. Introduction

In this paper, we propose the interference-unaware time-based receive transformation (IUTRT), which can be regarded as the time based version of the interference-unaware receive transformation (IURT) proposed in [1], [2] and can employ the time diversity provided by CCI and ISI in the SISO system. We consider the SISO system where the receiver uses the signals received during finite number of symbol periods to detect one data frame consisting of finite number of symbols. We reduce the symbol period (by increasing the symbol rate) so that the timing correlations of CCI can be constant over one data frame. As a result, the number of symbols that ISI affects (which is equal to the ratio of the delay spread of the channel to the symbol period) will be increased. We show that IUTRT can completely eliminate the effect of CCI and the inter-frame ISI (between different data frames) so that the time diversity provided by the intra-frame ISI (within the same data frame) can maximize the achievable rate. We also show that the achievable rate of the SISO system with IUTRT can be larger than the CCI-free and ISI-free SISO channel capacity and this advantage increases with the number of symbols per frame and/or the number of symbols that ISI affects. This implies that increasing the symbol rate can not only completely eliminate the effect of CCI but also increase the advantage of the achievable rate over the CCI-free and ISI-free SISO channel capacity.

II. System Model

Consider the SISO system where each symbol is transmitted serially, then at the kth symbol period, the received signal can be expressed as

$$r_k = \sum_{j=\max(k-L,0)}^{j=k} h_{kj} s_j + \eta_k + n_k , \tag{1}$$

where s_j denotes the jth transmitted symbol, η_k denotes the CCI from other transmitters, n_k denotes the received additive noise, L denotes the number of symbols that ISI affects and h_{kj} denotes the channel gain from the jth transmitted symbol.

The receiver uses the signals received during N symbol periods to detect one M-symbol data frame, thus on the basis of (1), the signals received during N symbol periods are modeled as follows:

$$\mathbf{r} = \mathbf{Hs} + \boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n} , \tag{2}$$

where $\mathbf{r} := (r_k)_{N\times 1}$ represents the vector for the received signals, $\mathbf{H} := (h_{kj})_{N\times M}$ represents the channel matrix for the considered M symbols received during N symbol periods, $\mathbf{s} := (s_k)_{M\times 1}$ represents the vector for the considered M-symbol frame, $\boldsymbol{\eta} := (\eta_k)_{N\times 1}$ represents the vector for the CCI from other transmitters, $\boldsymbol{\mu} := (\mu_k)_{N\times 1}$ represents the vector for the ISI from other frames and $\mathbf{n} := (n_k)_{N\times 1}$ represents the vector for the additive noises. Note that for the term $\sum_{j=\max(k-L,0)}^{j=k} h_{kj} s_j$ in (1), the part associated with the intra-frame ISI is integrated into $\mathbf{Hs}$ while the part associated with the inter-frame ISI is expressed as $\boldsymbol{\mu}$. Furthermore, the knowledge of the channel state information (CSI) is only available at the desired receiver but unavailable at all transmitters and other receivers.

III. IUTRT

In this section, we propose IUTRT for the SISO systems under CCI and ISI. By substituting the space for the time, we can regard (2) as the system model for the MIMO system under CCI, then as shown in [1], the transformation matrix $\mathbf{T}$ of IURT satisfies

$$\mathbf{TRT}^H = \mathbf{I} , \tag{3}$$

where $\mathbf{R} = E[(\boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n})(\boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n})^H]$, and it can make the achievable rate reach the MIMO channel capacity expressed as

$$\widetilde{C} = \log_2 \det(\mathbf{HSH}^H \mathbf{R}^{-1} + \mathbf{I}_N) , \tag{4}$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

where we let $\mathbf{S}=E[\mathbf{s}\mathbf{s}^H]$. Since the receiver of the considered SISO system uses the signals received during N symbol periods to detect one M-symbol data frame, the achievable rate for the SISO system with IUTRT can reach the SISO channel capacity expressed as

$$C=\frac{\log_2\det(\mathbf{H}\mathbf{S}\mathbf{H}^H\mathbf{R}^{-1}+\mathbf{I})}{N}, \tag{5}$$

which can be obtained by dividing (4) by N. For the case with $M=N$, $\mathbf{S}=p\mathbf{I}$, $\mathbf{R}=\alpha\mathbf{I}$ and $\mathbf{H}=h\mathbf{I}$, we have

$$C=\frac{\log_2\det((1+\frac{p|h|^2}{\alpha})\mathbf{I})}{N}=\frac{\log_2(1+\frac{p|h|^2}{\alpha})^N}{N}$$
$$=\log_2(1+\frac{p|h|^2}{\alpha}), \tag{6}$$

which gives the same expression as that for the CCI-free and ISI-free SISO channel capacity.

Note that we assume that precoding is not available at the transmitter, thus the optimal $\mathbf{S}$ for this assumption is given by $\mathbf{S}=p\mathbf{I}$, where p represents the transmit power. As a result, since $\mathbf{R}=E(\mathbf{r}\mathbf{r}^H)-\mathbf{H}\mathbf{S}\mathbf{H}^H$, we can employ $E(\mathbf{r}\mathbf{r}^H)$ and $\mathbf{H}$ (the local CSI) to evaluate $\mathbf{R}$ by $\mathbf{R}=E(\mathbf{r}\mathbf{r}^H)-p\mathbf{H}\mathbf{H}^H$. Alternatively, we can stop transmitting the data for a period so that we can evaluate $\mathbf{R}$ by $\mathbf{R}=E(\mathbf{r}\mathbf{r}^H)$ without the knowledge of the local CSI. As done in [1], we can use the eigenvalues and the eigenvectors of the evaluated $\mathbf{R}$ to solve (3) for $\mathbf{T}$. It is clear that the above realizations of IUTRT do not need the knowledge of CCI.

As implied by (3), IUTRT transforms interference plus noise into the noise uncorrelated over all symbol periods, hence IUTRT can be followed by any interference-neglected detection, and its operation is described in the following procedure.

Step 1: Impose IUTRT on the received signal and obtain the transformed received signal by

$$\mathbf{r}'=\mathbf{T}\mathbf{r}. \tag{7}$$

Step 2: Evaluate the transformed channel matrix by

$$\mathbf{H}'=\mathbf{T}\mathbf{H}. \tag{8}$$

Step 3: Use $\mathbf{r}'$ and $\mathbf{H}'$ to perform the detection for the M-symbol frame.

Note that the local CSI is needed in the procedure of the detection so that we have to perform the channel estimation under CCI. In fact, we can employ the timing correlations of the CCI to reduce the channel estimation error.

IV. Numerical Results

In this section, we simulate the SISO channel specified in Section 2 and insert the simulated channel gains into (5) to obtain the average achievable rates for the SISO system with IUTRT. As done in [3], we assume that all channel gains are independent random variables distributed as $CN(0, 1)$ with the real and imaginary parts each having variance 1/2. In the simulations, we assume that $N=M$. Furthermore, we assume that the timing correlations of CCI are constant over M symbol periods and there are 2 interferers with each interferer having the same transmit power as that of the considered transmitter and being equipped with one transmit antenna.

In Figure 1, we plot the average achievable rate (in bps/Hz) of the SISO system versus the number of symbols per frame (M), where the achievable rate of IUTRT is compared to the maximum rate and the CCI-free and ISI-free SISO channel capacity. It can be seen that the achievable rate of IUTRT converges to the maximum rate with the increase in M and its advantage over the CCI-free and ISI-free SISO channel capacity increases with M and/or L (the number of symbols that ISI affects). We also found that decreasing SNR (signal-to-noise ratio) and/or L can decrease the required M for making the achievable rate reach the maximum rate.

[1] J. T. Wang, "Decorrelation based receive transformation for MIMO system under multi-user co-channel interference," *IEEE Wireless Communications Letters*, vol. 3, no. 3, pp. 305-308, June 2014.

[2] J. T. Wang, "Interference-free criterion for interference-unaware receive transform in MIMO co-channel interference," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 210–213, April 2018.

[3] L. H. Grokop and D. N. C. Tse, "Diversity–multiplexing tradeoff in ISI channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 109-135, January 2009.

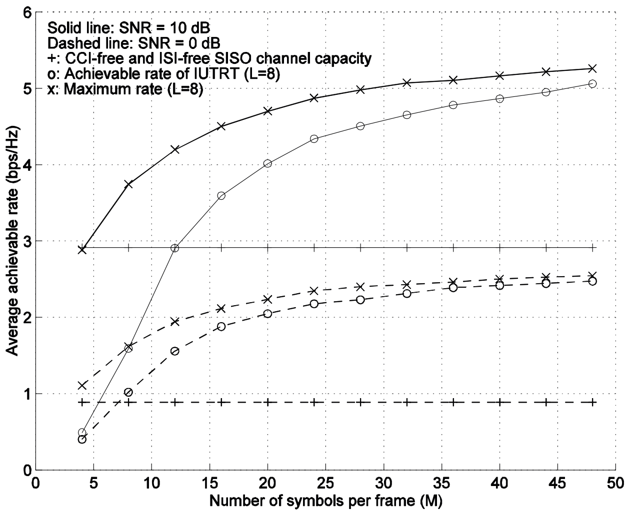


Figure 1. The average achievable rate (in bps/Hz) of the SISO system versus the number of symbols per frame (M).

Time-Space MIMO System with Interference-Unaware Time-Space Receive Transformation under Cochannel Interference and Intersymbol Interference

Jui Teng Wang
Department of Electrical Engineering
National Chi Nan University
Puli, Nantou, Taiwan
jtwang@ncnu.edu.tw

Tsun-Chia Wang
wtsunchia@yahoo.com

***Abstract*—In this paper, we introduce the time-space MIMO system (where the signals received at different time instants and spatial spots are jointly processed) and propose the interference-unaware time-space receive transformation (IUTSRT), which can integrate the time diversity with the spatial diversity for the time-space MIMO system under cochannel interferences (CCI) and intersymbol interference (ISI). We show that IUTSRT can completely eliminate the effect of CCI and the inter-frame ISI (between different data frames) so that the time diversity provided by the intra-frame ISI (within the same data frame) can maximize the achievable rate. We also show that the achievable rate of the time-space MIMO system with IUTSRT can be larger than the CCI-free and ISI-free MIMO channel capacity and this advantage increases with the symbol rate and/or the number of symbols per frame.**

Keywords*—*MIMO, CCI, ISI.

I. Introduction

We consider the time-space MIMO system where the receiver uses the signals received during finite number of symbol periods to detect one data frame consisting of finite number of symbols. We reduce the symbol period (by increasing the symbol rate) so that the timing correlations of CCI can be constant over one data frame. As a result, the number of symbols that ISI affects (which is equal to the ratio of the delay spread of the channel to the symbol period) will be increased. We show that IUTSRT, which can be regarded as the time-space version of the interference-unaware receive transformation (IURT) proposed in [1], [2], can completely eliminate the effect of CCI and the inter-frame ISI (between different data frames) so that the time diversity provided by the intra-frame ISI (within the same data frame) can maximize the achievable rate. We also show that the achievable rate of the time-space MIMO system with IUTSRT can be larger than the CCI-free and ISI-free MIMO channel capacity and this advantage increases with the number of symbols per frame and/or the number of symbols that ISI affects. This implies that increasing the symbol rate can not only completely eliminate the effect of CCI but also increase the advantage of the achievable rate over the CCI-free and ISI-free MIMO channel capacity.

II. System Model

Consider the time-space MIMO system where the signals received at different time instants and spatial spots are jointly processed. The receiver uses the signals received at N receive antennas during T symbol periods to detect M data frames (with L symbols per frame) from M transmit antennas, thus the signals received at N receive antennas during T symbol periods are modeled as follows:

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n}\,, \tag{1}$$

where $\mathbf{r}$ represents the vector for the received signals, $\mathbf{H}$ represents the channel matrix, $\mathbf{s}$ represents the vector for the data frames, $\boldsymbol{\eta}$ represents the vector for the cochannel interferences from other transmitters, $\boldsymbol{\mu}$ represents the vector for the received signals from other data frames and $\mathbf{n}$ represents the vector for the additive noises. Note that $\mathbf{r} = [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_N]$, $\mathbf{s} = [\mathbf{s}_1; \mathbf{s}_2; \cdots; \mathbf{s}_M]$, $\boldsymbol{\eta} = [\boldsymbol{\eta}_1; \boldsymbol{\eta}_2; \cdots; \boldsymbol{\eta}_N]$, $\boldsymbol{\mu} = [\boldsymbol{\mu}_1; \boldsymbol{\mu}_2; \cdots; \boldsymbol{\mu}_N]$, $\mathbf{n} = [\mathbf{n}_1; \mathbf{n}_2; \cdots; \mathbf{n}_N]$, where $\mathbf{r}_k = [r_{k1}; r_{k1}; \cdots; r_{kT}]$, $\mathbf{s}_k = [s_{k1}; \mathrm{s}_{k1}; \cdots; \mathrm{s}_{kL}]$, $\boldsymbol{\eta}_k = [\eta_{k1}; \eta_{k1}; \cdots; \eta_{kT}]$, $\boldsymbol{\mu}_k = [\mu_{k1}; \mu_{k1}; \cdots; \mu_{kT}]$, $\mathbf{n}_k = [n_{k1}; \mathrm{n}_{k1}; \cdots; \mathrm{n}_{kT}]$. Furthermore, $\mathbf{H}$ can be represented by a block matrix with N row partitions and M column partitions, where each sub-matrix has the same size $T \times L$ and the (k, j)th sub-matrix is denoted by $\mathbf{H}_{kj}$. As a result, we have

$$\mathbf{r}_k = \sum_{j=1}^{M} \mathbf{H}_{kj}\mathbf{s}_j + \boldsymbol{\eta}_k + \boldsymbol{\mu}_k + \mathbf{n}_k\,. \tag{2}$$

Note that the knowledge of the channel state information (CSI) is only available at the desired receiver but unavailable at all transmitters and other receivers.

III. IUTSRT

In this section, we propose IUTSRT for the time-space MIMO system under CCI and ISI. By letting $T = L = 1$ so that we can regard (1) as the system model for the space based MIMO system under CCI, then as shown in [1], the transformation matrix $\mathbf{T}$ of IURT satisfies

$$\mathbf{T}\mathbf{R}\mathbf{T}^H = \mathbf{I}\,, \tag{3}$$

where $\mathbf{R} = E[(\boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n})(\boldsymbol{\eta} + \boldsymbol{\mu} + \mathbf{n})^H]$, and it can make the achievable rate reach the space based MIMO channel capacity expressed as

$$\widetilde{C} = \log_2 \det(\mathbf{H}\mathbf{S}\mathbf{H}^H\mathbf{R}^{-1} + \mathbf{I})\,, \tag{4}$$

where we let $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H]$. Since the receiver of the considered time-space MIMO system uses the signals received at N receive antennas during T symbol periods to detect M data frames from M transmit antennas, the achievable rate of IUTSRT can reach the time-space MIMO channel capacity expressed as

$$C = \frac{\log_2 \det(\mathbf{HSH}^H \mathbf{R}^{-1} + \mathbf{I})}{T}, \tag{5}$$

which can be obtained by dividing (4) by T. For the case with $L = T$, $\mathbf{S} = p\mathbf{I}$, $\mathbf{R} = \alpha\mathbf{I}$ and $\mathbf{H}_{kj} = h_{kj}\mathbf{I}$, let $\mathbf{A}_{kj}$ denote the (k, j)th sub-matrix of $\mathbf{HH}^H$, then we have

$$\mathbf{A}_{kj} = \sum_{l=1}^{N} \mathbf{H}_{kl}\mathbf{H}_{jl}^* = \sum_{l=1}^{N} (h_{kl}\mathbf{I})(h_{jl}^*\mathbf{I}) = \sum_{l=1}^{N} (h_{kl}h_{jl}^*)\mathbf{I}, \tag{6}$$

as a result, it holds that

$$\begin{aligned} C &= \frac{\log_2 \det((p/\alpha)\mathbf{HH}^H + \mathbf{I})}{T} \\ &= \frac{\log_2[\det((p/\alpha)\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \mathbf{I})]^T}{T} \\ &= \log_2 \det((p/\alpha)\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \mathbf{I}), \end{aligned} \tag{7}$$

where $\hat{\mathbf{H}}$ is an $N \times M$ matrix with the (k, j)th element being h_{kj} and we use the fact that the eigenvalues of $\mathbf{HH}^H$ are the eigenvalues of $\hat{\mathbf{H}}\hat{\mathbf{H}}^H$ with multiplicity T. Note that the right side of (7) gives the CCI-free and ISI-free MIMO channel capacity.

We assume that precoding is not available at the transmitter, thus the optimal $\mathbf{S}$ for this assumption is given by $\mathbf{S} = p\mathbf{I}$, where p represents the transmit power of each transmit antenna. As a result, since

$$\mathbf{R} = E(\mathbf{rr}^H) - \mathbf{HSH}^H, \tag{8}$$

we can employ $E(\mathbf{rr}^H)$ and $\mathbf{H}$ (the local CSI) to evaluate $\mathbf{R}$ by

$$\mathbf{R} = E(\mathbf{rr}^H) - p\mathbf{HH}^H. \tag{9}$$

Alternatively, we can stop transmitting the data for a period so that we can evaluate $\mathbf{R}$ by $\mathbf{R} = E(\mathbf{rr}^H)$ without the knowledge of the local CSI. As done in [1], we can use the eigenvalues and the eigenvectors of $\mathbf{R}$ to solve (3) for $\mathbf{T}$. It is clear that the above realizations of IUTSRT do not need the knowledge of CCI.

As implied by (3), IUTSRT transforms the interference into the noise uncorrelated over all symbol periods and receive antennas, hence IUTSRT can be followed by any interference-neglected detection, and its operation is described in the following procedure.

Step 1: Impose IUTSRT on the received signal and obtain the transformed received signal by

$$\mathbf{r}' = \mathbf{Tr}. \tag{10}$$

Step 2: Evaluate the transformed channel matrix by

$$\mathbf{H}' = \mathbf{TH}. \tag{11}$$

Step 3: Use $\mathbf{r}'$ and $\mathbf{H}'$ to perform the detection for M data frames.

Note that the local CSI is needed in the procedure of the detection so that we have to perform the channel estimation under CCI. In fact, we can employ the timing correlations of CCI to reduce the channel estimation error.

IV. Numerical Results

In this section, we simulate the time-space MIMO channel specified in Section 2 and insert the simulated channel gains into (5) to obtain the average achievable rates for the time-space MIMO system with IUTSRT. As done in [3], we assume that all channel gains are independent random variables distributed as $CN(0, 1)$ with the real and imaginary parts each having variance 1/2. In the simulations, we assume that $T = L$. Furthermore, we assume that the timing correlations of CCI are constant over L symbol periods and there are 2 interferers with each interferer having the same transmit power as that of the considered transmitter and being equipped with one transmit antenna.

In Figure 1, we plot the average achievable rate (in bps/Hz) of the MIMO system versus the number of symbols per frame (L), where the achievable rate of IUTSRT is compared to the maximum rate and the CCI-free and ISI-free MIMO channel capacity. It can be seen that the achievable rate of IUTSRT converges to the maximum rate with the increase in L and its advantage over the CCI-free and ISI-free MIMO channel capacity increases with L and/or U (the number of symbols that ISI affects). We also found that decreasing SNR (signal-to-noise ratio) and/or U can decrease the required L for making the achievable rate reach the maximum rate.

[1] J. T. Wang, "Decorrelation based receive transformation for MIMO system under multi-user co-channel interference," *IEEE Wireless Communications Letters*, vol. 3, no. 3, pp. 305-308, June 2014.

[2] J. T. Wang, "Interference-free criterion for interference-unaware receive transform in MIMO co-channel interference," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 210–213, April 2018.

[3] L. H. Grokop and D. N. C. Tse, "Diversity–multiplexing tradeoff in ISI channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 109-135, January 2009.

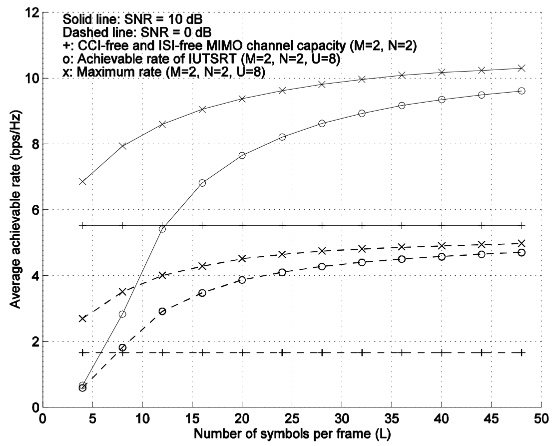


Figure 1. The average achievable rate (in bps/Hz) of the MIMO system versus the number of symbols per frame (L).

A Generation Method of a Two-Dimensional Optical ZCZ Sequence with the Zero-Correlation Zone $(4n-2)\times(4n-2)$

Takahiro Matsumoto *, Hideyuki Torii †, Yuta Ida *, Shinya Matsufuji *,
* Yamaguchi University, Ube, Yamaguchi 755–8611, Japan
† Kanagawa Institute of Technology, Atsugi, Kanagawa 243–0292, Japan

***Abstract*—In this paper, we propose a new generation method of a two-dimensional (2D) optical zero-correlation zone (ZCZ) sequence with the size of ZCZ $(4n-2)\times(4n-2)$. The 2D optical ZCZ sequence consists of pairs of a binary sequence which takes 1 or 0 and a bi-phase sequence which takes 1 or -1, and has a zero-correlation zone in the two-dimensional correlation function. Because of these properties, the 2D optical ZCZ sequence can be used for optical code-division multiple access (OCDMA) system using an LED array having a plurality of light-emitting elements arranged in a lattice pattern. The OCDMA system using the 2D optical ZCZ sequence can increase the data rate and can suppress interference by the light of adjacent LEDs. By using the proposed generation method, we can improve the peak value of the autocorrelation function of the sequence. This means that the BER performance of the OCDMA system using the sequence can be improved.**

***Index Terms*—optical wireless communication, optical code-division multiple access (OCDMA), optical zero-correlation zone (ZCZ) sequence, two-dimensional sequence, correlation function**

I. Introduction

The two-dimensional (2D) optical zero-correlation zone (ZCZ) sequence is a 2D sequence set consisting of pairs of a binary sequence which takes 1 or 0 and a bi-phase sequence which takes 1 or -1, and its correlation function has ideal correlation property in a certain shift zone around shift 0 which called a zero-correlation zone [1]. Because of these properties, the 2D optical ZCZ sequence can be used for optical code-division multiple access (OCDMA) system using an LED array having a plurality of light-emitting elements arranged in a lattice pattern. The OCDMA system using the 2D optical ZCZ sequence can increase the data rate and can suppress interference by the light of adjacent LEDs [2].

The 2D optical ZCZ sequence with the size of ZCZ 1×1 has been proposed [1], however, it is desirable to be able to select various sizes of ZCZ considering various applications. In addition, 2D optical ZCZ sequences can be easily generated by the product of 1D sequences, but the BER performance of OCDMA systems using the sequence is low performance because of the low peak value of the autocorrelation function of the sequence.

This work is supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (C)(18K04145), and the Telecommunications Advancement Foundation (TAF) Grant.

In this paper, we propose a new generation method of a 2D optical ZCZ sequence with the size of ZCZ $(4n-2)\times(4n-2)$ which has a high peak value of its autocorrelation function.

II. Definition of Two-Dimensional Optical ZCZ Sequence

Let Z be a 2D optical ZCZ sequence set of M pairs consisting of a bi-phase 2D sequence $z^{j}_{N_y,N_x}$ of length $N_y\times N_x$ whose elements take 1 or -1 and a binary 2D sequence $\hat{z}^{j,d}_{N_y,N_x}$ whose elements take 1 or 0, which can be expressed as

$$Z=\left\{(z^0_{N_y,N_x},\hat{z}^{0,d}_{N_y,N_x}),\cdots,(z^j_{N_y,N_x},\hat{z}^{j,d}_{N_y,N_x}),\right.$$
$$\left.\cdots,(z^{M-1}_{N_y,N_x},\hat{z}^{M-1,d}_{N_y,N_x})\right\}, \tag{1}$$

$$z^j_{N_y,N_x}=[z^j_{N_y,N_x,y,x}\in\{1,-1\}]$$
$$=\begin{bmatrix} z^j_{N_y,N_x,0,0} & \cdots & z^j_{N_y,N_x,0,N_x-1}\\ z^j_{N_y,N_x,1,0} & \cdots & z^j_{N_y,N_x,1,N_x-1}\\ \vdots & \ddots & \vdots\\ z^j_{N_y,N_x,N_y-1,0} & \cdots & z^j_{N_y,N_x,N_y-1,N_x-1}\end{bmatrix}, \tag{2}$$

$$\hat{z}^{j,d}_{N_y,N_x}=[\hat{z}^{j,d}_{N_y,N_x,y,x}\in\{1,0\}], \tag{3}$$

where j is the sequence number and $0\le j\le M-1$, d is the bit data and $d\in\{1,0\}$, M is the number of sequences and is called the family size, and y, x are the ordinal variable, $0\le y\le N_y-1$, $0\le x\le N_x-1$ and denotes a value of module N_y or N_x i.e., $y=y \bmod N_y$ and $x=x \bmod N_x$.

The periodic 2D correlation function between a bi-phase 2D sequence $z^j_{N_y,N_x}$ and a binary 2D sequence $\hat{z}^{j',d}_{N_y,N_x}$ for shifts y',x' of the optical 2D ZCZ sequence is given by

$$\rho_{z^j_{N_y,N_x},\hat{z}^{j',d}_{N_y,N_x},y',x'}=\sum_{y=0}^{N_y-1}\sum_{x=0}^{N_x-1} z^j_{N_y,N_x,y,x}\hat{z}^{j',d}_{N_y,N_x,y-y',x-x'}$$
$$=\begin{cases}(-1)^d w & ;x'=y'=0,j=j',\\ 0 & ;x'=y'=0,j\ne j',\\ 0 & ;|y'|\le Zcz_y,|x'|\le Zcz_x,\\ & \quad j\ne j',\end{cases} \tag{4}$$

where $0<w<N_yN_x$. In this paper, the above 2D correlation function $\rho_{z^j_{N_y,N_x},\hat{z}^{j',d}_{N_y,N_x},y',x'}$ is called the autocorrelation function for $j=j'$ and the cross-correlation function for $j\ne j'$.

III. Construction of 2D Optical ZCZ Sequence with ZCZ Size$(4n-2)\times(4n-2)$

A. Conventional Method

Let A be a 1D optical ZCZ sequence set [3] of length $N = N_1N_2$, $N_1 = 4n_1-1$ and $N_2 = 2^{n_2}$, family size $M = N_2-1$ and ZCZ size $4n_1-2$, which can be expressed as

$$A = \{(a_N^1, \hat{a}_N^{1,d}), \cdots, (a_N^j, \hat{a}_N^{j,d}), \cdots, (a_N^M, \hat{a}_N^{M,d})\}, \quad (5)$$

$$\begin{cases} a_N^j &= (a_{N,0}^j, \cdots, a_{N,i}^j, \cdots, a_{N,N-1}^j), \\ \hat{a}_N^{j,d} &= (\hat{a}_{N,0}^{j,d}, \cdots, \hat{a}_{N,i}^{j,d}, \cdots, \hat{a}_{N,N-1}^{j,d}), \end{cases} \quad (6)$$

where $a_{N,i}^j \in \{1,-1\}$ and $\hat{a}_{N,i}^{j,d} \in \{1,0\}$. In general, a 2D optical ZCZ sequence of ZCZ size $(4n_y-2)\times(4n_x-2)$ is obtained as a product of 1D optical ZCZ sequences of ZCZ sizes $4n_y-2$ and $4n_x-2$ of (5) and (6), and is given by

$$\begin{cases} z_{N_y,N_x}^j &= [z_{N_y,N_x,y,x}^j] = [a_{N_y,y}^{j_y} \cdot a_{N_x,x}^{j_x}], \\ \hat{z}_{N_y,N_x}^{j,d} &= [\hat{z}_{N_y,N_x,y,x}^{j,d}] = [\hat{a}_{N_y,y}^{j_y,d} \cdot \hat{a}_{N_x,x}^{j_x,d}], \end{cases} \quad (7)$$

where the ZCZ sizes of $(a_{N_y,y}^{j_y}, \hat{a}_{N_y,y}^{j_y,d})$ and $(a_{N_x,x}^{j_x}, \hat{a}_{N_x,x}^{j_x,d})$ are $4n_y-2$ and $4n_x-2$, respectively, and the family sizes of those sets are M_y and M_x, respectively, and sequence number j is given by $j = M_x(j_y-1)+(j_x-1)$. The periodic 2D correlation function of the optical 2D ZCZ sequence is given by

$$\rho_{z_{N_y,N_x}^j, \hat{z}_{N_y,N_x}^{j',d}, y', x'} = \begin{cases} \frac{(-1)^d N_yN_x(Zcz_y+2)(Zcz_x+2)}{16(Zcz_y+1)(Zcz_x+1)} & ; x'=y'=0, j=j', \\ 0 & ; x'=y'=0, j\neq j', \\ 0 & ; |y'| \leq 4n_y-2, \\ & |x'| \leq 4n_x-2, j \neq j'. \end{cases} \quad (8)$$

B. Proposed Method

Let $(m_{N_1}, \hat{m}_{N_1})$ be a perfect sequence pair of length N_1 whose autocorrelation function is impulse, which can be expressed as

$$\begin{cases} m_{N_1} &= (m_{N_1,0}, \cdots, m_{N_1,i}, \cdots, m_{N_1,N_1-1}), \\ \hat{m}_{N_1} &= (\hat{m}_{N_1,0}, \cdots, \hat{m}_{N_1,i}, \cdots, \hat{m}_{N_1,N_1-1}), \end{cases} \quad (9)$$

where a binary sequence m_{N_1} is a Legendre sequence [4] of length $N_1 = 4n_1-1$ or an M-sequence of length $N_1 = 2^{n_1+1}-1$ with positive n_1, $\sum_{i=0}^{N_1-1} m_{N_1,i} = 1$, $m_{N_1,i} \in \{1,-1\}$ and $\hat{m}_{N_1,i} = (1+m_{N_1,i})/2 \in \{1,0\}$. Let $(b_{N_{1y},N_{1x}}, \hat{b}_{N_{1y},N_{1x}})$ be a 2D perfect sequence pair of length $N_{1y}\times N_{1x}$, which can be expressed as

$$\begin{cases} b_{N_{1y},N_{1x}} &= [b_{N_{1y},N_{1x},y,x}] = [m_{N_{1y},y} \cdot m_{N_{1x},x}], \\ \hat{b}_{N_{1y},N_{1x}} &= [\hat{b}_{N_{1y},N_{1x},y,x}] = [\hat{m}_{N_{1y},y} \cdot \hat{m}_{N_{1x},x}]. \end{cases} \quad (10)$$

On the other hand, let $\mathbf{H}_{N_2}$ be an Hadamard matrix of length $N_2\times N_2$, which can be expressed as

$$\mathbf{H}_{N_2} = \left[h_{N_2}^0, \cdots, h_{N_2}^j, \cdots, h_{N_2}^{N_2-1}\right]^T, \quad (11)$$

$$h_{N_2}^j = \left(h_{N_2,0}^j, \cdots, h_{N_2,i}^j, \cdots, h_{N_2,N_2-1}^j\right), \quad (12)$$

where $N_2 = 2^{n_2}$ with positive n_2, $h_{N_2,i}^j \in \{1,-1\}$, $\sum_{i=0}^{N_2-1} h_{N_2,i}^j = 0$ and T denotes the matrix transposition, and $h_{N_2}^j$ is called a Hadamard sequence. Let $(c_{N_{2y},N_{2x}}^j, \hat{c}_{N_{2y},N_{2x}}^{j,d})$ be a 2D optical Hadamard sequence pair of length $N_{2y}\times N_{2x}$ and family size $M = (N_{2y}-1)(N_{2x}-1)$, which can be expressed as

$$\begin{cases} c_{N_{2y},N_{2x}}^j &= [c_{N_{2y},N_{2x},y,x}^j] = [h_{N_{2y},y}^{j_y} \cdot h_{N_{2x},x}^{j_x}], \\ \hat{c}_{N_{2y},N_{2x}}^{j,d} &= [\hat{c}_{N_{2y},N_{2x},y,x}^{j,d}] = [\frac{1+(-1)^d c_{N_{2y},N_{2x},y,x}^j}{2}], \end{cases} \quad (13)$$

where $j = (N_{2x}-1)(j_y-1)+(j_x-1)$, $1 \leq j_y \leq N_{2y}-1$, $1 \leq j_x \leq N_{2x}-1$ and $0 \leq j \leq M-1$. Finally, a 2D optical ZCZ sequence of ZCZ size $(4n_{1y}-2)\times(4n_{1x}-2)$ is obtained as a product of a 2D perfect sequence pair of (10) and a 2D optical Hadamard sequence pair of (13), is given by

$$z_{N_y,N_x}^j = [z_{N_y,N_x,y,x}^j]$$
$$= [b_{N_{1y},N_{1x},y \bmod N_{1y},x \bmod N_{1x}} \cdot c_{N_{2y},N_{2x},y \bmod N_{2y},x \bmod N_{2x}}^j], \quad (14)$$

$$\hat{z}_{N_y,N_x}^{j,d} = [\hat{z}_{N_y,N_x,y,x}^{j,d}]$$
$$= [\hat{b}_{N_{1y},N_{1x},y \bmod N_{1y},x \bmod N_{1x}} \cdot \hat{c}_{N_{2y},N_{2x},y \bmod N_{2y},x \bmod N_{2x}}^{j,d}]. \quad (15)$$

The periodic 2D correlation function of the optical 2D ZCZ sequence is given by

$$\rho_{z_{N_y,N_x}^j, \hat{z}_{N_y,N_x}^{j',d}, y', x'} = \begin{cases} \frac{(-1)^d N_yN_x(Zcz_y+2)(Zcz_x+2)}{8(Zcz_y+1)(Zcz_x+1)} & ; x'=y'=0, j=j', \\ 0 & ; x'=y'=0, j\neq j', \\ 0 & ; |y'| \leq 4n_y-2, \\ & |x'| \leq 4n_x-2, j \neq j'. \end{cases} \quad (16)$$

Therefore, from (8) and (16), the peak value of the autocorrelation function of the sequence by the proposed method can be doubled compared with that of the sequence by the conventional method.

IV. Conclusion

In this paper, we have proposed a new generation method of a two-dimensional (2D) optical zero-correlation zone (ZCZ) sequence with the size of ZCZ $(4n-2)\times(4n-2)$. By using the proposed method, we can improve the peak value of the autocorrelation function of the sequence. This means that the BER performance of the OCDMA system using the sequence can be improved.

References

[1] T. Matsumoto, H. Torii, Y. Ida, and S. Matsufuji, "A generation method of a two-dimensional optical ZCZ sequence with the smallest zero-correlation zone," Proc. of the 2019 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing, pp. 172–175, 2019.

[2] R. Shimizu, K. Nomura, T. Matsumoto, H. Torii, Y. Ida, and S. Matsufuji, "Research on optical CDMA system using two-dimensional optical ZCZ sequence set," Poc. of the 2019 International Workshop on Smart Info-Media Systems in Asia, pp. 109–114, 2019.

[3] S. Matsufuji, T. Matsumoto, Y. Tanada, and N. Kuroyanagi, "ZCZ Codes for ASK-CDMA System," IEICE Trans. Fundamentals, vol. E89-A, no.9, pp. 2268–2274, 2006.

[4] P. Z. Fan, and M. Darnell, "Sequence Design for Communications Applications," Research Studies Press, 1996.

Repetition Control Method Using Terminal Mobility for Uplink Grant-Free URLLC

Shinichi OZAKU, Yukiko SHIMBO, Hirofumi SUGANUMA, and Fumiaki MAEHARA
Graduate School of Fundamental Science and Engineering, Waseda University
3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169-8555, Japan
E-mail: fumiaki_m@waseda.jp

***Abstract*—This paper proposes a repetition control method using terminal mobility to alleviate the packet collision for uplink grant-free (GF) ultra-reliable and low-latency communications (URLLC). Because the time diversity benefit within repetitions is basically dependent on a fading correlation, the proposed approach determines the number of necessary repetitions using a fading correlation easily obtained from the terminal speed. Using the proposed repetition control, the number of repetitions can be reduced especially in low fading correlations, alleviating packet collision. The effectiveness of our proposed method is demonstrated in terms of latency by computer simulations.**

I. Introduction

Demand for mobile broadband services has increased exponentially and fifth-generation mobile communications systems (5G) are slated to provide diversified services, including internet of things (IoT) and intelligent transport system (ITS) applications [1]. In 5G, ultra-reliable and low-latency communications (URLLC) are considered to be a new use case especially for ITS, which requires a target block error rate of 1×10^{-5} or less within 1 ms [2].

For uplink URLLC, grant-free (GF) transmission is considered to be an attractive approach because latency can be reduced by eliminating scheduling requests and dynamic grant required for grant-based transmission [3]. Moreover, to improve reliability with low latency, the K-repetitions method which transmits the same packet a predetermined number of times until a feedback from the base station (BS) has been adopted for 3GPP Release 15 [4]. However, this approach has a potential problem in that the improved reliability is restricted by packet collision among mobile stations (MSs) with excessive repetitions [5]. Therefore, the new method must reduce the packet collision while retaining the improved reliability.

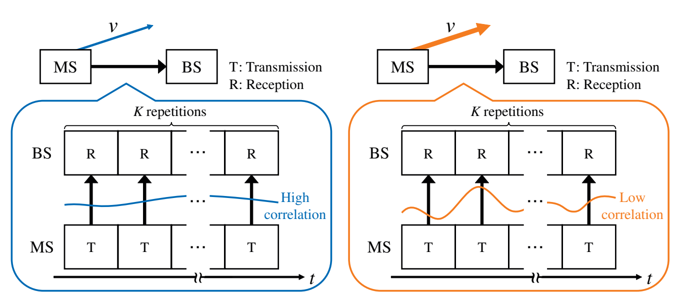
In this paper, we propose a repetition control method using terminal mobility to alleviate packet collision for uplink GF URLLC. The proposed method is inspired by the fact that the effect of the K-repetitions depends on fading correlation determined by terminal mobility. Considering the recent advances in sensor-related technologies in mobile terminals, including smart phones and vehicles, terminal speed can be easily and precisely detected, which can be used to theoretically derive the fading correlation. Taking advantage of this correlation, the proposed method reduces the number of repetitions under a relatively low fading correlation, contributing to packet collision reduction. The effectiveness of the proposed method is demonstrated in terms of latency by computer simulations.

II. Proposed Method

In the K-repetitions method for uplink GF URLLC [4], a MS generally transmits the same packet K times until a feedback is received from the BS, which improves reliability with a time diversity benefit. However, if the repetitions are excessive, intra-cell interference increases due to packet collision among the MSs, resulting in decreased reliability. Thus, excessive repetitions must be prevented while retaining the improved reliability. Herein, we propose a repetition control method using terminal mobility to reduce packet collision. The proposed approach determines the number of necessary repetitions using a fading correlation obtained from the terminal speed because the time diversity benefit depends on the fading correlation.

Figure 1 shows the K-repetitions method under a different terminal mobility. As shown in Fig. 1(b), high terminal mobility causes large fading variation within the K-repetitions, which can exploit the large time diversity benefit. In this case, the number of repetitions K can be reduced while retaining reliability.

Figure 2 shows the overall concept of the proposed method, where T_s is the symbol duration. The MS calculates the fading correlation from the terminal speed and determines the number of repetitions K using the correlation. Assuming Rayleigh fading channel, the fading correlation is given by $\rho(\Delta t) = J_0(2\pi f_D \Delta t)$ [6], where Δt and $J_0(\cdot)$ denote the time difference and Bessel function of the first kind of order zero, respectively. Moreover, the maximum Doppler frequency f_D



(a) Low terminal mobility (b) High terminal mobility

Fig. 1. K-repetitions method under a different terminal mobility.

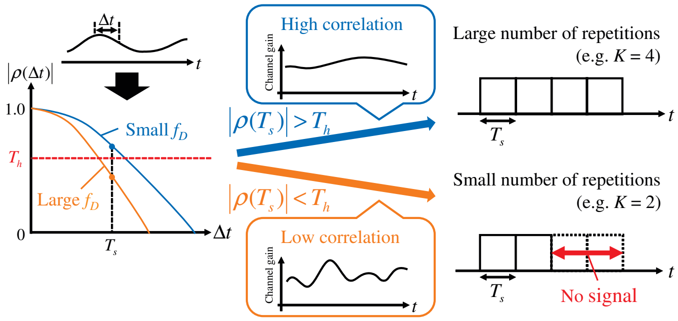


Fig. 2. Concept of the proposed method.

TABLE I
Simulation Parameters

Modulation	QPSK / OFDM
Number of sub-carriers N	512
Guard interval length T_G	64
Forward error correction (FEC)	Turbo coding / Max-log-MAP decoding ($R = 1/2$)
Channel model	16-ray exponentially decaying Rayleigh fading
Delay spread τ_{rms}	$1.0T_{sam}$
Normalized maximum Doppler frequency $f_D T_s$	0.001–0.2
Packet combining scheme	Chase combining
Combining criterion	Maximal ratio combining (MRC)
Channel estimation	Perfect

is determined by the terminal speed. In the proposed method, if the fading correlation $|\rho(T_s)|$ exceeds the threshold T_h, a large number of repetitions are required (e.g. $K = 4$) because the time diversity benefit cannot be exploited. In contrast, if $|\rho(T_s)|$ is less than T_h, the number of repetitions can be reduced (e.g. $K = 2$) thanks to the large time diversity benefit. In this manner, the proposed method properly controls the number of repetitions to prevent packet collision.

III. Numerical Results

In this section, we verify the effectiveness of the proposed method in comparison with the traditional K-repetitions method by computer simulations. Table I shows the simulation parameters used herein. In the performance evaluation, the maximum Doppler frequency normalized by OFDM symbol duration $f_D T_s$ is uniformly distributed from 0.001 to 0.2 and the average CNR is set to be 16 dB.

Figure 3 shows the complementary cumulative distribution function (CCDF) of latency normalized by T_s in the proposed method. It can be seen from Fig. 3 that our proposed method improves the latency performance with decreased threshold T_h. Table II shows the use rate of the number of repetitions K in each threshold T_h. The proposed method increases the use rate of $K = 2$ by 17.5% while retaining the target block error rate of 1×10^{-5} and latency within $4T_s$.

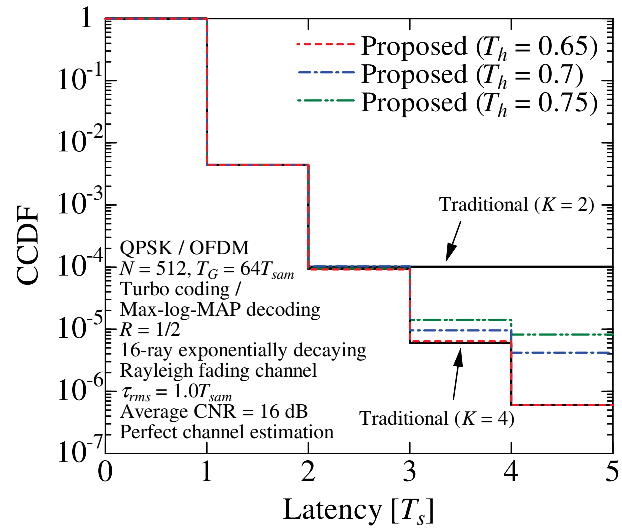


Fig. 3. Complementary cumulative distribution function (CCDF) of latency.

TABLE II
Use Rate of the Number of Repetitions K

	$K = 2$	$K = 4$
Traditional ($T_h = 0.0$)	0.0%	100.0%
Proposed ($T_h = 0.65$)	1.0%	99.0%
Proposed ($T_h = 0.70$)	9.0%	91.0%
Proposed ($T_h = 0.75$)	17.5%	82.5%

IV. Conclusion

In this paper, we proposed a repetition control method which featured terminal mobility for uplink GF URLLC. The number of necessary repetitions is determined using a fading correlation easily obtained from the terminal speed, which can prevent excessive repetitions especially under high terminal mobility. The numerical results showed that the method proposed herein can reduce the number of repetitions by setting the appropriate threshold for the fading correlation while achieving the target block error rate and latency. Hence, the proposed method has the potential to prevent packet collision among MSs.

References

[1] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, “5G: A tutorial overview of standards, trials, challenges, deployment, and practice,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[2] 3GPP TR 38.913 v14.2.0, “Study on scenarios and requirements for next generation access technologies,” Mar. 2017.

[3] C. Wang, Y. Chen, Y. Wu, and L. Zhang, “Performance evaluation of grant-free transmission for uplink URLLC services,” *Proc. IEEE 85th Veh. Technol. Conf. (VTC 2017-Spring)*, pp. 1–6, June 2017.

[4] 3GPP TS 38.214 v15.2.0, “NR; Physical layer procedures for data,” June 2018.

[5] T. Jacobsen, R. B. Abreu, G. Berardinelli, K. I. Pedersen, P. E. Mogensen, I. Koves, and T. Kozlova, “System level analysis of uplink grant-free transmission for URLLC,” *2017 IEEE Globecom Workshops (GC Wkshps 2017)*, pp. 1–6, Dec. 2017.

[6] J . G. Proakis and M. Salehi, *Digital Communications*, 5th ed., McGraw-Hill, 2007.

BER Detection of A2G Wireless Communication in Rician *K*-Factor Fading Channel for Massive IoT Connectivity Network

Sarun Duangsuwan, Anna Chusongsang* Chakree Teekapakvisit and Sathaporn Promwong*
Information Engineering, Department of Engineering
King Mongkut's Institute of Technology Ladkrabang Prince of Chumphon Campus
**Telecommunication Engineering, Faculty of Engineering*
King Mongkut's Institute of Technology Ladkrabang
1 Chalongkrung Rd., Ladkrabang, Bangkok, Thailand
sarun.du@kmitl.ac.th, annaannoy@hotmail.com, t.chakree@gmail.com, kpsathap@kmitl.ac.th

***Abstract*—In this paper, we present a performance of bit error rate detection in air-to-ground (A2G) wireless communication over Rician *K*-factor fading channel for massive the internet of things (IoT) connectivity network. The massive IoT connectivity deploys to the cyber-physical system (CPS). We analyze the BER by using the asymptotic simulation of zero-forcing (ZF) and minimum mean square error (MMSE) detection. The system model employs the effect of Rician *K*-factor channel between the transmitter antenna (drone) and IoT users. Finally, we discuss the performance of BER curves.**

Keywords—A2G wireless communication, BER detection, Massive IoT connectivity network, Rician K-factor channel, Public safety.

I. INTRODUCTION

The massive IoT connectivity is a network of mobile broadband (MBB), ultra-reliable low-latency communication (URLLC) and massive machine type communication (MTC). MBB is a natural unfolding of long-term evolution (LTE), where the primary goal is to increase the user data rate. In addition, URLLC and MTC, are the cornerstones of machine-type traffic and thus enablers of various types of IoT connectivity [1].

In surveillance of smart city, unmanned aerial vehicles (UAV) or drone become an important role, it is used not only the public safety, but also to provide for wireless link communication of services. Drone in public safety scenario was illustrated in [2] (therein the figure 2). In order to secure in the natural disasters, drone can be used to be cellular base station and ground communications infrastructure. In such scenarios, there is a vital need for public safety communication between first responder and victims for search and rescue operation. Fig 1 shows drone for IoT connectivity networks such as in smart city, smart farming and smart industrial, these are challenging to apply drone for a new communication link [2].

To challenge the research work, wireless propagation is affected by the medium between the transmitter and the receiver. The channel model is called that A2G wireless communication has been interestingly studied. In [3], they presented an overview of existing research related to A2G channel modeling. In [4], the authors provided both simulation and measurement results for path loss, delay spread, and fading in A2G wireless communication. While the path loss model is derived to A2G channel modeling as expressed in [4], it is worth noting that the small-scale fading in A2G wireless communication can be characterized by Rician fading channel model. The Rician *K*-factor that represents the strength of line of sight (LOS) component is a function of elevation angle and the drone altitude.

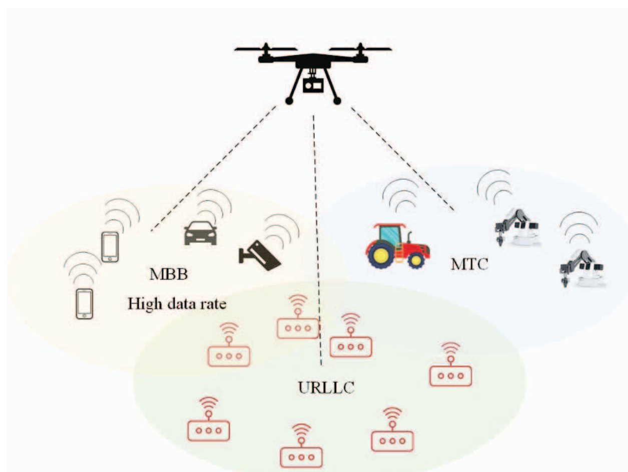


Fig. 1. Model of Drone based IoT connectivity network.

In this paper, we consider the performance of detection of uplink A2G wireless communication. By using the ZF and MMSE detector, we analyze the BER performance over Rician *K*-factor fading channel.

The remainder of the paper is organized as follows. In Section II, the A2G communication over Rician fading channel model is presented. The BER result presents in Section III, and conclusion of the paper in Section IV.

II. A2G Wireless Communication over Rician K-Factor Fading Channel

A. Signal Model

The uplink received signal of drone is

$$\begin{aligned}\mathbf{y}_{(i)} &= \sum_{i=1}^{N} \mathbf{h}_{(i)} \sqrt{p_t}\, x_{(i)} + \mathbf{n}_{(i)} \\ &= \sqrt{p_t}\, \mathbf{H}_{(i)} \mathbf{x}_{(i)} + \mathbf{n}_{(i)}\end{aligned} \tag{1}$$

where p_t is the average power constraint, which depends on the path loss constant and distance of air-to-ground [4]. $x_{(i)}$ is the transmitted signal. $\mathbf{n}_{(i)}$ is the noise component, whose elements are complex Gaussian random variables with zero mean and unit variance.

The uplink channel gain $\mathbf{H}_{(i)} = \left[\mathbf{h}_{(1)} \cdots \mathbf{h}_{(N)}\right]$ can be re-written in term of Rician K-factor fading as

$$\mathbf{H}_{(i)} = \sqrt{\frac{K}{K+1}}\hat{\mathbf{H}}_{(LOS)} + \sqrt{\frac{1}{K+1}}\hat{\mathbf{H}}_{(NLOS)} \tag{2}$$

B. ZF Detection

The optimal detection matrix [1] is

$$\hat{\mathbf{y}} = \left(\mathbf{H}_{(i)}^{H}\mathbf{H}_{(i)}\right)^{-1} \mathbf{H}_{(i)}^{H}\mathbf{y}_{(i)} \tag{3}$$

where $(\cdot)^H$ is the matrix transpose.

C. MMSE Detection

The optimal detection matrix [1] can written as

$$\hat{\mathbf{y}} = \left(\mathbf{H}_{(i)}^{H}\mathbf{H}_{(i)} + \gamma \mathbf{I}\right)^{-1} \mathbf{H}_{(i)}^{H}\mathbf{y}_{(i)} \tag{4}$$

where γ is the uplink signal to noise ratio (SNR) and $\mathbf{I}$ denotes the identical matrix.

III. Simulation Result

This section describes the simulation result of BER performance by using ZF and MMSE detection. Herein, we perform quadrature phase shift keying (QPSK) modulation, the conditional error probability is

$$P\left(e \mid \hat{\mathbf{y}}\right) = a\,\mathrm{erfc}\left(\sqrt{b\hat{\mathbf{y}}}\right) \tag{5}$$

where a = 0.5 and b = 0.5, therein [3]. The $\mathrm{erfc}(x)$ is the error function. Additionally, the complementary error function can be obtained as

$$\mathrm{erfc}(x) = \left(\frac{2}{\sqrt{\pi}}\right)\int_{x}^{\infty} \exp\left(-x^2\right) dx \tag{6}$$

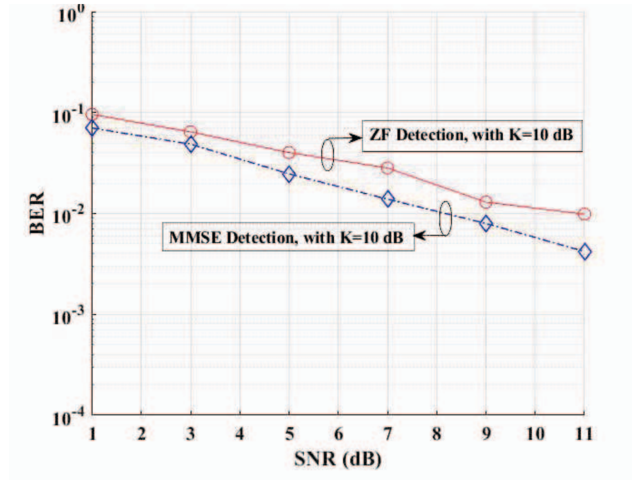


Fig. 2. BER performance by using ZF and MMSE, where K = 10 dB.

The simulation result of BER performance is shown in Fig. 2, where K = 10 dB. Herein, our scenario of the simulation was performed by using Monte Carlo simulation, where n = 1500 users of IoT connectivity network and we assume that the distance of air-to-ground was 400 m. In Fig. 2, we observe that the average of BER both ZF and MMSE detection is level as 10^{-2} in Rician K-factor fading channel. Additionally, the MMSE detection is optimal better than ZF detection. However, it depends on the distance of air-to-ground, path loss and transmits power constraint.

IV. Conclusion

This paper has introduced the BER detection by using the existing of linear detection to analyze the performance of A2G wireless communication over Rician K-factor fading channel. We confirm that ZF and MMSE detection can recover the communication system of A2G, the distance is higher than 200 m., as well as BER performance has been discussed. In future work, the channel estimation method will be considered in A2G channel model.

References

[1] B. M. Lee, and H. Yang, "Massive MIMO for industrial internet of things in Cyber-physical systems," *IEEE Transactions on Industrial Informatics*, pp. 2631 – 2652, vol. 14, no. 6, Jun. 2018.

[2] M. Mazaffari, W. Saad, M. Bennis, Y-H. Nam, M. Debbah, "A tutorial on UAV for wireless networks: Applications, challenges, and open problem," *IEEE Communications Surveys & Tutorials*, pp. 2334 – 2360, vol. 21, no. 3, 2019.

[3] A. M. Hayajneh, S. A. R. Zaidi, D. C. Mclernon, and M. Ghogho, "Optimal dimensioning and performance analysis of drone-based wireless commuications," in *IEEE Globecom Workshops*, Washington, DC, USA, Dec. 2016.

[4] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platform in urban environment," in *IEEE Globecom Workshops*, Austin, TX, USA, Dec. 2014.

Performance Evaluation of IDMA-Based Random Access Considering User Detection and Channel Estimation

Taiki Tomizawa†, Yoshihisa Kishiyama‡, and Kenichi Higuchi†
†Graduate School of Science and Technology, Tokyo University of Science
‡5G Laboratory, NTT DOCOMO, INC.

Abstract—**This paper evaluates the performance of interleave division multiple access (IDMA)-based random access in order to support massive machine-type communications (mMTC) in the fifth generation mobile communication system taking into account actual user detection and channel estimation using a preamble attached to the packet. To suppress the packet loss due to contention and achieve multi-packet reception, accurate user detection and channel estimation, which are used in the successive interference canceller (SIC)process in IDMA, are essential. Multiple preamble sequences, each of which has a one-to-one relationship with the interleaver pattern prepared in the system, are generated by applying a cyclic time shift and orthogonal masking operations to Zadoff-Chu (ZC) sequences. Based on the threshold decision on the correlation between the received signal and each preamble sequence, user detection and channel estimation are performed. Computer simulation results show the impact of the preamble overhead, miss detection and false alarm in the user detection, channel estimation error, and the case when multiple users simultaneously select the same interleaver on the achievable performance of IDMA-based random access.**

I. Introduction

In the fifth-generation (5G) mobile communication system [1], massive machine-type communications (mMTC) will be supported to actualize the Internet-of-Things (IoT). To support the mMTC uplink, a grant-free and contention-based multiple access scheme is essential to reduce the control signaling overhead and transmission latency. To suppress packet loss due to collisions and to achieve multi-packet reception, non-orthogonal multiple access (NOMA) with interference cancellation [2] at the base station receiver is essential. In this paper, we focus on interleave division multiple access (IDMA) [3] based random access. In IDMA, multiple packets transmitted simultaneously from different terminals are decomposed by multiuser detection such as by using a successive interference canceller (SIC) at the base station receiver in which terminal-specific channel interleavers are utilized. The IDMA signal does not increase the peak power and its transmission performance is robust against the error in time/frequency synchronization among terminals. These are desirable properties in the mMTC uplink.

In IDMA-based random access, to suppress the packet loss due to collision and to achieve multi-packet reception, accurate user detection, in which the set of users who transmit the packet are identified, and channel estimation for detected user are essential by using preamble attached to the packet. When miss detection occurs in user detection, the received packet of the miss detected user is directly discarded. On the other hand, when a false alarm occurs where a user who has not sent a packet is incorrectly determined to have sent a packet, additional useless packet decoding is performed in the SIC, which causes an increase in the computational complexity and inter-packet interference.

In this paper, multiple preamble sequences, each of which has a one-to-one relationship with the interleaver pattern prepared in the system, are generated by applying a cyclic time shift and orthogonal masking operations to Zadoff-Chu (ZC) sequences [4]. Based on the threshold decision on the correlation between the received signal and each preamble sequence, user detection and channel estimation are performed. Computer simulation results show the impact of the preamble overhead, miss detection and false alarms in the user detection, channel estimation error, and the case when multiple users simultaneously select the same interleaver on the achievable performance of IDMA-based random access.

The remainder of the paper is organized as follows. First, Section II describes the preamble design and methods for channel estimation and user detection. Section III presents numerical results based on computer simulations. Section IV concludes the paper.

II. Preamble Sequence Design and Methods for User Detection and Channel Estimation

We assume a discrete Fourier transform (DFT)-spread orthogonal frequency division multiplexing (OFDM)-based single-carrier transmission, and one packet consists of $B_{\rm t}$ DFT blocks. Out of the $B_{\rm t}$ DFT blocks, $B_{\rm p}$ DFT blocks are used for transmitting the preamble and the remaining $B_{\rm d}$ (= $B_{\rm t} - B_{\rm p}$) DFT blocks are used for transmitting encoded data. As $B_{\rm p}$ is increased, the accuracies of the channel estimation and user detection are improved at the cost of increasing the channel coding rate of the packet data.

We use ZC sequences as a preamble because they have a good correlation property. ZC sequences of length N can be represented as

$$a_u(i) = \exp\left(\frac{-j2\pi u}{N}\cdot\frac{i(i+1)}{2}\right),\quad i = 0,1,...,N-1 \tag{1}$$

where u denotes the sequence index, which is an arbitrary integer relatively prime to N. When N is a prime number, the number of available sequences is N-1. Note that N is basically the same as the DFT size, and a preamble based on ZC sequences of length N is transmitted in one DFT block.

In order to increase the number of orthogonal preamble sequences, a cyclic time shift and orthogonal masking operations are applied to the ZC sequences. Since the autocorrelation of the ZC sequence of length N is an ideal delta function, we can generate $\lfloor N/\Delta \rfloor$ orthogonal preamble sequences from one ZC sequence by applying a cyclic time shift with an integer multiple of Δ samples. We note that Δ should be greater than the received timing difference among user terminals to distinguish the received multipath from different users.

Orthogonal masking is applied when $B_{\rm p}$ is greater than one. Orthogonal masking coefficient $o_n(m)$ is represented as $\exp(-j2\pi(n-1)(m-1)/M)$ for $n = 1, \ldots, B_{\rm p}$ and $m = 1, \ldots, B_{\rm p}$. We assume that the preamble sequence with length N generated from the u-th ZC sequence with the l-th cyclic shift is denoted as $a_{u,l}(i)$. By using orthogonal masking, from the same $a_{u,l}(i)$, $B_{\rm p}$ preamble sequences with length $B_{\rm p}N$ are generated as

$$\begin{array}{cccc} & \text{For DFT block 1} & & \text{For DFT block } B_{\rm p} \\ \tilde{a}_{u,l,1} = & o_1(1)\cdot\{a_{u,l}(i)\} & \cdots & o_1(B_{\rm p})\cdot\{a_{u,l}(i)\} \\ \vdots & & & \\ \tilde{a}_{u,l,B_{\rm p}} = & o_{B_{\rm p}}(1)\cdot\{a_{u,l}(i)\} & \cdots & o_{B_{\rm p}}(B_{\rm p})\cdot\{a_{u,l}(i)\} \end{array}. \tag{2}$$

Channel estimation and user detection using the preamble are performed in the time domain. First, we take the correlation between the time-domain received signal and candidate preamble sequences. When B_p is greater than one, the correlation values between multiple DFT blocks are coherently combined with orthogonal masking coefficient $o_n(m)$ to suppress the noise. We assume that the squared norm of the correlation at the received timing, τ (τ = 0, …, $\Delta-1$), for the preamble based on the u-th ZC sequence with the l-th cyclic shift and m-th orthogonal masking is denoted as $R_{u,l,m}[\tau]$. Term $\bar{R}$ is the average of $R_{u,l,m}[\tau]$.

User detection is performed as described hereafter. Threshold T_UD for user detection is defined as $T_\mathrm{UD} = C_\mathrm{UD}\bar{R}$ where C_UD is a positive constant. For each candidate combination of $\{u,l,m\}$, which corresponds to preamble = user index, $R_{u,l,m}[\tau]$ is sorted for τ = 0, …, $\Delta-1$ and the highest Q $R_{u,l,m}[\tau]$ values are averaged and compared to T_UD. If it exceeds T_UD, the user detection declares that the received signal contains the packet from user $\{u,l,m\}$. The user detection accuracy is expected to improve using an appropriately large Q value due to the path diversity effect.

When the packet from user $\{u,l,m\}$ is estimated to be received, the channel estimation of that packet is performed. Path detection threshold T_CE in the channel estimation is defined as $T_\mathrm{CE} = C_\mathrm{CE}\bar{R}$ where C_CE is a positive constant. We compare $R_{u,l,m}[\tau]$ with T_CE at each τ, and detected correlation values at all τ where $R_{u,l,m}[\tau]$ is lower than T_CE are set to zero to remove the noise paths. In this paper, a frequency-domain SIC [5] is used. The time-domain correlation profile after noise suppression, which corresponds to the observed channel impulse response, is converted into a frequency-domain channel response using the DFT.

III. Numerical Results

We assume a DFT-spread OFDM-based single-carrier transmission. The number of subcarriers (= DFT size) is 307 with the subcarrier spacing of 15 kHz, which corresponds to a 4.6-MHz transmission bandwidth. Term B_t is set to 7, which corresponds to the packet length of 0.5 ms including the cyclic prefix. For the channel coding, we use a combination of a rate-1/3 turbo code and repetition code. QPSK data modulation is assumed. Term B_p is parameterized. The number of repetitions in the repetition coding is set to 10, 8, 7, and 5 for B_p of 1, 2, 3, and 4, respectively, so that the number of information bits per packet is fixed to approximately 115 regardless of B_p for fair comparison. The number of interleavers prepared by the system, which is equal to the number of preambles, is set to 100. Ten users transmit packet simultaneously. Six-path block Rayleigh fading with the rms delay spread of 1 μs is assumed. Four-branch receiver antenna diversity is employed. Max-Log MAP (maximum a posteriori) decoding with eight iterations is used to decode the turbo code. The maximum number of iterations in the SIC process is set to eight.

Fig. 1 shows the average packet error rate (PER) as a function of the signal-to-noise ratio (SNR) when user detection is ideally performed. Term B_p is parameterized. Here, B_p of three achieves the lowest PER thanks to the best tradeoff between the accuracy of the channel estimation and the repetition coding rate.

Fig. 2 shows the miss detection and false alarm probabilities as a function of Q. The target false alarm probability is set to 10^{-2}, and we adjusted the C_UD value at each Q so that the target false alarm probability is achieved. The SNR is set to −13 dB. The miss detection probability is lowest when Q is approximately four to six.

Fig. 3 shows the average PER as a function of the SNR where the effect of collision in the selected interleaver among users is taken into account. The case where no collision occurs in the selected interleaver among users is also tested. The PER when ideal user detection and channel estimation are assumed is also shown as a reference. The investigated user detection and channel estimation works well. The performance degradation from the ideal case is within 2 dB. Meanwhile, the impact of the collision in the interleaver selection (in this evaluation, each of 10 users randomly selects an interleaver from 100 candidates) is rather large. This indicates that developing a strategy to lower the collision probability and reducing the complexity of user detection and channel estimation per candidate interleaver are important issues.

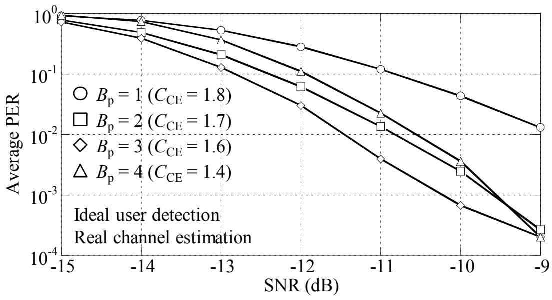


Fig. 1. Average PER as a function of SNR for various B_p.

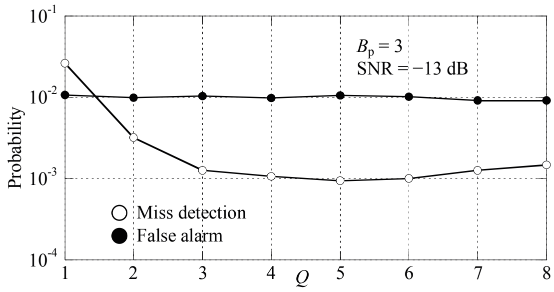


Fig. 2. Miss detection and false alarm probabilities as a function of Q.

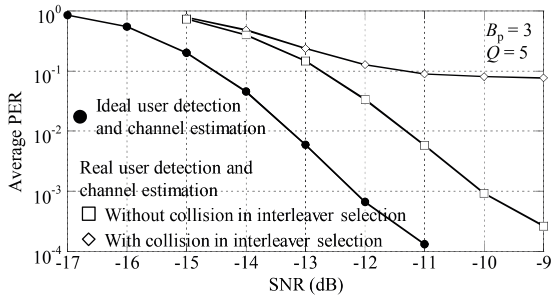


Fig. 3. Performance when collision occurs in interleaver selection.

IV. Conclusion

IDMA-based random access with the investigated preamble structure and user detection/channel estimation works well. Developing a strategy for lowering the collision probability of interleaver selection and a way to reduce the complexity of user detection and channel estimation per candidate interleaver are important issues to be addressed in the future.

References

[1] ITU-R, "IMT Vision − Framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation M.2083-0, Sept. 2015.

[2] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," IEICE Trans. Commun. vol. E98-B, no. 3, pp. 403-414, March 2015.

[3] Y. Hu, C. Xu, and L. Ping, "NOMA and IDMA in random access systems," in Proc. IEEE VTC2018-Spring, Porto, Portugal, June 2018.

[4] T. Kawamura, Y. Kishiyama, K. Higuchi, and M. Sawahashi, "Orthogonal pilot channel using combination of FDMA and CDMA in single-carrier FDMA-based evolved UTRA uplink," IEICE Trans. Commun., vol. E91-B, no. 7, pp. 2299-2309, July 2008.

[5] M. Kawata, K. Tateishi, and K. Higuchi, "Investigation on structure of interference canceller for IDMA-based random access," in Proc. IEEE VTC2018-Fall, Chicago, U.S.A., 27-30 Aug. 2018.

Image Regularization with Morphological Gradient Priors Using Optimal Structuring Elements for Each Pixel

Shoya Oohara*, Hirotaka Oka*, Mitsuji Muneyasu*, Soh Yoshida* and Makoto Nakashizuka†
*Faculty of Engineering Science, Kansai University
3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680 Japan
Email: {k818000, k889917, muneyasu, sohy}@kansai-u.ac.jp
†Faculty of Engineering, Chiba Institute of Technology
2-17-1 Tsudanuma, Narashino, Chiba, 275-0016 Japan
Email: nkszk@sky.it-chiba.ac.jp

***Abstract*—As an image prior for image restoration, the use of the sum of the morphological gradient for an image has been proposed. In this paper, we propose a method of optimizing the structuring element (SE) for each pixel, in particular, by greatly reducing the computational burden for optimization by adopting a new evaluation method for the SE in simulated annealing. By optimizing the SE for each pixel, edges of the image can be faithfully evaluated, and the improvement of restoration accuracy can be expected. An experimental result shows the effectiveness of the proposed method.**

***Index Terms*—morphology, morphological gradient, total variation, regularization, image restoration, simulated annealing**

I. Introduction

The image restoration problem in which an original image from only the observation image is estimated, such as a noisy image, is an ill-conditioned problem. Therefore, a regularization term defined from prior information of an image is used. The objective function of the restoration problem is defined by the sum of the regularization term and the fidelity term, that is, the sum of the squares of the differences between a restored image and an observed image, and the restored image is obtained by minimizing this sum.

As a regularization term, the total variation (TV) norm [1] and an image prior consisting of a morphological gradient (MG) [2] have been proposed. In the conventional method [3], the combination of the TV norm and the MG is adopted. The use of the MG by selecting an appropriate structuring element (SE) can make the regularization term more suitable for an image and may bring about an improvement in restoration accuracy. However, in these methods, only the edges of a specified direction in the whole image are evaluated. Therefore, it is ill-suited to evaluating various textures and edges in the image.

In this paper, we propose a method for restoring an image via SE optimization for each pixel. By optimizing of the SE for each pixel, it is possible to select the optimum SE for various edges. Therefore an improvement in restoration accuracy can be expected. We use simulated annealing (SA) for SE optimization. The experimental result shows the effectiveness of the proposed method in image restoration.

II. Regularization with morphological gradients

Let an $N \times N$ original image and the observed image be denoted as $\boldsymbol{g}_o$ and $\boldsymbol{y}$, respectively. The observed image is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{g}_o + \boldsymbol{e} \ , \tag{1}$$

where $\boldsymbol{H}$ is a degradation matrix and $\boldsymbol{e}$ is additive noise. The image restoration problem is the inverse problem of estimating $\boldsymbol{g}_o$ from $\boldsymbol{y}$.

Part of this research was financially supported by a JSPS Grant-in-Aid for Scientific Research (C) (No. 17K06450).

If $\boldsymbol{H}$ is known and $\boldsymbol{e}$ is Gaussian noise, the estimated image $\hat{\boldsymbol{g}}$ is obtained by solving the following regularization problem:

$$\hat{\boldsymbol{g}} = \arg\min_{\boldsymbol{g}} \left\{ \frac{1}{2}\|\boldsymbol{H}\boldsymbol{g} - \boldsymbol{y}\|_2^2 + \lambda P(\boldsymbol{g}) \right\} \ , \tag{2}$$

where $\hat{\boldsymbol{g}}$ is a restored image. In (2), the first term is a fidelity term consisting of the sum of squares of differences between the restored image and the observed image, and the second term includes an image prior $P(\boldsymbol{g})$ and a regularization parameter λ. As $\lambda P(\boldsymbol{g})$, we adopt

$$\lambda P(\boldsymbol{g}) = \lambda_{TV} P_{TV}(\boldsymbol{g}) + \lambda_{MG} P_{MG}(\boldsymbol{g}) \ , \tag{3}$$

where $P_{TV}(\boldsymbol{g})$ and $P_{MG}(\boldsymbol{g})$ are a TV norm [1] and an image prior that consists of the MG [2], respectively.

Let $\boldsymbol{S}$ be the set of coordinates constituting a flat SE. The MG, $G_{MG} \circ g_{i,j}$ in $\boldsymbol{S}$, is defined as $G_{MG} \circ g_{i,j} = D_{\boldsymbol{S}} \circ g_{i,j} - E_{\boldsymbol{S}} \circ g_{i,j}$, where the dilation $D_{\boldsymbol{S}} \circ g_{i,j}$ and erosion $E_{\boldsymbol{S}} \circ g_{i,j}$ are represented as $D_{\boldsymbol{S}} \circ g_{i,j} = \max_{(k,l)\in\boldsymbol{S}} g_{i+k,j+l}$ and $E_{\boldsymbol{S}} \circ g_{i,j} = \min_{(k,l)\in\boldsymbol{S}} g_{i+k,j+l}$, and "min" and "max" indicate the minimum and maximum values of an element, respectively. The sum of the MG is given as

$$P_{MG}(\boldsymbol{g}) = \sum_{1\leq i\leq N} \sum_{1\leq j\leq N} G_{MG} \circ g_{i,j} \ . \tag{4}$$

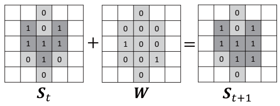
III. Proposed method

The conventional method [3] employed a genetic algorithm (GA) to optimize an SE, however, the GA cannot be applied for the proposed method because of its large computational burden. In particular, an evaluation of the SE (the calculation of fitness) in [3] requires much computation. To reduce the computational burden, we employ SA instead of the GA, and a new method for the evaluation of the SE is proposed.

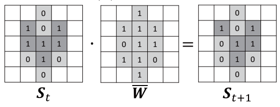
A. New evaluation method in SE optimization

1) Definition of solution space: Exploration of the search space by SA is insufficient compared with that by the GA. Therefore, a probability of finding the optimal solution in the solution space is lower in SA. If a restricted search space that includes the optimal solution can be given, we may be able to find the optimal solution by SA. In the proposed method, we assume that the value of the MG obtained using the most appropriate SE is lower than that obtained using other SEs.

Under this assumption, we set the search range as follows. The two-pixel SE, which has a pixel at the center position and the other pixel in an $L \times L$ window, is considered. We calculate the sum of the MG of an image using each SE and select M SEs in descending order. The search range is created by superimposing these SEs.



(a) Case 1



(b) Case 2

Fig. 1: Examples of generating solution candidate.

2) Evaluation of SE: To reduce the computational burden, the evaluation method of the SE is changed. First, we obtain an initial estimation image, $\tilde{\boldsymbol{g}}$, by minimizing

$$1/2 \cdot \|\boldsymbol{Hg} - \boldsymbol{y}\|_2^2 + \lambda P_{TV}(\boldsymbol{g}) \ . \tag{5}$$

Using the initial estimation image, the evaluation value is obtained by

$$E(\boldsymbol{S}_{i,j}) = \frac{1}{2}\|\boldsymbol{H}\tilde{\boldsymbol{g}}_{i,j} - \boldsymbol{y}_{i,j}\|_2^2 + \lambda_{TV} P_{TV}(\tilde{\boldsymbol{g}}_{i,j}) + \lambda_{MG} P_{MG}(\tilde{\boldsymbol{g}}_{i,j}, \boldsymbol{S}_{i,j}), \tag{6}$$

where $\boldsymbol{S}_{i,j}$ is the SE for the local region whose center is (i,j), $\tilde{\boldsymbol{g}}_{i,j}$ is a $D \times D$ local region of $\tilde{\boldsymbol{g}}$, centered around (i,j), and $\boldsymbol{y}_{i,j}$ is also a $D \times D$ local region of the observed image $\boldsymbol{y}$, centered around (i,j). In (6), the first and second terms are made constants by fixing $\tilde{\boldsymbol{g}}_{i,j}$. Then, we evaluate $E(\boldsymbol{S}_{i,j}) = P_{MG}(\tilde{\boldsymbol{g}}_{i,j}, \boldsymbol{S}_{i,j})$.

3) Generation of solution candidate: We describe a method of generating a solution candidate. For a current candidate, we select one of the following three cases with equal probability: 1) add the pixel to the current candidate; 2) remove the pixel from the current candidate; or 3) change only the position of pixels in the current candidate.

In cases 1 and 2, the mask array is used for generating the solution candidate. For each element in the search range, the probability p_m, which indicates whether or not each element is included in the solution candidate, is given by $p_m = \frac{V - G_m}{\sum_{k=1}^{M}(V - G_k)}$, where $V = \sum_{k=1}^{M} G_k$ and G_m is the sum of the MG using the SE that consists of the center pixel and mth pixel.

Using p_m, an $L \times L$ mask array $\boldsymbol{W}$ for generating the solution candidate is considered. The value of the mask at index m, $\{\boldsymbol{W}\}_m$, is defined as

$$\{\boldsymbol{W}\}_m = \begin{cases} 1, & r_m \le p_m \\ 0, & r_m > p_m \end{cases} , \tag{7}$$

where r_m is a random number that satisfies $0 \le r_m \le 1$. The solution candidate $\boldsymbol{S}_{t+1}$ is generated from the current candidate $\boldsymbol{S}_t$ using $\boldsymbol{W}$. For case 1, $\boldsymbol{S}_{t+1}$ is given by the logical sum of $\boldsymbol{S}_t$ and $\boldsymbol{W}$ for each element. For case 2, $\boldsymbol{S}_{t+1}$ is given by the logical product of $\boldsymbol{S}_t$ and $\boldsymbol{W}$ for each element. Examples of both cases for $M = 10$ are shown in Fig. 1.

B. Algorithm of image restoration

The proposed method consists of the optimization of the SE by SA and the image restoration for each pixel using the obtained SE. The procedure for the optimization of the SE is shown.

1) The solution space is defined by the method above. Let $t = 1$ and $k = 1$. The inital temperature T_1 and the initial solution candidate $\boldsymbol{S}_1$ are given. Equation (5) is minimized to obtain an initial image, and the evaluation value E_1 of $\boldsymbol{S}_1$ is calculated using this initial image.
2) According to the method above, a new solution candidate $\boldsymbol{S}_{t+1}$ is generated and its evaluation value is calculated.
3) The acceptance rate of $\boldsymbol{S}_{t+1}$, $A(E_t, E_{t+1}, T_k)$, is calculated using the Metropolis standard given as
$$A(E_t, E_{t+1}, T_k) = \begin{cases} 1, & E_{t+1} - E_t < 0 \\ \exp(-\frac{E_{t+1} - E_t}{T_k}), & Otherwise \end{cases} . \tag{8}$$
If the solution candidate is accepted, then go to 4), otherwise let $\boldsymbol{S}_{t+1} = \boldsymbol{S}_t$, $E_{t+1} = E_t$, and go to 4).
4) Let $t = t + 1$. If the number of iterations is within the predetermined threshold, then let $T_{k+1} = \gamma T_k$, $k = k + 1$, and go to 5). Otherwise, go to 2).
5) If the termination condition is satisfied, the algorithm is terminated. If not, go to 2).

After the optimization of the SE for each pixel, we carry out the image restoration for each pixel using the obtained SE. That is, the image of the $D \times D$ local region $\tilde{\boldsymbol{g}}_{i,j}$ is restored by minimizing (6) using the obtained SE. Let the center pixel value of $\tilde{\boldsymbol{g}}_{i,j}$ be the pixel value at (i,j) in the whole restored image; then we obtain the whole restored image.

TABLE I: MSEs of restored images in the experimental result.

Image	TV [1]	[3]	Proposed
Lenna	53.24	50.61	50.34
Barbara	131.9	129.0	106.0
Hill	71.69	71.65	69.97
Bark	167.4	167.5	167.3
Brick wall	116.8	112.8	110.5
Straw	205.7	188.9	181.1

IV. Experiment

Experiments were conducted to demonstrate the effectiveness of the proposed method. In this experiment, we treated the restoration problem for the images degraded by Gaussian noise. The images used for the experiment were grayscale images of 256 levels with a size of 512×512 pixels. As additive noise, white Gaussian noise with mean $\mu = 0$ and standard deviation $\sigma = 20$ was used. The parameters were set to $D = 25$, $M = 10$, $T_K = 1000$, $L = 5$, and $\gamma = 0.8$. Table I shows the mean square error (MSE) of the restored images in the experimental result. It is found that the restoration accuracy improved for all images.

V. Conclusions

In this paper, we proposed a method of SE optimization for each pixel by SA. The proposed method can restore an image while preserving the signal of various textures, edges and details. By optimizing the SE for the local regions that include each pixel, we obtained excellent results compared with the conventional method in an experiment of reducing of Gaussian noise.

References

[1] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear Total Variation Based Noise Removal Algorithm,” Physica. D, Nonlinear Phenomena, vol. 60, nos. 1–4, pp. 259–268, Jan. 1992.

[2] M. Nakashizuka, “Image Regularization with Higher-Order Morphological Gradient,” Proc. of 2015 23rd European Signal Processing Conference, pp. 1820–1824, Nice, France, Aug. 31–Sept. 4, 2015.

[3] S. Oohara, M. Muneyasu, S. Yoshida, and M. Nakashizuka, “Image Regularization with Morphological Gradient Priors Using Optimization of Structuring Element,” Proc. 2018 International Symposium on Intelligent Signal Processing and Communication Systems, Ishigaki, Okinawa, Nov. 27–31, pp. 498–503, 2018.

Atrial Fibrillation Detection in Spectrogram Based on Convolution Neural Networks

1st Jing-Ming Guo
Department of Electrical Engineering
National Taiwan University of Science and Technology
Taiwan, Taipei
jmguo@seed.net.tw

2nd Chiao-Chun Yang
Department of Electrical Engineering
National Taiwan University of Science and Technology
Taiwan, Taipei
n8224268@gmail.com

3rd Zong-Hui Wang
Department of Electrical Engineering
National Taiwan University of Science and Technology
Taiwan, Taipei
ol986532@gmail.com

4th Chih-Hsien Hsia
Department of Computer Science and Information Engineering
National Ilan University
Taiwan, Ilan
cchsia625@gmail.com

5th Li-Ying Chang
Department of Electrical Engineering
National Taiwan University of Science and Technology
Taiwan, Taipei
russalchang@gmail.com

***Abstract*—Nowadays, the computerized electrocardiogram (ECG) interpretation is the best available tool for the detection of heart diseases. The symptoms of atrial fibrillation, one of the heart disease, are the most challenging task of heart disease that relies on the cardiologist to read the ECG. However, this task involves a time-consuming process, leading to fatigue-induced medical errors. In this paper, a novel atrial fibrillation diagnostic algorithm based on deep learning architecture is proposed which incorporates the spectrogram to further improve the accuracy performance. As shown in the experimental results, the proposed method on Phy-sioNet /CinC Challenge 2017, which contains 8528 of single-lead ECG recordings, achieves the 10-fold cross-validation set performance of 78%.**

Keywords—deep learning, ResNet, atrial fibrillation, AF Challenge 2017

I. INTRODUCTION

Atrial fibrillation (AF), which is an irregular and often rapid heart rate. According to the study [1], the primary symptom of AF disease is causing the stroke, leading to life-threatening. Nowadays, an electrocardiography (ECG) which records the physiological activity signals of the heart in time units using the magnitude of the voltage as the strength of the heartbeats is the best available diagnostic equipment for detecting the AF disease [2].

The general procedure of cardiac arrhythmia detection and classification methods can roughly divide into four steps including pre-processing, segmentation, feature extraction and classify. In the pre-processing and segmentation, the short-time Fourier transform (STFT) [3] and wavelet transform (WT) [4] are common methods on ECG signal for features extraction. To ensure that ECG features [5] can be more stable and reliable throughout the extraction procedure, the segmentation process plays an important role in the use of classification. Recently, the deep learning architecture can guarantee the selection of the best feature, leading to better classification performance such as convolution neural network (CNN) [6], long-short time memory (LSTM) form recursive neural network (RNN) and ResNet [7]. However, in the ECG classification, how to gather dense information is an important issue which can improve the network performance. ResNet is one of the most famous and novel deep learning architectures to address this problem by proposing the residual block scheme. In the residual block scheme, the ResNet applies the skip-connections which add the identity mapping mechanism into the convolution layer to address this issue. The key contributions of this paper are as follows: (1) To address the lack of global context information, this paper proposed utilizes short-time Fourier transform as the input through the shallow ResNet network. (2) To address the problem of the learning from the unexpected signal, the proposed network applies max-pooling instead of the shortcut directly from ResNet network. (3) To improve the high accuracy rate, the proposed network can achieve better result (F1-score=78%) compare with previous works.

II. METHOD

This paper proposed a high-performance cardiac arrhythmia detection technique based on convolutional neural network. In the contract to the former methods, this algorithm learns spectrogram features which are extracted fixed-length ECG signals to improve the accuracy rate and adopts the max-pooling layers capabilities to enhance training networks. Thus, this section introduces our proposed network architecture and explains the overall details of proposed cardiac arrhythmia detection.

A. Preprocessing

Preprocessing is an essential step in deep learning method as the good quality and useful information of data that can be directly affected the ability of the model to learn. In the public database [8], each of single-lead ECG signal does not contain an equal length of data. Our proposed preprocessing method adopts nine seconds long signal which is divided from each example. To convert the ECG signal into a spectrogram, the short-time Fourier transform is utilized in this study which can extract the features on the frequency domain and obtain more detail information than time domain. Finally, this study sets the window size of short-time Fourier transform to 256 samples with overlap 80%.

B. Network Architecture

Deep neural networks [9] have recently achieved remarkable success in various computer vision research areas, including medical signal computing. With the limited ECG training data, searching optimal hyperparameters for each layer is a daunting task and leads to poor performance. This paper proposes a high-performance automatic classification of cardiac arrhythmia system to overcome overfitting and limited training data situation.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

TABLE I. RECALL, PRECISION AND F1 SCORE RESULTS ON OUR PROPOSED METHOD

	Classes[a]				
	N	***A***	***O***	**~**	***overall***
Recall	0.78	0.83	0.82	0.38	0.81
Precision	0.74	0.94	0.61	0.41	0.76
F1 score	0.76	0.88	0.70	0.39	0.78

[a] The capital "N" indicates the normal signal. The capital "A" indicates the atrial fibrillation signal. The capital "O" indicates the other signal. The mark "~" indicates the noise signal.

TABLE II. F1 SCORES ON DIFFERENT METHODS OF NEURAL NETWORK

	Classes				
	N	***A***	***O***	**~**	***overall***
Fabio [14]	0.84	0.61	0.57	0.54	0.64
Santiago [15]	0.86	0.76	0.64	0.46	0.76
Fernando [7]	0.77	0.72	0.75	0.51	0.72
Shadi [16]	0.88	0.78	0.67	0	0.77
Simon [17]	0.86	0.79	0.64	0	0.77
Our Net	0.78	0.88	0.68	0.42	0.78

The architecture based on ResNet with 18 layers and replaces skip-connections into the max-pooling layers to focus on low-frequency area and implements in the last residual block. Finally, batch normalization [10] and ReLU activation function [11] are followed after each convolution layer in residual block.

C. *Training*

In network architecture, we employ cross-entropy loss as our training objective function, and optimizer is Adam with default parameters recommended in [12]. Also, we reduce learning rate on plateau with factor 0.1.

III. EVALUATION

To compare the quantitative performance of the proposed cardiac arrhythmia detection, this study commonly used confusion matrix [13], precision-recall and F1 score to measure our system. A confusion matrix is constructed to evaluate the classification performance. As recommended by PhysioNet/CinC 2017 the overall evaluation matrix is the average among normal, AF and other signals. As stated above, the classification performance measured is shown in Table. 1. So, our overall F1 score is 78.06%.

In this study, to compare between our method and previous approaches, including 3-layer neural network [14], Resnet-34 with filter added in preprocessing [7], etc. Our proposed method can achieve 78.06% on F1 score. Finally, the details of comparison are shown in Table. 2.

IV. CONCLUSION AND DISCUSSION

The three main contributions of the proposed deep learning architecture in this study are as follow. First, to convert the dimensional signal of ECG into the spectrogram, which can provide more significant information to the network on higher domain. Second, the proposed network architecture utilizes the max-pooling layer to address the overfitting problem. Finally, the single-lead ECG signal of Phy-sioNet /CinC Challenge 2017 database were utilized to evaluate the performance of the proposed algorithm, and compared against the state-of-the-art cardiac arrhythmia detection algorithms. Results show that the proposed method can achieve superior performance (78%) than the part of the deep learning-based algorithms in 2-dimension in terms of the F1.

REFERENCES

[1] Philip A. Wolf, Janet B. Mitchell and Colin S. Baker, "Impact of Atrial Fibrillation on Mortality, Stroke, and Medical Costs," Arch Intern Med, Feb, 1998. 158(3):229-234J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] A.I. Hernandez, J. Dumont, M. Altuve, A. Beuch ´ ee, and G. Carrault, "Evolutionary Optimization of ECG Feature Extraction Methods: Applications to the Monitoring of Adult Myocardial Ischemia and Neonatal Apnea Bradycardia Events," Springer, London, 2012.

[3] Kıymık, M. K., Güler, İ., Dizibüyük, A. and Akın, M, "Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application." Computers in biology and medicine 35.7, 603-616, 2005.

[4] Saritha, C., V. Sukanya, and Y. Narasimha Murthy, "ECG signal analysis using wavelet transforms," Bulg. J. Phys 35.1, 68-77, 2008.

[5] Israel, S. A., Irvine, J. M., Cheng, A., Wiederhold, M. D. and Wiederhold, B. K., "ECG to identify individuals," Pattern recognition 38.1, 133-142, 2005.

[6] Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C. and Ng, A. Y, "Cardiologist-level arrhythmia detection with convolutional neural networks." arXiv preprint arXiv:1707.01836, 2017.

[7] Fernando Andreotti, Oliver Carr, Marco A. F. Pimentel, Adam Mahdi and Maarten De Vos, "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG," 2017 Computing in Cardiology (CinC). IEEE, 2017.

[8] G. D. Clifford, C. Liu, B. Moody, L. w. H. Lehman, I. Silva, Q. Li,A. E. Johnson and R. G. Mark, "AF classification from a short single lead ECG recording: The physionet/computing in cardiology challenge 2017," 2017 Computing in Cardiology (CinC), 2017, pp. 1–4.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, "Deep residual learning for image recognition," Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.

[10] Ioffe, Sergey; Szegedy and Christian, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.

[11] Nair, Vinod and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines," Proceedings of the 27th international conference on machine learning (ICML-10), 2010.

[12] Kingma, Diederik P. and Jimmy Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[13] Ting K.M. "Confusion Matrix," Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning and Data Mining, Springer, Boston, MA, 2017.

[14] Fabio Hernández, Dilio Méndez, Lusvin Amado and Miguel Altuve, "Atrial Fibrillation Detection in Short Single Lead ECG Recordings Using Wavelet Transform and Artificial Neural Networks," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul, 2018.

[15] Santiago Jiménez-Serrano, Jaime Yagüe-Mayans, Elena Simarro-Mondéjar1, Conrado J. Calvo,Francisco Castells1 and José Millet, "Atrial Fibrillation Detection Using Feedforward Neural Networks and Automatically Extracted Signal Features," 2017 Computing in Cardiology (CinC), Sep, 2017.

[16] Shadi Ghiasi, Mostafa Abdollahpur, Nasimalsadat Madani, Kamran Kiani and Ali Ghaffari, "Atrial fibrillation detection using feature based algorithm and deep convolutional neural network," 2017 Computing in Cardiology (CinC). IEEE, 2017.

[17] Simon Geirnaert, Griet Goovaerts, Sibasankar Padhy, Martijn Bousse´, Lieven De Lathauwer, Sabine Van Huffel, "Tensor-based ECG signal processing applied to atrial fibrillation detection," 2018 52nd Asilomar Conference on Signals, Systems, and Computers, IEEE, Oct, 2018

Design and Implementation of Ultra-Low-Latency Video Encoder Using High-Level Synthesis

Kosuke Fukaya
Graduate School of Engineering and Technology, Nihon University
Koriyama, Japan
ceko19018@g.nihon-u.ac.jp

Kaito Mori
Graduate School of Engineering and Technology, Nihon University
Koriyama, Japan
ceka17021@gmail.com

Kousuke Imamura
Institute of Science and Engineering
Kanazawa University
Kanazawa, Japan
imamura@ec.t.kanazawa-u.ac.jp

Yoshio Matsuda
Institute of Science and Engineering
Kanazawa University
Kanazawa, Japan
matsuda@ec.t.kanazawa-u.ac.jp

Tetsuya Matsumura
College of Engineering,
Nihon University
Koriyama, Japan
matsumura@cs.ce.nihon-u.ac.jp

Seiji Mochizuki
Graduate School of Engineering and Technology, Nihon University
Koriyama, Japan
cese16001@gmail.com

***Abstract*— For real-time applications such as autonomous driving and virtual reality (VR), we previously proposed an ultra-low-latency video coding method, which adopts line-based processing for Full-HD video. In this paper, we newly propose a design and implementation of the ultra-low-latency video encoder. In order to reduce the hardware amount, image-prediction specification is optimized for our previous work. Applying a high-level synthesis (HLS) design methodology for Xilinx FPGA, the implementation results of logic count with 10,677 LUTs, 3,714FFs and 66 DSPs is obtained. The implemented video encoder achieves less than 1.0 µs low-latency and compression to 39.4% without significant visual degradation. As a result, cost effective ultra-low-latency video encoder is implemented for low cost FPGA.**

Keywords—video coding; low latency; autonomous driving; virtual reality (VR), FPGA

I. Introduction

Recently, autonomous driving for vehicles and virtual reality/mixed reality (VR/MR) for Internet of Things (IoT) devices are studied and developed actively. According to this trend, we previously proposed an ultra-low-latency video coding method [1]. This method is line-based coding to reduce the latency on the order of microseconds, while the current standard video codec such as MPEG-2 [2] and H.264 [3] have the latency on the order of milliseconds. On the other hand, the line-based compression system which performs low-latency transmission was reported [4], but it is not sufficient from the viewpoint of compression ratio and latency.

In this paper, we propose design and implementation of ultra-low-latency video coding on the basis of our previous work [1]. Furthermore, in order to reduce the hardware amount, optimized line-based prediction method is verified and implemented. As a result, cost effective low-latency video encoder is realized with small hardware amount compared to conventional line-based video transmission system [4]. This result is expected to the image recognition system with ultra-low-latency as future autonomous image sensing devices and IoT devices with low cost FPGA.

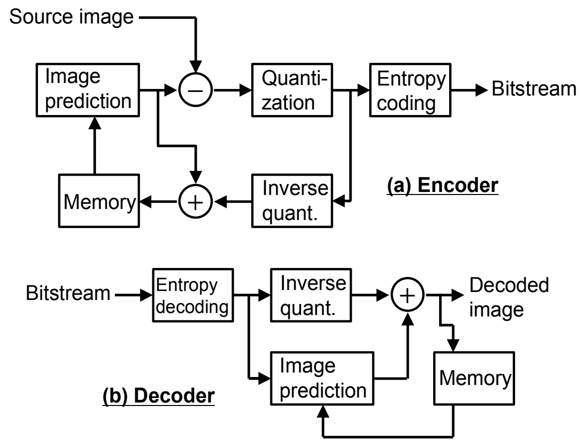


Fig. 1. Block diagrams of encoder and decoder.

TABLE I. Evaluation Results for Basic and Optimized Method

Evaluation item	Average of basic method [1]	Average of optimized method
RPD (Reference prediction) Spec.	-16 pixel ~ +15 pixel *(Half-pel precision)*	-4 pixel ~ +3 pixel *(Integer-pel precision)*
CR (%)	39.3	39.4
PSNR Y (dB)	45.4	45.1
PSNR U (dB)	42.1	41.6
PSNR V (dB)	39.7	39.2

II. Algorithm & Performance of Proposed Method

Figure 1(a) and (b) show a block diagram of line-based Full-HD ultra-low-latency video encoder and decoder that supports the proposed method previously [1]. This algorithm basically utilizes a processing unit of 16 pixels by 1 line, called a compression block (CB). The proposed method is designed for a pixel depth of 8bits/pixel and the YUV420 color format. The coding method is composed of the following tools: Line-based prediction such as prediction from the neighboring pixels (NPD) and reference prediction (RPD), image-adaptive quantization, optimized entropy coding, and rate control. In the previous report [1], the proposed method achieves compression to 39.1% data size and image quality of

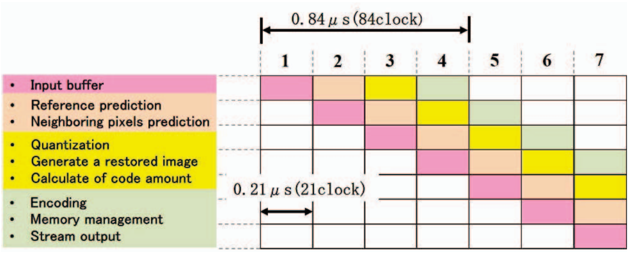


Fig. 2. Pipeline structure

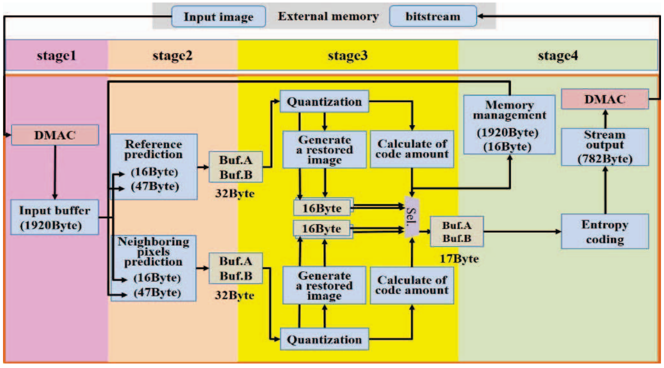


Fig. 3. Block diagram of video encoder

45.4dB. The rate distortion curves show the proposed basic algorithm achieve compression to 33% without significant visual degradation. With the proposed video coding, the autonomous vehicles and VR devices implementing can transfer HD video using 20% of the bandwidth of the source video without significant latency or visual degradation.

Table I shows the evaluation results for the previously proposed basic method and optimized method without rate control. The compression ratio (CR) and image quality were evaluated using four test video sequences of Full-HD resolution (1920 pixels * 1080 lines) at 60fps. In the optimized method, in order to reduce the hardware amount, the search range of the RPD is improved from original -16/+15 for half–pel precision to -4/+3 for integer-pel precision with maintaining image quality. The average CR is 39.4%, and the average PSNR is 45.1dB. Thus, the proposed method achieves 40% data reduction without visual degradation.

III. Design and Implementatoin

Figure 2 shows a pipeline structure realizing less than 1µs delay. The pipeline is designed to four stage structure. Since input rate of the pixels for Full-HD resolution at 30fps, whose dotclock is 74.25MHz, is 0.22 µs per 16 pixels for a CB, all stages are designed to perform each processing stage less than 0.21 µs. The first stage executes image input control using DMA controller and Input buffer. Next, the second stage performs line-based prediction such as prediction from the neighboring pixels and reference prediction. The third stage executes quantization and restore image generation. Finally, the fourth stage carries out the entropy coding, and stream output is managed by DMA output controller. Figure 3 shows the detailed block diagram of ultra-low-latency video encoder. According to the pipeline stage, each processing unit is arranged as shown in Fig. 2.

TABLE II. Implementation results

Item	Previous method[1]				Optimized method			
	Cycles (clock)	Logic amount			Cycles (clock)	Logic amount		
		FF	LUT	DSP		FF	LUT	DSP
RPD	20	752	3948	0	19	635	2126	0
NPD	19	700	1022	2	19	700	1022	2
Quantization	7	780	1670	32	7	780	1670	32
Restored image	7	860	2050	32	7	860	2050	32
Code amount cal.	2	92	1047	0	2	92	1047	0
Entropy Coding	17	616	2119	0	17	616	2119	0
Mem. management	1	31	643	0	1	31	643	0
Total	-	3831	12499	66	-	3714	10677	66

The ultra-low-latency video encoder is implemented using Xilinx Vivado HLS environment. In order to reduce synthesized logic amount and perform the low-latency, original C source code is optimized for high-level synthesis. Table □ shows an implementation result that lists the clock cycles for operation and generated logic amount for each module at the clock frequency of 100 MHz. The clock cycle of each stage is less than 20 clocks, which results in keeping design constraints of low-latency. The video encoder is implemented with 10,677 LUTs, 3,714FFs and 66 DSPs, which is very competitive result regardless of the advantage of compression ratio and latency compared with conventional line-based compression system [4]. This means a cost effective implementation has been accomplished as a sensing device solution for autonomous application.

IV. Conclution

The ultra-low-latency video encoder is designed with four stage pipeline structure and implemented using Vivado HLS environment for Xilinx FPGA. The encoder achieves less than 1.0 µs latency for Full-HD video by adopting line-based processing without significant visual degradation. Furthermore, adopting optimized reference prediction method, the video encoder is implemented with 10,677 LUTs, 3,714FFs and 66 DSPs, which result in competitive hardware amount regardless of the advantage of compression ratio and latency compared to conventional line-based compression system [4]. This result is expected to realize an image recognition system with ultra-low-latency as future autonomous image sensing devices and IoT devices with low cost FPGA.

Acknowledgment

This work is supported by JSPS KAKENHI Grant Number JP17K06444.

References

[1] S. Mochizuki, et al, “Ultra-low-latency video coding method for autonomous vehicles and virtual reality devices,” Proceedings of IEEE IOTAIS 2018, pp.155-161, Nov. 2018.

[2] ISO/IEC 13818-2:2013 Information technology -- Generic coding of moving pictures and associated audio information -- Part 2: Video, October 2013, https://www.iso.org/standard/61152.html.

[3] ITU-T Rec. H.264 (02/2016) Advanced video coding for generic audiovisual services, Feb. 2016, https://www.itu.int/rec/T-REC-H.264.

[4] T. Inatsuki et al., "An FPGA implementation of low-latency video transmission system using lossless and near-lossless line-based compression," Proceedings of IEEE Digital Signal Processing, pp. 10621066, July 2015

A Convolutional Dictionary Learning based l_1 Norm Error with Smoothed l_0 Norm Regression

Kaede Kumamoto(*) Shinnosuke Matsuo Yoshimitsu Kuroki(**)
National Institute of Technology (KOSEN), Kurume College, Fukuoka, Japan
(*)a3611kk@kurume.kosen-ac.jp (**)kuroki@kurume-nct.ac.jp

***Abstract*—General sparse representation schemes for image signals divide given images into non-overlapped patches, and discuss dictionaries and/or sparse coefficients for the patches. A convolutional sparse representation provides an alternative approach, in which an entire image is represented by the sum of convolutions with dictionary filters. This work tackles designing of the dictionary filters with the l_1 norm error criterion in the data fidelity term to improve robustness against outliers. This paper also applies smoothed l_0 norm to estimate sparseness of coefficient vectors.**

***Index Terms*—Convolutional sparse representation, l_1 error fidelity, smoothed l_0 norm regression**

I. Introduction

Sparse representations [1] [2] are widely used for very broad fields of signal and image processing applications, in which an entire image is generally divided into non-overlapped patches, and the patches are represented with linear combination of dictionary vectors. A convolutional sparse representation provides an alternative approach, which models an entire signal or image as the sum of convolutions with dictionary filters. In both patch-based and convolutional sparse representation, the cost function to design dictionaries and to obtain coefficient vectors consists of the weighted sum of two terms: data fidelity and regularization terms. The former and the latter are generally estimated with l_2 norm error and the l_1 norm of coefficient vectors respectively. To improve the robustness of the dictionaries against outliers, this study employs l_1 norm error for the fidelity term and smoothed l_0 norm regression for the sparseness to fast implementation.

II. Preliminaries

For a given signal $\boldsymbol{s}_k \in \mathbb{R}^N$, where N is the number of pixels, a convolutional sparse representation expresses $\boldsymbol{s}_k$ with the sum of the set of M dictionary filters $\boldsymbol{d}_m \in \mathbb{R}^N (m = 1, 2, \ldots, M)$ and coefficients $\boldsymbol{x}_{k,m} \in \mathbb{R}^N$ as

$$\boldsymbol{s}_k = \sum_{m=1}^{M} \boldsymbol{d}_m * \boldsymbol{x}_{k,m}. \tag{1}$$

We update $\boldsymbol{d}_m$ and $\boldsymbol{x}_{k,m}$ alternatively as the following procedures [3]. For a given set of K images for dictionary learning, the convolutional dictionary learning is defined as

$$\underset{\{\boldsymbol{d}_m\}}{\arg\min} \frac{1}{2} \sum_{k=1}^{K} \left\| \sum_{m=1}^{M} \boldsymbol{x}_{k,m} * \boldsymbol{d}_m - \boldsymbol{s}_k \right\|_2^2 \ s.t.\ \boldsymbol{d}_m \in C_{PN}, \tag{2}$$

$$C_{PN} = \{\boldsymbol{y} \in \mathbb{R}^N : (I - PP^T)\boldsymbol{y}, ||\boldsymbol{y}||^2 \leq 1\}. \tag{3}$$

Equation (3) means that the zero-padded dictionary $\boldsymbol{d}_m$ is same as $PP^T\boldsymbol{d}_m$ and is projected onto the unit sphere. Here, $P \in \mathbb{R}^{N \times B}$ is the zero-padding matrix for a original filter $P^T\boldsymbol{d}_m \in \mathbb{R}^B$. Rewriting Eq.(2) and (3), we obtain

$$\underset{\{\boldsymbol{d}_m\}}{\arg\min} \frac{1}{2} \sum_{k=1}^{K} \left\| \sum_{m=1}^{M} \boldsymbol{x}_{k,m} * \boldsymbol{d}_m - \boldsymbol{s}_k \right\|_2^2 + \sum_{m=1}^{M} \imath_{C_{PN}}(\boldsymbol{g}_m) \quad s.t.\ \boldsymbol{d}_m - \boldsymbol{g}_m = 0, \tag{4}$$

where $\imath_{C_{PN}}(\boldsymbol{g}_m)$ is 0 when $\boldsymbol{g}_m \in C_{PN}$, ∞ otherwise. We solve this problem with respect to $\boldsymbol{d}_m$ and $\boldsymbol{g}_m$ alternatively with Alternating Direction Method of Multipliers (ADMM), one of convex solvers.

The l_0 norm is discontinuous and indifferentiable; then, [4] approximates the l_0 norm to a smoothed l_0 norm with a Gaussian function as

$$||\boldsymbol{x}||_0 \approx N - \sum_{n=1}^{N} e^{-\frac{x_n{}^2}{2\sigma^2}}. \tag{5}$$

Since the approximation goes to equal when $\sigma \to 0$, σ decreases as the update to $\boldsymbol{d}_m$ and $\boldsymbol{x}_{k,m}$ update proceeds. The update of $\boldsymbol{x}_{k,m}$ is shown as

$$\underset{\{x_{k,m,n}\}}{\arg\min} KMN - \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{N} e^{-\frac{x_{k,m,n}{}^2}{2\sigma^2}} \quad s.t.\ \sum_{m=1}^{M} \boldsymbol{d}_m * \boldsymbol{x}_{k,m} = \boldsymbol{s}_k. \tag{6}$$

We solve Eq.(6) with Stochastic Gradient Descent (SGD), and the approximation error is evaluated by

$$\underset{\{\boldsymbol{x}'_{k,m}\}}{\arg\min} \frac{1}{2} \sum_{m=1}^{M} ||\boldsymbol{x}'_{k,m} - \boldsymbol{x}_{k,m}||_2^2 + \frac{\lambda}{2} \left\| \sum_{m=1}^{M} \boldsymbol{d}_m * \boldsymbol{x}'_{k,m} - \boldsymbol{s}_k \right\|_2^2. \tag{7}$$

The first term constrained $\boldsymbol{x}'_{k,m}$ to the neighbor of $\boldsymbol{x}_{k,m}$ because without this term Eq. (4) does not change $\boldsymbol{d}_m$.

III. Proposed Method

In the conventional method, the influence of outliers appears in the dictionary filters. To solve this problem, our method evaluates the data fidelity term with the l_1 norm.

By applying l_1 norm error, Eq. (4) is changed:

$$\begin{aligned}&\arg\min_{\boldsymbol{d}} ||\boldsymbol{X}\boldsymbol{d}-\boldsymbol{s}||_1 + \imath_{C_{PN}}(\boldsymbol{d}), \qquad (8)\\ &\boldsymbol{X}\boldsymbol{d} = IFFT(\hat{\boldsymbol{X}}\hat{\boldsymbol{d}}),\\ &\boldsymbol{d} = \left(\boldsymbol{d}_1\ \boldsymbol{d}_2 \cdots \boldsymbol{d}_M\right)^T,\ \boldsymbol{s} = \left(\boldsymbol{s}_1\ \boldsymbol{s}_2 \cdots \boldsymbol{s}_M\right)^T,\end{aligned}$$

where $\hat{\boldsymbol{X}}$ means $\boldsymbol{X}$ in the Fourier domain. We again formulate this to solve by ADMM as

$$\arg\min_{\boldsymbol{g}_0,\boldsymbol{g}_1} ||\boldsymbol{g}_0||_1 + \imath_{C_{PN}}(\boldsymbol{g}_1)\ s.t.\ \begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{I}\end{pmatrix}\boldsymbol{d} - \begin{pmatrix}\boldsymbol{g}_0\\ \boldsymbol{g}_1\end{pmatrix} = \begin{pmatrix}\boldsymbol{s}\\ \boldsymbol{0}\end{pmatrix}. \qquad (9)$$

ADMM obtains the dictionary for fixed coefficients as the following iterative manner

$$\boldsymbol{d}^{(j+1)} = \arg\min_{\boldsymbol{d}} \left|\left|\boldsymbol{X}\boldsymbol{d} - \boldsymbol{g}_0^{(j)} - \boldsymbol{s} + \boldsymbol{h}_0^{(j)}\right|\right|_2^2 + \left|\left|\boldsymbol{d} - \boldsymbol{g}_1^{(j)} + \boldsymbol{h}_1^{(j)}\right|\right|_2^2 \qquad (10)$$

$$\boldsymbol{g}_0^{(j+1)} = \arg\min_{\boldsymbol{g}_0} ||\boldsymbol{g}_0||_1 + \frac{\rho}{2}\left|\left|\boldsymbol{X}\boldsymbol{d}^{(j+1)} - \boldsymbol{g}_0 - \boldsymbol{s} + \boldsymbol{h}_0^{(j)}\right|\right|_2^2 \qquad (11)$$

$$\boldsymbol{g}_1^{(j+1)} = \arg\min_{\boldsymbol{g}_1} \imath_{C_{PN}}(\boldsymbol{g}_1) + \frac{\rho}{2}\left|\left|\boldsymbol{d}^{(j+1)} - \boldsymbol{g}_1 + \boldsymbol{h}_1^{(j)}\right|\right|_2^2 \qquad (12)$$

$$\boldsymbol{h}_0^{(j+1)} = \boldsymbol{h}_0^{(j)} + \boldsymbol{X}\boldsymbol{d}^{(j+1)} - \boldsymbol{g}_0^{(j+1)} - \boldsymbol{s} \qquad (13)$$

$$\boldsymbol{h}_1^{(j+1)} = \boldsymbol{h}_1^{(j)} + \boldsymbol{d}^{(j+1)} - \boldsymbol{g}_1^{(j+1)}. \qquad (14)$$

The optimization problem to obtain the coefficients $\boldsymbol{x}_k$ for the dictionary is shown as

$$\begin{aligned}&\arg\min_{\boldsymbol{x}'_k} \frac{1}{2}||\boldsymbol{x}'_k - \boldsymbol{x}_k||_2^2 + \lambda||\boldsymbol{D}\boldsymbol{x}'_k - \boldsymbol{s}_k||_1, \qquad (15)\\ &\boldsymbol{D}\boldsymbol{x} = IFFT(\hat{\boldsymbol{D}}\hat{\boldsymbol{x}}),\\ &\boldsymbol{x}_k = \left(\boldsymbol{x}_{k,1}\ \boldsymbol{x}_{k,2}\cdots\boldsymbol{x}_{k,M}\right)^T,\ \boldsymbol{x}'_k = \left(\boldsymbol{x}'_{k,1}\ \boldsymbol{x}'_{k,2}\cdots\boldsymbol{x}'_{k,M}\right)^T,\end{aligned}$$

and is also rewritten to solve with ADMM as

$$\arg\min_{\boldsymbol{x}'_k} \frac{1}{2}||\boldsymbol{x}'_k - \boldsymbol{x}_k||_2^2 + \lambda||\boldsymbol{y}_k||_1\ \ s.t.\ \boldsymbol{D}\boldsymbol{x}'_k - \boldsymbol{y}_k = \boldsymbol{s}_k. \qquad (16)$$

The solution is given by the following iteration

$$\boldsymbol{x}'^{(j+1)}_k = \arg\min_{\boldsymbol{x}'_k} \frac{1}{2}||\boldsymbol{x}'_k - \boldsymbol{x}_k||_2^2 + \frac{\gamma}{2}\left|\left|\boldsymbol{D}\boldsymbol{x}'_k - \boldsymbol{y}_k^{(j)} - \boldsymbol{s}_k + \boldsymbol{u}_k^{(j)}\right|\right|_2^2 \qquad (17)$$

$$\boldsymbol{y}_k^{(j+1)} = \arg\min_{\boldsymbol{y}_k} \lambda||\boldsymbol{y}_k||_1 + \frac{\gamma}{2}\left|\left|\boldsymbol{D}\boldsymbol{x}'^{(j+1)}_k - \boldsymbol{y}_k - \boldsymbol{s}_k + \boldsymbol{u}_k^{(j)}\right|\right|_2^2 \qquad (18)$$

$$\boldsymbol{u}_k^{(j+1)} = \boldsymbol{u}_k^{(j)} + \boldsymbol{D}\boldsymbol{x}'^{(j+1)}_k - \boldsymbol{y}_k^{(j+1)} - \boldsymbol{s}_k. \qquad (19)$$

IV. Experiments

To evaluate the proposed method, we design the dictionary filters for five images. The size of every image is 512×512 pixels in gray scale quantized to 8 bits. The images are preprocessed the average of pixel values to be zero. The filter size is 8×8, and its number is 64. The initial dictionary is generated by the standard normal Gaussian. The parameter σ is changed 15 times by multiplying by 0.9. The designed dictionary is applied for Lenna, which is not used for dictionary design; then, PSNR versus l_0 norm performance is compared in Fig. 1 and 2. The parameters of the conventional method are $\mu = 1$, $\lambda = 0.1$, and $\rho = 0.1$; and those of proposed method are $\mu = 1$, $\gamma = 10$, $\lambda = 0.1$, and $\rho = 1$.

(a) Original image (b) Restored by conventional method (c) Restored by proposed method

Fig. 1: Original image and restored images

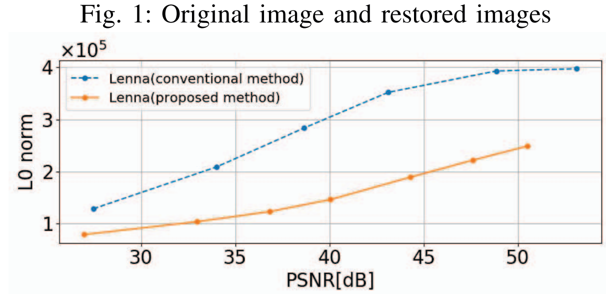


Fig. 2: The relationship between PSNR and the l_0 norm

The experiment results show that our method restore high PSNR images with less l_0 norm.

V. Conclusion

We improve the robustness of dictionaries against outliers by employing l_1 norm for the fidelity term, and this method increases sparseness in the same PSNR. Future works include development of variable length filters.

References

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," SIAM Review, vol. 51, no. 1, pp. 34-81, 2009. DOI:10.1137/060657

[2] J. Mairal, F. Bach, and J. Ponce, "Sparse modeling for image and vision processing," Foundations and Trends in Computer Graphics and Vision, vol. 8, no. 2-3, pp. 85-283, 2014. doi:10.1561/0600000

[3] B. Wohlberg, "Efficient Algorithms for Convolutional Sparse Representations," IEEE Trans. on Image Processing, vol.25, no.1, pp.301-315, Jan 2016. DOI:10.1109/TIP.2015.2495260.

[4] H. Mohimani, M. Babie-Zadeh, and C. Jutten, "A fast approach for over complete sparse decomposition based on smoothed l_0 norm," IEEE Trans. on Siganl Processing, vol.57, no.1, pp.289-301, Jan 2009. DOI: 10.1109/TSP.2008.2007606.

Convolutional sparse dictionary learning with smoothed l_0 norm and projected gradient descent

Kazuki Kitajima, Akira Sugano, and Yoshimitsu Kuroki
Natinal Institute of Technology (KOSEN), Kurume College, Fukuoka, Japan
a3609kk@kurume.kosen-ac.jp, s54215sa@kurume.kosen-ac.jp, kuroki@kurume-nct.ac.jp

***Abstract*—Convolutional sparse representation extracts features of given signals. Conventional methods employ Alternating Direction Method of Multipliers (ADMM) to design dictionaries. However, their convergence is slow and computational cost is burden. This paper proposes a faster design method of the dictionaries by using Projected Gradient Descent (PGD) instead of ADMM, where the l_0 norm of coefficient vectors is approximated by a smoothed l_0 norm function similar to conventional methods. Experimental results demonstrate that the proposed method is faster and less iterations than a conventional method.**

***Index Terms*—convolutional dictionary learning, smoothed l_0 norm, projected gradient decent**

I. Introduction

Sparse representations [1] are widely used for very broad fields of signal and image processing, in which non-overlapped blocks of a given image are generally approximated by linear combination of dictionary vectors. A convolutional sparse representation provides an alternative approach, which models an entire signal as the sum of convolutions with dictionary filters.

II. Preliminaries

For a given image $\boldsymbol{s}_k \in \mathbb{R}^N$, a convolutional sparse representation expresses $\boldsymbol{s}_k$ by convolutions of a set of M dictionary filters $\boldsymbol{d}_m \in \mathbb{R}^N (m = 1, 2, \cdots, M)$ and coefficients $\boldsymbol{x}_{k,m}$ as

$$\underset{\{\boldsymbol{x}_{k,m}\}}{\arg\min} \sum_{m=1}^{M} L\left(\boldsymbol{x}_{k,m}\right) \quad \text{s.t.} \quad \sum_{m=1}^{M} \boldsymbol{d}_m * \boldsymbol{x}_{k,m} = \boldsymbol{s}_k, \tag{1}$$

where $L(\boldsymbol{x}_{k,m})$ is a function to estimate the sparseness of $\boldsymbol{x}_{k,m}$, and the following smoothed l_0 norm is used:

$$\|\boldsymbol{x}\|_0 \approx L(\boldsymbol{x}) = N - \sum_{i=1}^{N} e^{-x_i^2/2\sigma^2}. \tag{2}$$

The dictionary $\boldsymbol{d}_m$ and its coefficients $\boldsymbol{x}_{k,m}$ are updated alternately so that $\boldsymbol{x}_{k,m}$ are sparse. Please note that the original dimension of $\boldsymbol{d}_m$ is less than N but is padded with zeros to be $\boldsymbol{d}_m \in \mathbb{R}^N$.

To optimize $\boldsymbol{x}_{k,m}$ for a fixed $\boldsymbol{d}_m$ in Eq.(1), it is rewritten in Fourier domain as

$$\underset{\{\boldsymbol{x}_{k,m}\}}{\arg\min} \sum_{m=1}^{M} L\left(\boldsymbol{x}_{k,m}\right) \quad \text{s.t.} \quad \hat{D}\hat{\boldsymbol{x}}_{k,m} = \hat{\boldsymbol{s}}_k, \tag{3}$$

where

$$\hat{D} = \left(\operatorname{diag}\left(\hat{\boldsymbol{d}}_1\right) \cdots \operatorname{diag}\left(\hat{\boldsymbol{d}}_M\right)\right), \hat{\boldsymbol{x}}_{k,m} = \begin{pmatrix} \hat{\boldsymbol{x}}_1 \\ \vdots \\ \hat{\boldsymbol{x}}_M \end{pmatrix}. \tag{4}$$

The vectors $\hat{\boldsymbol{x}}_{k,m}, \hat{\boldsymbol{d}}_m$, and $\hat{\boldsymbol{s}}_k \in \mathbb{C}^N$ are Fourier domain vectors of $\boldsymbol{x}_{k,m}, \boldsymbol{d}_m$, and $\boldsymbol{s}_k \in \mathbb{R}^N$ respectively. The convolution $\boldsymbol{x}_{k,m} * \boldsymbol{d}_m$ is expressed as $\operatorname{diag}\left(\hat{\boldsymbol{x}}_{k,m}\right)\hat{\boldsymbol{d}}_m$. The sparseness term in Eq.(3), namely $L(\boldsymbol{x}_{k,m})$, is updated with PGD[3], and the gradient descent is

$$x_i \leftarrow x_i - \mu x_i e^{\frac{x_i^2}{2\sigma^2}}, \tag{5}$$

where x_i is each component of $\boldsymbol{x}_{k,m}$, and μ is a learning rate. If $\boldsymbol{x}_{k,m}$ satisfies $\hat{D}\hat{\boldsymbol{x}}_{k,m} = \hat{\boldsymbol{s}}_k$, the dictionary can not be changed at the dictionary update procedure. Then, [2] does not obtain the solution of $\hat{D}\hat{\boldsymbol{x}}_{k,m} = \hat{\boldsymbol{s}}_k$, and the convolutional approximation is replaced with the following formula using a weight parameter λ as

$$\underset{\hat{\boldsymbol{x}}'_{k,m}}{\arg\min} \left\|\hat{\boldsymbol{x}}'_{k,m} - \hat{\boldsymbol{x}}_{k,m}\right\|_2^2 + \lambda \left\|\hat{D}\hat{\boldsymbol{x}}'_{k,m} - \hat{\boldsymbol{s}}_k\right\|_2^2. \tag{6}$$

The dictionary optimization for K training images is shown

$$\underset{\{\boldsymbol{d}_m\}}{\arg\min} \frac{1}{2} \sum_{k=1}^{K} \left\| \sum_{m=1}^{M} \boldsymbol{x}_{k,m} * \boldsymbol{d}_m - \boldsymbol{s}_k \right\|_2^2 \quad \text{s.t. } \boldsymbol{d}_m \in C_{PN}, \tag{7}$$

$$C_{PN} = \left\{\boldsymbol{x} \in \mathbb{R}^N : (I - PP^T)\boldsymbol{x} = \boldsymbol{0}, \|\boldsymbol{x}\|_2 \leq 1\right\}, \tag{8}$$

where C_{PN} means that a zero-padded dictionary $\boldsymbol{d_m} \in \mathbb{R}^N$ is same as $PP^T\boldsymbol{d_m}$ and is projected onto the unit sphere. Here, $P \in \mathbb{R}^{N \times B}$ is the zero-padding matrix for a original filter $P^T\boldsymbol{d_m} \in \mathbb{R}^B$. Equation (7) is solved with ADMM in the Fourier domain. This computational cost is the sum of two terms: FFT and the solution of Eq.(7) using ADMM. The former is $\mathcal{O}(KMN \log N)$ and the latter depends on the update procedure of $\boldsymbol{d}_m$: $\mathcal{O}(KM^3N)$ in Gaussian Elimination or $\mathcal{O}(K^2MN)$ in Sherman-Morrison Algorithms. Please see [1] for details. The conventional method iterates Eq.(5), the solution of (6), and that of (7) with ADMM.

III. Proposed Method

This paper replaces ADMM for the dictionary update with PGD, and reduces the computational burden. Firstly, the proposed method replaces Eq.(6) with

$$\underset{\hat{\boldsymbol{x}}'_{k,m}}{\arg\min} \left\|\hat{\boldsymbol{x}}'_{k,m} - \hat{\boldsymbol{x}}_{k,m}\right\|_2^2 \text{ s.t. } \left\|\hat{D}\hat{\boldsymbol{x}}'_{k,m} - \hat{\boldsymbol{s}}_k\right\|_2^2 = 0, \tag{9}$$

978-1-7281-3038-5/19/$31.00

to obtain the point projected from the previous solution $\hat{\boldsymbol{x}}_{k,m}$ onto the space satisfying $\hat{D}\hat{\boldsymbol{x}}_{k,m} = \hat{\boldsymbol{s}}_k$ in the Fourier domain. The solution is given by the Lagrange multiplier as

$$\hat{\boldsymbol{x}}_{k,m} \leftarrow \hat{\boldsymbol{x}}_{k,m} - \hat{D}^H \left(\hat{D}\hat{D}^H\right)^{-1} (\hat{D}\hat{\boldsymbol{x}}_{k,m} - \hat{\boldsymbol{s}}_k). \tag{10}$$

Next, we apply PGD to the approximation error shown in Eq.(7) in the Fourier domain, which is expressed as

$$\hat{f}(\hat{\boldsymbol{d}}) = \frac{1}{2} \sum_{k=1}^{K} \left\| \hat{X}_k \hat{\boldsymbol{d}} - \hat{\boldsymbol{s}}_k \right\|_2^2 \tag{11}$$

where

$$\hat{X}_k = \left(\mathrm{diag}\left(\hat{\boldsymbol{x}}_{k,1}\right) \cdots \mathrm{diag}\left(\hat{\boldsymbol{x}}_{k,M}\right)\right), \hat{\boldsymbol{d}} = \begin{pmatrix} \hat{\boldsymbol{d}}_1 \\ \vdots \\ \hat{\boldsymbol{d}}_M \end{pmatrix}, \tag{12}$$

and this gradient is

$$\nabla_{\hat{\boldsymbol{d}}} \hat{f} = \sum_{k=1}^{k} \hat{X}_k^H \left(\hat{X}_k \hat{\boldsymbol{d}} - \hat{\boldsymbol{s}}_k\right). \tag{13}$$

We obtain the gradient $\nabla_{\boldsymbol{d}_m} f$ by Inverse-FFT of Eq.(13), and the gradient descent of dictionary $\boldsymbol{d}_m$ is given by

$$\boldsymbol{d}_m \leftarrow \boldsymbol{d}_m - \alpha \nabla_{\boldsymbol{d}_m} f, \tag{14}$$

where α is the learning rate. The projection of $\boldsymbol{d}_m$ onto the set of C_{PN} in Eq.(8) is shown as

$$\boldsymbol{d}_m \leftarrow \begin{cases} \frac{PP^T \boldsymbol{d}_m}{\|PP^T \boldsymbol{d}_m\|_2} & (\|PP^T \boldsymbol{d}_m\|_2 > 1) \\ PP^T \boldsymbol{d}_m & (\text{ otherwise }) \end{cases}. \tag{15}$$

The whole update process is the iteration of the update procedures of Eq.(5), (14), (15), and (9). Please note that the error $\| \sum_{m=1}^{M} \boldsymbol{d}_m * \boldsymbol{x}_{k,m} - \boldsymbol{s}_k \|_2^2$ is not zero at the first three steps, namely from Eq.(5) to (15); then, $\boldsymbol{d}_m$ can be changed in this process. Also, the computational cost at each iteration is given by $O(KMN \log N)$ of FFT.

Our method varies the parameter σ at every iteration as $\sigma \leftarrow c\sigma$ with a constant $c < 1$. We give σ a maximum of initial coefficients as initial value like [3].

IV. Experiments

We designed the $M = 256$ dictionary filters for five images "Barbara", "kodim", "monarch", "sail", and "tulips" to evaluate the validity of our method by using [4]. The five images are 256×256 arrays quantized to 8-bit gray scale. The number of dictionary filters is $B = 12 \times 12 = 144$, and the filters are quantized to signed 8 bit gray scale. We optimized the sparse coefficients for "Lenna" and "couple" using the dictionary designed for the above five images. Figure 1 compares the iterations of the proposed and the conventional methods showing the similar l_0 norm and SSIM, where inverse SSIM of its vertical axis expresses $1 -$ SSIM. Table 1 shows the execution times of both methods. The parameter of the conventional method is set $\lambda = 0.01$, and those of proposed method are set $\mu = 1$ and $\alpha = 1$. Figure 3 depicts that the dictionary filters of both methods, and they are different. However, we have not estimate the effect of the dictionaries yet.

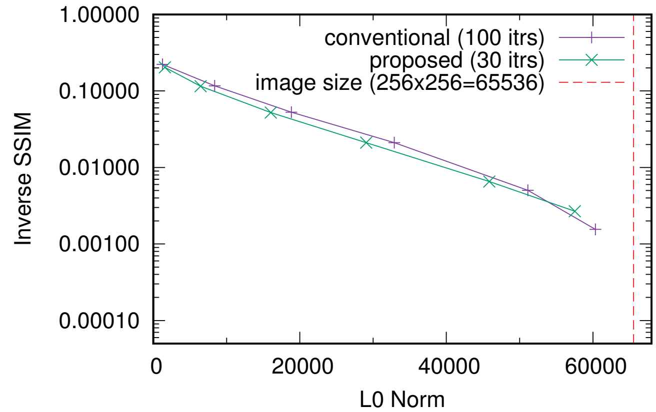
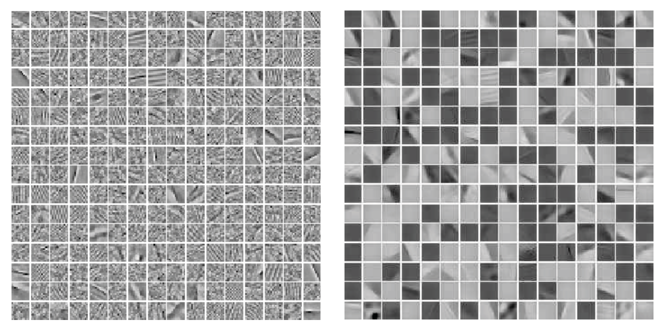


Fig. 1: Comparison of relation of l_0 norm and SSIM

Fig. 2: Comparison of execution time

	proposed method (30 itrs)	conventional method (100 itrs)
execution time	$177.3s$	$1000.1s$



(a) conventional method (b) proposed method

Fig. 3: Comparison of dictionary filters

V. Conclusion

This paper proposes a first convolutional sparse dictionary learning method. Experimental results demonstrate that the proposed method is faster and less iterations than a conventional method. The authors would like to validate the proposed method for a large scale data set including high definition images in the future.

References

[1] B.Wohlberg, "Efficient algorithms for convolutional sparse representation," IEEE Trans. on Image Processing, vol.25, no.1, pp.301-315, Jan. 2016.

[2] Shogo IIZUKA and Makoto NKASHIZUKA, "Convolutional Dictionary Learning based on Smoothed L0 Regularization," 33rd SIP SYMPOSIUM P-14(in Japanese), Nov.2018.

[3] H.Mohimani, M.Babaie-Zadeh,and C.Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed l0 norm.", IEEE Trans. on Signal Processing, vol.57, no.1, pp.289-301, Jan. 2009.

[4] B.Wohlberg, "SPORCO", http://brendt.wohlberg.net/software/SPORCO/, Feb.2019.

Improvements on Data Insertion Technique in Encrypted Image against Lossy Compression

Ryoma Ito, Kiyoshi Tanaka
Shinshu University
Nagano, Japan
{17w2801d, ktanaka}@shinshu-u.ac.jp

KokSheik Wong
Monash University Malaysia
Selangor, Malaysia
wong.koksheik@monash.edu

Simying Ong
University of Malaya
Kuala Lumpur, Malaysia
simying.ong@um.edu.my

***Abstract*—This paper proposes two techniques in enhancing the robustness of a joint image encryption and data insertion method against lossy compression [1]. Safety zone is first introduced to ensure that pixel values representing '0' are separated from those representing '1' with a gap in between. This empty gap serves as a buffer so that pixels from both groups do not cross and eventually affect the extraction process. Next, majority vote is also adopted so that error can be further suppressed, although reducing number of bits that can be inserted. Results suggest that both strategies can suppress error caused by lossy compression.**

***Index Terms*—JEDI, safety zone, majority vote, bit error rate**

I. Introduction

Recently researchers consider various means to achieve joint encryption and data insertion (JEDI) method [1] for image. JEDI method allows content administration without knowing its actual semantic. For example, in cloud storage, image is encrypted to protect owner's privacy, while the inserted data can be extract from the encrypted image for indexing or hyper-linking purpose. Similar, in video surveillance, video is encrypted to protect unauthorized viewing, while the inserted authentication data can be extracted for verification purpose [2].

For such purposes, instead of maintaining quality, Ong et al. [3] treat data embedding with the aim to distort image. Associated histogram bins are first defined, then each image pixel value is modified to the corresponding pixel value based on the data to be inserted. Liu et al. [4] encrypt an input image by transferring the MSB planes to LSB planes through complex swapping, followed by the subblocks permutation. The vacated space is then utilized for data embedding. Recently, we proposed a separable encryption and data insertion method [1], where an image is divided into 2 non-overlapping parts for manipulation purpose. Specifically, the first part (i.e., bit planes of high significance value) is manipulated to mask the perceptual semantic of the image, while the other part (i.e., the remaining bit planes) is manipulated to hide data using histogram matching technique [5]. While the aforementioned methods are able to hide data into encrypted image, robustness of the inserted data against compression is not discussed, i.e., the encrypted image may undergo lossy compression. Therefore, in this work, 2 techniques are put forward to improve our previous JEDI work [1], with the aim to make the inserted data robust against lossy compression.

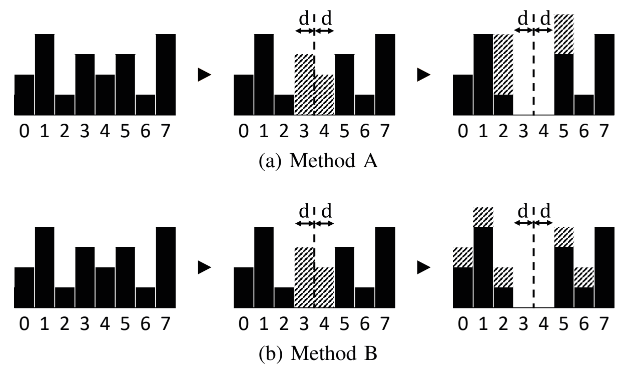


Fig. 1: Illustration of safety zone for $\tau = 3$.

II. Proposed Method

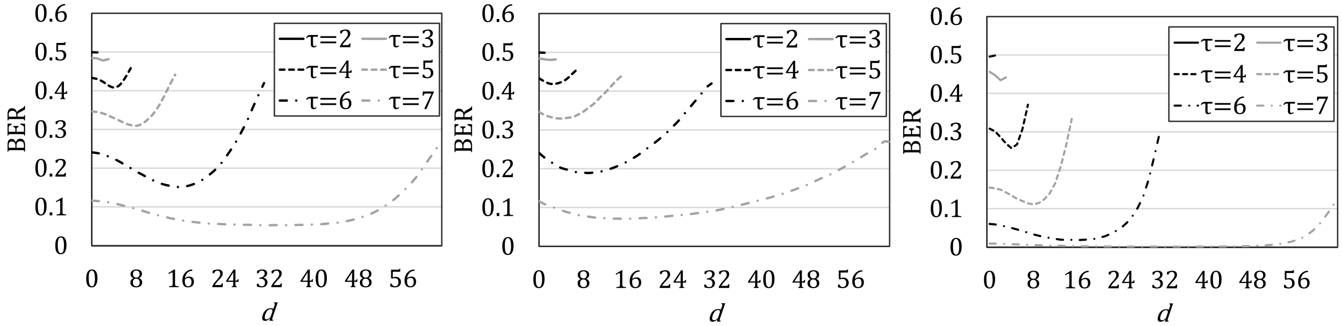
A. Safety Zone

Lossy compression often causes loss of image data, which deteriorates the quality of the inserted data in the encrypted image. As a result, the inserted data, which is of low quality, might lead to failure in authentication, ownership dispute, etc. To address this issue, this paper introduces a region called *safety zone* in the histogram of pixel values. Safety zone is an emptied area defined as a range of pixel value, where no pixel can assume any value in the defined range. said range any value these values. Recall that in [1] the adjustable parameter τ is the number of bit planes utilized for data insertion. Let d be the size of safety zone for $0 \leq d \leq 2^{\tau}/2$, and two implementations of safety zone are shown in Fig.1 for $\tau = 3$. Specifically, 2 selected bins (i.e., 3 and 4 in Fig. 1) are vacated by means of distributing their elements (i.e., pixels) to other bins, i.e., changing the pixel values. Specifically, Method A restricts the distribution process to the immediate adjacent bins, while Method B uniformly distributes the elements to all other non-safety-zone bins. When a pixel value is changed due to lossy image compression or filtering, it is unlikely to spill over the boundaries due to the introduced safety zone (viz., barrier).

B. Majority Vote

Our previous method [1] considers a binary image as the data to be inserted. Therefore, the original image and the data to be inserted are of the same size. To further suppress bit

978-1-7281-3038-5/19/$31.00 ©2019 IEEE



(a) Safety Zone - Method A. (b) Safety Zone - Method B (c) Majority Vote

Fig. 2: BER for different d value when creating safety zone.

TABLE I: Lowest BER and its corresponding d value for each τ considered.

τ	1	2	3	4	5	6	7
BER$_{(a)}$	0.4992	0.4989	0.4784	0.4084	0.3095	0.1515	0.0526
d	0	1	2	4	8	16	31
BER$_{(b)}$	0.4992	0.4989	0.4806	0.4190	0.3293	0.1890	0.0709
d	0	1	2	2	4	9	15
BER$_{(c)}$	0.5043	0.4955	0.4341	0.2573	0.1111	0.0182	0.0013
d	0	0	2	4	8	16	27, 32

error rate of the inserted data after lossy image compression, the data size is first reduced to 25% and repeatedly inserted into the image for 4 times. Data extracted are then corrected by using these 4 sets of data. With this construct, *majority vote* is introduced here. If there is a tie, a majority result is taken over a 3×3 subimage around the target pixel. In the case when there is undecided pixel, a random value, 0 or 1, is considered.

III. Experiments

The proposed method is implemented in Matlab (version R2018b Update 4 9.5.0.1067069). The image processed by our method (i.e., encrypted and contains data) undergoes JPEG compression with a quality factor setting of 75. The Lenna image (512×512) is considered here to verify the basic performance of the proposed method. The data extracted from the compressed image is compared with its original counterpart to compute the bit error rate (BER), i.e., the ratio of *total error bit* to *data size*. Here, smaller BER value implies better performance, and vice versa.

Figure 2 shows the BER values calculated from the original and extracted data for different d. Note that $d = 0$ implements our original method [1]. For both Method A and B, results suggest that BER is reduced when safety zone is introduced. In addition, BER decreases when d increases, but increases again after a certain point, depending on τ. The corresponding d yielding the lowest BER, denoted by $\hat{d}$, is recorded in Table I for each τ. Results suggest $\hat{d}$ occurs around half of the largest d value that can be considered, which agrees with our design. Furthermore, it is observed in Fig. 2(a) and (b) that BER reduces when safety zone is adopted. Moreover, BER is further reduced when majority vote is adopted as suggested by Fig. 2(c).

IV. Conclusions

Two strategies are put forward in this work to improve the robustness of the inserted data against image processing such as lossy compression. Although the size of the binary image data is reduced to 25% of the original size, the bit error rate observed in the extracted data is empirically shown to be improved after applying the proposed strategies. As future work, we want to extend our strategies to address other form of attacks to make the inserted data more robust against image processing.

Acknowledgement

This research is supported by the E-Science fund under the project - Innovative High Dynamic Range Imaging - from Information Hiding to Its Applications (Grant No. 01-02-10-SF0327).

References

[1] R. Ito, K. Wong, S. Ong, and K. Tanaka, "Encryption and data insertion technique using region division and histogram manipulation," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Nov 2018, pp. 1118–1121.

[2] Yiqi Tew, KokSheik Wong, Raphael C.-W. Phan, and King Ngi Ngan, "Separable authentication in encrypted hevc video," *Multimedia Tools and Applications*, vol. 77, no. 18, pp. 24165–24184, Sep 2018.

[3] Simying Ong, Koksheik Wong, and Kiyoshi Tanaka, "A scalable reversible data embedding method with progressive quality degradation functionality," *Image Commun.*, vol. 29, no. 1, pp. 135–149, Jan. 2014.

[4] Zi-Long Liu and Chi-Man Pun, "Reversible data-hiding in encrypted images by redundant space transfer," *Information Sciences*, vol. 433-434, pp. 188 – 203, 2018.

[5] D. Coltuc, P. Bolon, and J. . Chassery, "Exact histogram specification," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1143–1152, May 2006.

An Efficient Speech Recognition Algorithm for Small Intelligent Electronic Devices

Zhichao Zheng[1], Xiaotao Lin[1], Weiwei Zhang[1], Jianqing Zhu[1*] and Huanqiang Zeng[2]
[1]College of Engineering, Huaqiao University, 269 Chenghua North Road, Quanzhou, Fujian, China
Email:{1695111058, jqzhu}@hqu.edu.cn
[2]College of Information Science and Engineering, Huaqiao University, 668 Jiemei Avenue, Xiamen, Fujian, China
Email: zeng0043@hqu.edu.cn

***Abstract*—The speech recognition technology makes it possible for people to communicate with intelligent electronic devices. However,existing speech recognition algorithms are overly complex for small intelligent electronic devices (e.g., mini speakers, intelligent toys, intelligent remote controls, etc.). For this, an efficient speech recognition algorithm is proposed. Firstly, the *Mel-scale Frequency Cepstral Coefficients* (MFCC) is applied to extract features of voices. Secondly, the *Support Vector Machines* (SVM) is used to train speech classification models. Finally, a speech database is collected to validate the proposed algorithm. The speech database contains 500 audio files of 10 speech commands for an electric motor car driving assistant and 550 audio files of 11 speech commands for a intelligent remote control. The proposed method is evaluated via a 5-fold cross-validation, and experiments show that the propose method acquires 94.20% and 88.73% average accuracy rates for the electric motor car driving assistant and the intelligent remote control, respectively.**

***Index Terms*—MFCC, SVM, Speech Recognition**

I. Introduction

The speech recognition technology allows people to communicate friendly with intelligent electronic devices. But, general speech recognition algorithms have a large computation load that usually completed on a remote server, which are not suitable for small intelligent electronic devices (e.g., mini speakers, intelligent toys and intelligent remote controls, etc.). Because these small intelligent electronic devices are cost controlled so that they can not accept a large computation load and additional communication chips. Therefore, it is urgent to design an efficient speech recognition algorithm for small intelligent electronic devices.

Erlin et al. [1] analyzed discriminative features of voices in detail, including loudness and tonal harmony, and designed the audio classifier according to the *nearest neighbor* (NN) criterion. Guo and Li [2] designed a SVM based multi-stage audio classifier. Lu et al. [3] proposed a hidden Markov model based audio classification method. Based on the work of Guo and Li [2], Lin used the wavelet transform to extract sub-band energy and pitch frequency [4] as features of voices.

In this paper, an efficient speech recognition algorithm is proposed. It uses *Mel-scale Frequency Cepstral Coefficients* (MFCC) for representing voices and applies *Support Vector Machines* (SVM) to classify speeches. Moreover, a speech database containing 500 audio files of 10 speech commands for an electric motor car driving assistant and 550 audio files of 11 speech commands for a intelligent remote control is collected to validate the proposed algorithm via a 5-fold cross-validation. The experiments show that the proposed method acquires 94.20% and 88.73% average accuracy rates for the electric motor car driving assistant and the intelligent remote control, respectively.

* Jianqing Zhu is the corresponding author.

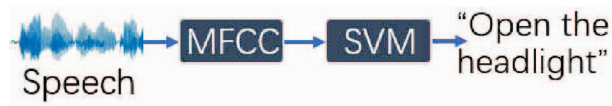


Fig. 1. The flow chart of the proposed method.

II. Method

As shown in Fig. 1, the proposed approach consists of two steps(i.e., MFCC and SVM), which are introduced as follows.

A. MFCC

Due to MFCC [5] having an anti-noise ability, it is applied to extract features of speeches in this paper. Due to the paper length limitation, we can not detailly introduce MFCC. The details of MFCC can be found in [5]. However, one important step of MFCC should be noted. Because of the differences in speed and content of speech commands, the data lengths of speech commands are different, which is harmful to the following SVM. For this, during the MFCC step, frame energy is descending sorted, and only top-128 frames are selected to construct features of speeches, making each speech holds the same dimensional feature vector.

B. SVM

In practice, the interaction between human and small intelligent electronic devices does not require complex speech commands. The kinds of speech commands are limited to a specific electronic product. Therefore, we transform the speech recognition task as a classification task that aims to predict the class label of an input speech file. Considering that the SVM [6] has high efficiency and accuracy, it is applied to finish the speech classification task for small intelligent electronic devices in this paper.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

TABLE I
THE ENGLISH TRANSLATIONS PF SPEECH COMMANDS IN THE SPEECH DATABASE.

electric motor car driving assistant (Part I)		
1.the intelligent assistant	2.intelligent photography	3.start recoding video
4.stop recoding video	5.start the recoding	6.stop the recoding
7.electronic rearview mirror	8.lock video	9.return to the video
10.yadi yadi		

intelligent remote control (Part II)		
1.open the lock	2.closed lock	3.open the headlight
4.close the headlight	5.hello, xiao bao	6.please bright red
7.please bright yellow	8.please bright green	9.please bright blue
10.please bright purple	11.please run lantern mode	

TABLE II
ACCURACY RATE (%) OF THE PROPOSED METHOD.

dataset	fold-1	fold-2	fold-3	fold-4	fold-5	average
Part I	97.00	97.00	98.00	89.00	90.00	94.20
Part II	90.90	90.00	94.55	85.45	82.73	88.73

III. EXPERIMENT

A. Database

In this paper, a new speech database is collected to validate the proposed method. This speech database includes two part, the Part I is for an electric motor car driving assistant and Part II is for an intelligent remote control. As shown in Table I, both Part I and Part II consists of simple speech commands, and the language is the Chinese mandarin. 50 volunteers (i.e., 25 men and 25 women) from different provinces of China contribute their speeches to construct this database. As a result, the Part I contains 500 audio files of 10 speech commands, and the Prat II includes 550 audio files of 11 speech commands.

B. Implementation Detail

The hardware is a notebook with an i5-6300HQ CPU and 8.00 GB memory. The software is MATLAB 2016b. Regarding to the MFCC step, the cepstrum coefficient is 12; alpha value is 0.97; the frequency range is 100-800 Hz. Regarding to the SVM step, the libsvm [6] toolbox is applied. The linear kernel is used and the penalty factor C is set as 1.5. The 5-fold cross-validation is implemented to evaluate the proposed method.

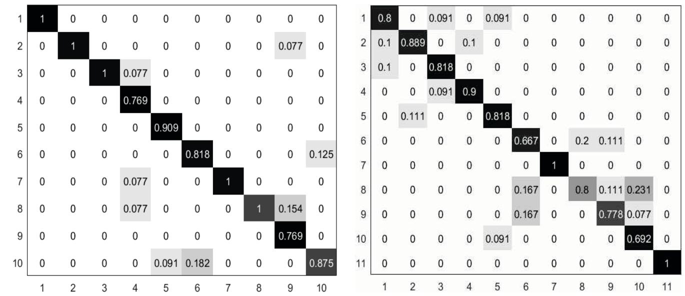
C. Experimental Result

As shown in Table II, on the Part I (i.e., electric motor car driving assistant), the average accuracy rate of the proposed method is 94.20%, and on the Part II (i.e., intelligent remote control), the average accuracy rate of the proposed method is 88.73%. Moreover, Fig. 2 shows the confusion matrixs of Part I and Part II. It can find that only a small number of samples are not on the diagonal, which shows that most of samples are classified correctly.

In addition to the accuracy rate, for processing one sample, the average time cost of the MFCC step is 12.80 ms and the SVM step is 2.87 ms. Consequently, the total time for processing one sample is only about 16 ms, which shows that the proposed method is able to be expected to run on small intelligent electronic devices.

IV. CONCLUSION

In this paper, an efficient speech recognition algorithm using *Mel-scale Frequency Cepstral Coefficients* (MFCC) and *Support Vector Machines* (SVM) is proposed. Experiments are evaluated on a newly collected speech database by a 5-fold cross-validation. The speech database contains 500 audio files of 10 speech commands for an electric motor car driving assistant and 550 audio files of 11 speech commands for a intelligent remote control. The proposed method acquires 94.20% and 88.73% average accuracy rates for the electric motor car driving assistant and the intelligent remote control, respectively.



(a) (b)

Fig. 2. Confusion matrixs. (a) is for the Part I (electric motor car driving assistant) and (b) is for the Part II (intelligent remote control).

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under the Grants 61976098, 61602191, 61871434, 61802136 and 61876178, in part by the Natural Science Foundation of Fujian Province under the Grant 2018J01090, in part by the Natural Science Foundation for Outstanding Young Scholars of Fujian Province under the Grant 2019J06017, in part by the Open Foundation of Key Laboratory of Security Prevention Technology and Risk Assessment, People's Public Security University of China under the Grant 18AFKF11, in part by the Science and Technology Bureau of Quanzhou under the Grants 2018C115R, 2017G027 and 2017G036, in part by the Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University under the Grants ZQN-PY418 and ZQN-YX403, and in part by the Scientific Research Funds of Huaqiao University under the Grants 16BS108.

REFERENCES

[1] E. Wold, T. Blum, D. Keislar, and J. Wheaten, "Content-based classification, search, and retrieval of audio," *IEEE MultiMedia*, vol. 3, no. 3, pp. 27–36, Fall 1996.

[2] G. Guo and S. Z. Li, "Content-based audio classification and retrieval using svm learning," in *IEEE Pacific-Rim Conference on Multimedia(PCM)*, vol. 8, no. 5, 2000, pp. 619–625.

[3] L. Jian, C. Yi-song, S. Zheng-xing, and Z. Fuyan, "Automatic audio classification by using hidden markov model," *Journal of Software*, vol. 13, no. 8, pp. 1593–1597, 2002.

[4] Chien-Chang Lin, Shi-Huang Chen, Trieu-Kien Truong, and Yukon Chang, "Audio classification and categorization based on wavelets and support vector machine," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 644–651, Sep. 2005.

[5] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of mfcc," *Journal of Computer Science and Technology*, vol. 16, no. 6, pp. 582–589, 2001.

[6] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

A Transformation for Polar Code BP Decoding

Yi-Ru Chen
Grad. Inst. of Communication Engineering,
National Taiwan University
Taipei 106, Taiwan.
(e-mail: r07942123@ntu.edu.tw)

Huang-Chang Lee
Dept. of Electrical Engineering,
Chang Gung University,
Taoyuan City 33302, Taiwan.
(e-mail: akula.lee@gmail.com)

Mao-Chao Lin
Grad. Inst. of Communication Engineering,
National Taiwan University
Taipei 106, Taiwan.
(e-mail: mclin@ntu.edu.tw)

***Abstract*—A transformation technique has been designed, where the encoding result of the polar code can be transfered to an low-density parity-check (LDPC) code. The belief-propagation decoding can be applied to the received sequence according to the parity-check matrix of the LDPC code. The transfered polar code can provide the same error-rate of the LDPC code, and its performance can be further improved if the success-cancellation decoding is avaliable.**

***Index Terms*—Belief propagation, error-control codes, polar code, low-density parity-check (LDPC) codes.**

I. Introduction

Infinite-length polar code [1] have been proven to achieve the capacity for binary-input symmetric memory-less channels, but it cannot outperform low-density parity-check (LDPC) [2] code with limited code length using successive cancellation (SC) decoding. Although the error-rate performance can be improved by using CRC-list-SC decoding [3], these SC-based decoding schemes suffer from longer latency, and the throughput is limited due to their serial nature. On the other hand, belief propagation (BP) [4] decoding has the intrinsic advantage of parallel processing, and can be a potential solution for high throughput. However, the original BP decoding applied to polar code is based on the bipartite graph of encoding, it cannot provide the parity-check feature and achieves only limited performance.

In [5] and [6], a short LDPC code is used as an outer code and is concatenated with an inner polar code. Its error-rate performance is comparable to that of the CRC-list-SC decoding with eight CRC bits and list size equals to two (CRC-8, L=2), but the rate of overall concatenated code is lower than the original polar code, and both polar and LDPC decodings must be processed in order to achieve the final decoding results.

In this paper, an alternative solution has been proposed. A transformation matrix has been designed and serves as an inner code. The encoding result of the outer polar code can be transfered to an equivalent LDPC code, then the BP decoding for LDPC code can be directly applied to the received sequence. If the equivalent LDPC code cannot be successively decoded, SC decoding of the polar code can be used and the remaining error bits are possibly solved.

II. Design of The Proposed Transfer Matrix

The system model is presented in 1. The generator matrix of the polar code is denoted as $\mathbf{G}_{polar}$, it can be achieved after removing the rows corresponding to the frozen bits. For a (N, K) polar code, the size of $\mathbf{G}_{polar}$ is K-by-N. The parity-check matrix of the LDPC code is denoted as $\mathbf{H}_{ldpc}$, its size is $(N-k)$-by-N. The transfer matrix is denoted as $\mathbf{T}$, its size is N-by-N, then the overall concatenated code can be kept as $R = K/N$.

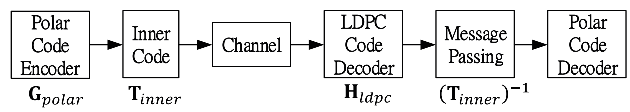


Fig. 1. System model of the proposed polar code decoder.

$$\mathbf{G}_{polar}\mathbf{T}_{inner}\mathbf{H}_{ldpc}^{\mathrm{T}} = \mathbf{0} \tag{1}$$

In order to decode the received sequence using BP decoding of LDPC code, the condition of (1) must be satisfied. Consider the polar code of $R = 1/2$. $\mathbf{G}_{polar}$ can be represented as a combination of two $N/2$-by-$N/2$ sub-matrices $\mathbf{G}_L$ and $\mathbf{G}_R$. A similar representation is shown for $\mathbf{H}_{ldpc}$, whose left and right helf parts are denoted as $\mathbf{H}_L$ and $\mathbf{H}_R$, respectively. Then (1) can be re-writen as (2), the transfer matrix $\mathbf{T}_{inner}$ is separated into four $N/2$-by-$N/2$ sub-matrices, the upper-right and lower-left ones are all zero matrices, and the upper-left and lower-right ones are denoted as $\mathbf{U}$ and $\mathbf{D}$, respectively.

$$\begin{bmatrix}\mathbf{G}_L & \mathbf{G}_R\end{bmatrix}\begin{bmatrix}\mathbf{U} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{H}_L^{\mathrm{T}}\\ \mathbf{H}_R^{\mathrm{T}}\end{bmatrix} = \mathbf{0}, \tag{2}$$

$$\mathbf{G}_L\mathbf{U}\mathbf{H}_L^{\mathrm{T}} + \mathbf{G}_R\mathbf{D}\mathbf{H}_R^{\mathrm{T}} = \mathbf{0}, \tag{3}$$

$$\mathbf{G}_L\mathbf{U}\mathbf{H}_L^{\mathrm{T}} = \mathbf{G}_R\mathbf{D}\mathbf{H}_R^{\mathrm{T}}. \tag{4}$$

$$\mathbf{U} = (\mathbf{G}_L)^{-1}\,\mathbf{G}_R\mathbf{D}\mathbf{H}_R^{\mathrm{T}}\left(\mathbf{H}_L^{\mathrm{T}}\right)^{-1} \tag{5}$$

$$\mathbf{U} = (\mathbf{G}_L)^{-1}\,\mathbf{G}_R\mathbf{H}_R^{\mathrm{T}}\left(\mathbf{H}_L^{\mathrm{T}}\right)^{-1} \tag{6}$$

Follow the process from (2) to (6), and set $\mathbf{D}$ as an identity matrix, the transfer matrix that can convert a polar code to an LDPC code is derived.

In this work, the transfer matrix $\mathbf{T}_{inner}$ surves as a rate-1 inner code. If the received sequence cannot be successfully

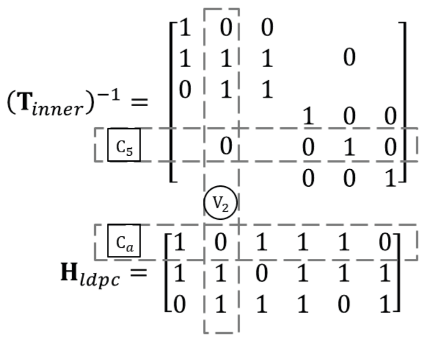


Fig. 2. The representation of the $(\mathbf{T}_{inner})^{-1}$ transformation.

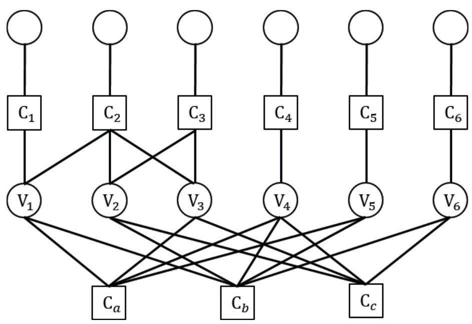


Fig. 3. Graph connection of $(\mathbf{T}_{inner})^{-1}$ and $\mathbf{H}_{ldpc}$.

decoded using the BP decoder based on $\mathbf{H}_{ldpc}$, the LLR values of the temparly decoding result can be transfered following the message passing algorithm according to the bipartite graph corresponding to $(\mathbf{T}_{inner})^{-1}$, and the SC decoding of polar code can be used to further improve the performance. The operations of the message passing based on $(\mathbf{T}_{inner})^{-1}$ is shown in Fig. 2 and Fig. 3, where the circle and square nodes respectively denote the processes of variable nodes and check nodes in LDPC code.

III. Numberical Evaluation

When a (1024,512) polar code is evaluated in an AWGN channel, the performance of block-error rate (BLER) and bit-error rate (BER) are respectively presented in Fig. 4 and Fig. 5. The simulation results shows that if only LDPC decoding of 20-iterations is used, the the proposed method can provide a similar FER performance compared to that of [6] and CRC-list-SC with L=2. The SC decoding can slightly improve the BLER performance. If the 50 iterations is avaliable, the BLER can be significantly improved. However, as shown in Fig. 5, due the error propagation, the BER of SC decoding is even worse than that of only use LDPC decoding.

IV. Conclusion

In this work, we present a design flow to transfer a polar code to an LDPC code. Although the application of this design flow is limit to rate-1/2 polar code, the simulation results show that the high throughput BP decoding is possible to achieve a competitive error-rate performance compared to that of the CRC-list-SC decoding.

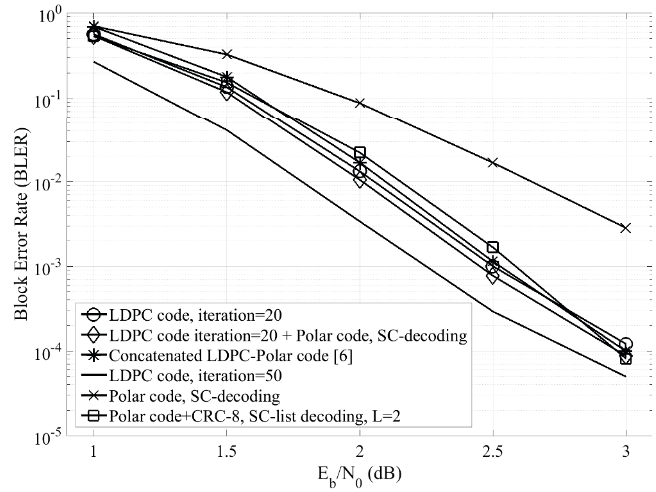


Fig. 4. The BLER performance for the (1024,512) polar code.

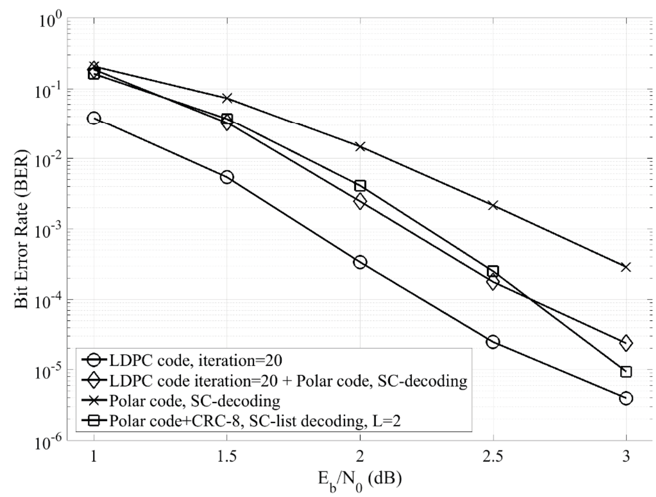


Fig. 5. The BER performance for the (1024,512) polar code.

References

[1] E. Arkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory,* vol. 55, no. 7, pp. 3051-3073, Jul. 2009.S.

[2] R. G. Gallager, "Low-density parity-check codes," M.I.T. Press, 1963.

[3] I. Tal and A. Vardy, "List decoding of polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*,2011,pp.1-5.

[4] Guiping Li, Jianjun Mu, Xiaopeng Jiao, Junjun Guo, and Xiaohang Liu, "Enhance belief propagation decoding of polar codes by adapting the parity-check matrix," Eurasip journal on wireless communication and network.

[5] J. Guo M. Qin A. Guillen I Fabregas P. H. Siegel "Enhanced belief propagation decoding of polar codes through concatenation," *Proc. IEEE International Symposium on Information Theory,* pp. 2987-2991 2014.

[6] S. M. Abbas Y. Fan J. Chen C. Y. Tsui "Concatenated LDPC-polar codes decoding through belief propagation," *IEEE Int. Symp. on Circuits and Systems (ISCAS),* pp. 1-4 2017.

Subjective Interpupillary Distance of Measurement Technique

Jih-Yi Liao
Chief Executive Officer
Loocadi ,INC. No.29199990
Taipei, Taiwan
id4chair1616@hotmail.com

Shih-Tsung Chang
Optometry Science and Technolgy
Feng Chia University
Taichung, Taiwan
cheung2379@yahoo.com

Chao-Han Wu
Department of Optometry
Mackay Junior College of Medicine, Nursing, and Management
Taipei, Taiwan
s521@mail.mkc.edu.tw

Der-Chin Chen
Department of Electrical Engineering
Feng Chia University
Taichung, Taiwan.
kanatasan.tw@yahoo.com.tw

Chung-Ping Chen
Graduate Institute of Electronics Engineering
National Taiwan University
Taipei, Taiwan
cpchen@ntu.edu.tw

Cheng-Ke Hsu
College of Nursing and Health Sciences
Vision Science Master Program
Da-Yeh University
Changhua, Taiwan
given4026@yahoo.com.tw

***Abstract*—In this paper, a new optical technique, two movable slit plate modules and the subjective optometry method was developed to examine Interpupillary distance. The evaluation of binocular vision are dependent on the three-dimensional structure of natural scenes. The eye position and the fusion image of the eyes are fine-tuned and the alignment is assured by through the vision of subject with the trial frame and a movable slit plate module after full optometry correction procedure. This novel method is to success and replace of new optical lenses, the error of prismatic imbalance estimated has an accuracy of $0.12^{\triangle}$. It also changes the refraction test from an objective method to a subjective method in order to meet the subject's most comfortable binocular visual fusion.**

Keywords—Subjective, Interpupillary Distance, Rotation Center of Eye, Binocular vision.

I. INTRODUCTION

Interpupillary Distance (IPD) is the distance between the pupils when seeing a distant target. Near PD is the distance between the pupils when seeing a near target. When fitting a subject for glasses it is essential that the subject's pupils are aligned with the optical centers of the lenses. In this study the use of the slit unit combined with the subjective refraction to measure the IPD.

II. HUMAN PERCEPTION OF PARALLAX

A. 3D Visual Discomfort

To maintain single and clear binocular vision, neural processes that accomplish accommodation and vergence are performed cooperative, cross-coupled processes vergence-accommodation (VA) and accommodation-vergence (AV).

Visual discomfort including improper stereography [1], or unusual extreme statistical distribution of depths in the (3D) content, especially in salient regions [2]–[4], or from rapid movements in depth [5]–[6]. This kind of visual discomfort could(give) rise to or manifest as several physiological symptoms, including eyestrain, a feeling of pressure in the eyes, reduced sensitivity of visual perception, reduced ability of AV , headaches and even neck pain.

B. On The Horopter

If locus of points does not fall on the horopter, then it may have a large retinal disparity, thereby leading to a failure of fusion. If the retinal disparity falls beyond Panum's fusion area, stereoscopic fusion does not take place but an effort may still be made to achieve fusion, which may trigger diplopia and/or asthenopia (eye strain) , shown in Fig. 1.

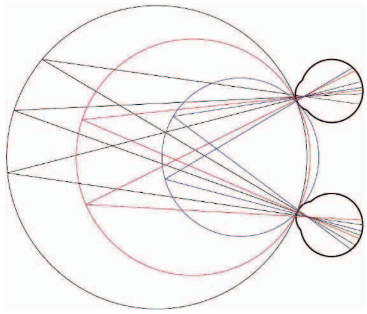


Fig. 1. Panum's fusion area

III. METHODS

This research is to verify the eye position and the fusion image of the eyes through the vision of subject with the trial frame and a movable slit plate module after full optometry correction procedure.

A. Stenopeic slit

The field of view (FOV) reduction technology adopts the "far-distance method" for measurement PD. As can seen the far-distance method leads to a small FOV, whereas the near-distance method leads to a large FOV. Firstly, the subject sees a large field of view of letter chart at close range, and then observes the letter chart at far-sighted range to reduce the field of view for improving measurement accuracy as shown in Fig. 2.

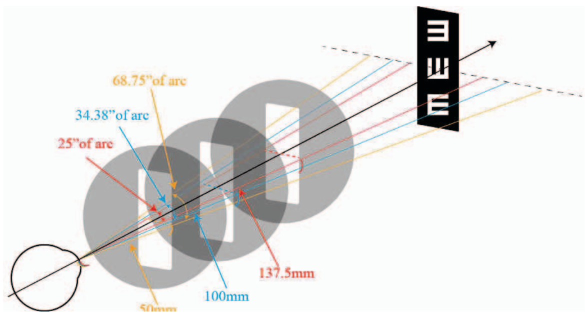


Fig. 2. Stenopeic slit with the field of view expectation value is 3 times the letter chart size, that is, 25' angle, 2 times the letter chart size, which angle is 16.67'.

B. *System Architecture*

The system architecture is to install the trial frame module in front of the subject, and the subject must complete the optometry procedure before performing the PD measurement. The system includes one subject, a rangefinder module, an letter chart detection module, one test frame module, and one CCD camera. After full optometry correction procedure, "error correction" are finished and then a movable slit plate has been added to the trial frame. It can verify the right eye position and the image fusion of the subject by distance method, shown in Fig. 3.

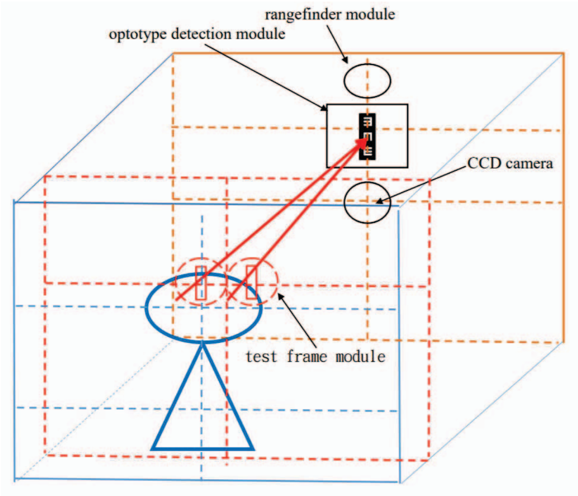


Fig. 3. The Ocularcenter of Rotation exercise Positioning system

IV. Results

Compared with the traditional objective method of examining PD as the basic procedure of initial optometry procedure, subjective optometry method is better to reduce anomalous binocular vision. However, in order to determine the optometry prescription of the subject, there is still no good then subject's self-perception. The result of our research with subject's self-perception will solve by using the shortcomings of the traditional optometry.

A. *The subjective and objective optometry method*

The objective optometry method of examining PD value cannot be obtained before the lens assembling. Therefore, the lens shaped is irreversible after assembling, and the corrected refractive error of subject causes different tolerance values. The optometry method is to treat the vision of subject as the first consideration in order to lower the tolerance of the final examination and then to enhance the fusion of the binocular vision. The subjective and objective (ANSI Z80.1) optometry method listed in Table I [7]. on horizontal or vertical. Degree tolerance data.

TABLE I. Relative Uncertainty of Horizontal and Vertical Prism Imbalance [7]

Axis / type	Vertical axis 0~±3.375D	Vertical axis >±3.375D	horizontal axis 0~±2.75D	horizontal axis >±2. 75D
Objective (ANSI Z80.1)	≦0.33△	≦1.0mm reference point	≦0.67△	≦1.0mm Pupil distance gap
Subjective (Stenopeic Slit)	≦0.12△	≦1.0mm reference point	≦0.12△	≦1.0mm Pupil distance gap

Mostly, the traditional objective method of measuring PD is judged by the optometrist or the optometry instruments. In this research, the subjective (self-sense) method of PD measurement is judged by the subject after the optometry correction, and the PD is judged by the subjective feeling (self-sense). We analysis five PD measurement methods in this experiment. The Method 1 is a collimated measurement without any optical component. Method 2, 3, and 4 are non-collimated measurements method by other optical components. And base on the result of the testing distance (600cm) in Fig. 4. It shows that the new method has certain regularity with the traditional PD methods which explants the empirical research is a feasible method.

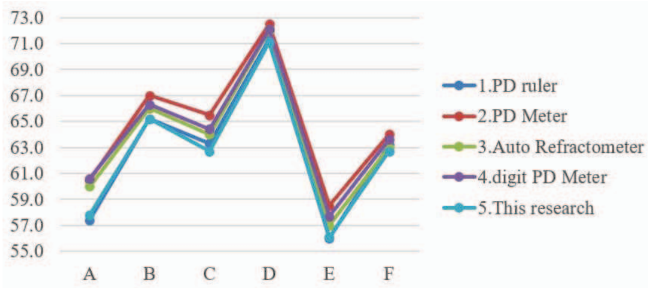


Fig. 4. Analysis of results

B. *Future prospects*

1) The IPD data from the subjective interpupillary distance measurement method are adapted to the achtual visual experience of subject. The optical inspection tools is low cost.

2) When the line of sight of both eyes passes through the two lens and the **Stenopeic slits,** and then the one chart is observed by eyes, the pupil distance and the center position of eye rotation can be obtained through the calculation of the similar triangle.

3) At present, patent applications have been filed in Taiwan and China. We will develop optical inspection tools for future optometry personnel.

References

[1] F. Zilly, J. Kluger, and P. Kauff, "Production rules for stereo acquisition," Proc. IEEE, vol. 99, no. 4, pp. 590–606, 2011. 7, Mar. 2013.

[2] H. Sohn, Y. J. Jung, S. Lee, and Y. M. Ro, "Predicting visual discomfort using object size and disparity information in stereoscopic images," IEEE Trans. Boradcast., vol. 59, no. 1, pp. 28–3

[3] K. Lee, A. K. Moorthy, S. Lee, and A. C. Bovik, "3D Visual activity assessment based on natural scene statistics," IEEE Trans. Image Process., vol. 23, no. 1, pp. 450–465, Jan. 2014.

[4] H. Kim, S. Lee, and A. C. Bovik, "Saliency measurement on stereoscopic videos," IEEE Trans. Image Process., vol. 23, no. 4, pp. 1476–1490, Apr. 2014.

[5] I. Juricevic, L. Land, A. J. Wilkins, and M. A. Webster, "Visual discomfort and natural image statistics," Perception, vol. 39, no. 7, pp. 884–899, 2010.

[6] T. Kim, J. Kang, S. Lee, and A. C. Bovik, "Multimodal interactive continuous scoring of subjective 3D video quality of experience," IEEE Trans. Multimedia, vol. 16, no. 2, pp. 387–402, Feb. 2014.

[7] Clifford W. BrooksandIrvin M. Borish,"System for ophthalmic dispensing, 3rded,"Elsevier Inc, 2007.

Cross Conditional Network for Speech Enhancement

Haruki Tanaka[†1‡], Yosuke Sugiura[†2], Nozomiko Yasui[†2], Tetsuya Shimamura[†2], Ryoichi Miyazaki[†1]
[†1]National Institute of Technology, Tokuyama College, Yamaguchi, Japan,
[†2]Saitama University, Saitama, Japan
[‡]i14tanaka@tokuyama.kosen-ac.jp

***Abstract*—In the signal processing field, there is a growing interest in speech enhancement. Recently, a lot of speech enhancement methods based on the deep neural network have been proposed. Mostly, these networks, such as SEGAN, Wave-U-Net, adopt the autoencoder structure. In this paper, we propose the cross conditional network for speech enhancement based on SEGAN architecture. The proposed network has two Auto-Encoder, where the mutual latent vector is composed of the concatenated vector of these encoder outputs. In the experiments, we show that the proposed method exceeds SEGAN in terms of the objective evaluation measure by PESQ.**

***Index Terms*—Speech Enhancement, Deep Neural Network, Autoencoder, Generative Adversarial Nets**

I. Introduction

The speech enhancement is often used as the front-end processing for speech communication devices such as music players, smartphones, and hearing aids. The speech enhancement is an important signal processing technique which aims to improve the speech quality and intelligibility of noisy speech. Even now, many researchers are striving to establish fundamental solutions for this problem.

Recently, a speech enhancement architecture using Deep Neural Network (DNN) [1] has been proposed. DNN automatically learns the representation of the relationship between data and label data which are paired by a lot of data set. It was reported that the speech enhancement using DNN can provide higher speech quality than the conventional model-based methods. Among speech enhancement using DNN, the waveform-based autoencoder such as SEGAN [2], Wave-U-Net [3], and so on, is an excellent architecture providing particularly outstanding performance.

In this paper, we propose a new structure of the Auto-Encoder to gain the speech enhancement performance.

II. Speech Enhancement Generative Adversarial Network

In this paper, Speech Enhancement Generative Adversarial Network (SEGAN) [2] is used as a base architecture. SEGAN, which has been proposed by S. Pascual et al., is one kind of the supervised adversarial generative model modified for the speech enhancement task.

SEGAN consists of two DNNs, which are the generator and the discriminator. The generator is based on the end-to-end autoencoder, which outputs a waveform directly from the input waveform. The discriminator outputs "0" or "1" correspondingly to whether the input is the real distribution or the generator output. In most of the DNN based speech enhancement including SEGAN, model training is based on only the information of the clean speech waveform. In other words, there is no mechanism to learn any other features such as the information of noise. Therefore, we aim to improve the speech enhancement performance by introducing a new network to learn the noise information.

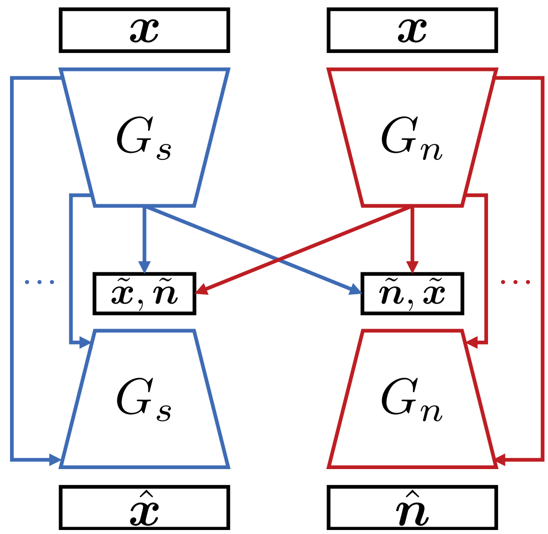


Fig. 1. Generator Architecture.

III. Cross Conditional Network for Speech Enhancement

We show the proposed network architecture in the Fig. 1. The generator consists of two symmetric autoencoders. The two networks play the roles of the clean speech generator and the containing noise generator, respectively. The left-side network, G_s, generates the estimated speech $\hat{\boldsymbol{x}}$ from the input noisy speech $\boldsymbol{x}$. The right-side one, G_n, generates the estimated noise $\hat{n}$ from the input noisy speech. The latent vector in each network is composed of not only the own encoder output but also the another encoder output. Focusing only on the one-side network, this concatenation of the latent vector is similar to conditional adversarial network [4]. Thus we call the proposed architecture cross conditional network. The number of dimensions per layer, being it samples by features, is

TABLE I
PESQ MEASUREMENT.

	Noisy	SEGAN	Proposed
PESQ	1.97	2.16	2.22

16382×1, 8192×16, 4096×32, 2048×32, 1024×64, 512×64, 256×128, 128×128, 64×256, 32×256, 16×512, 8×1024. The activation function is the Parametric Rectified Linear Unit (PReLU) except for the output in the generator, and Leaky ReLU in the discriminator. At the output layer, the activation function is tanh function.

On the basis of SEGAN, the objective function of the proposed network is defined by the followings.

$$\begin{aligned}\min_{G_s,G_n} V(G_s,G_n) &= \frac{1}{2}E_{\boldsymbol{x},\boldsymbol{x}_c\sim p_{\mathrm{data}}(\boldsymbol{x})}\left[\|D_s(G_s(\boldsymbol{x},\tilde{\boldsymbol{n}}),\boldsymbol{x}_c)-1\|_1\right] \\ &+ \frac{1}{2}E_{\boldsymbol{x},\boldsymbol{n}\sim p_{\mathrm{data}}(\boldsymbol{x})}\left[\|D_n(G_n(\boldsymbol{x},\tilde{\boldsymbol{x}}),\boldsymbol{n})-1\|_1\right] \\ &+ \lambda\|G_s(\boldsymbol{x},\tilde{\boldsymbol{n}})-\boldsymbol{x}\|_1 \\ &+ \lambda L_1+\lambda_2 L_2+\lambda_3 L_3, \qquad (1)\end{aligned}$$

where

$$\begin{aligned}L_1 &= \|G_n(\boldsymbol{x},\tilde{\boldsymbol{x}})-\boldsymbol{n}\|_1, \\ L_2 &= \|\|G_s(\boldsymbol{x},\tilde{\boldsymbol{n}})+G_n(\boldsymbol{x},\tilde{\boldsymbol{x}})\|_1-\|\boldsymbol{x}\|_1\|_1, \\ L_3 &= \|\tilde{\boldsymbol{x}}-\tilde{\boldsymbol{x}}_c\|_1+\|\tilde{\boldsymbol{n}}-\tilde{\boldsymbol{x}}_n\|_1. \qquad (2)\end{aligned}$$

In the above (1) and (2), $\boldsymbol{x}$ is noisy speech, $\boldsymbol{x}_c$ is clean speech, and $\boldsymbol{n}$ is noise. The variable with the tilde $\tilde{\boldsymbol{a}}$ means the generator output with the input $\boldsymbol{a}$. G_s, G_n are the generators for speech and noise, respectively, and D_s, D_n are the discriminators for speech and noise, respectively. λ is the regularization factor. The objective function of the proposed method consists of three loss functions as well as the original loss function of SEGAN, where the three losses are the reconstruction error of speech, L_1, the reconstruction error of noise, L_2, and the approximation error of latent vectors, L_3. p_{data} indicates the dataset distribution. $E[\cdot]$ is the expectation operator. Unlike SEGAN, the first two terms on the right side are absolute errors.

The settings for training in this experiment are the followings: the number of epochs is 100 and the batch size is 200. The optimization method is AdaBound [7], the learning rate of the generators is 0.0005 and one of the discriminators is 0.00001, the weight decay rate is 0.0005. Training of discriminators follow the reference [2].

IV. Experiment

A. Setup

In this experiment, we evaluate the speech quality of the enhanced speech. The same data set as SEGAN is used as follows. As the speech data, we use 30 speakers from the Voice Bank Corpus [5] for learning and 2 speakers for test. The noise data for training is constructed by 8 types of noise selected from [6] and 2 types of artificially generated noise. The noisy speech for training is artificially created by mixing the speech and noise so that Signal-to-Noise ratio (SNR) is 15, 10, 5, and 0 [dB], respectively. The test dataset is a pair of noise-mixed speech and its clean speech that are artificially added to the SNR of 17.5, 12.5, 7.5, and 2.5 [dB] from [6], respectively. All of the sounds are down-sampled from 48 [kHz] to 16 [kHz]. The two datasets used in training and test are different.

For evaluating the quality of the estimated speech, we adopt the broadband version of Perceptual Evaluation of Speech Quality (PESQ) [8] standardized by ITU-T. PESQ takes from -0.5 to 4.5, and 4.5 indicates no deterioration from the original signal. In order to verify the effectiveness of our proposed method, we compare PESQ for noisy speech, SEGAN, and the proposed method.

B. Result

Table I shows PESQ for noisy speech, SEGAN, and the proposed method. From the table I, PESQ of the proposed method is 2.22, which is higher than SEGAN by 0.06 points. This indicates that the speech quality can be improved by learning noise information.

V. Conclusion

In this paper, by learning not only the speech information but also the noise information, the speech quality of the proposed method exceeded that of SEGAN.

Future work, we aim to improve speech enhancement performance by using other speech features and changing the discriminator structure. Furthermore, we will apply our cross conditional architecture to Wave-U-Net.

References

[1] X. Lu, Y. Tsao, S. Matsuda and C. Hori, "Speech enhancement based on deep denoising autoencoder." Proc. of Interspeech 2013, pp.436-440, 2013.
[2] S. Pascual, A. Bonafonte and J. Serr, "SEGAN: Speech enhancement generative adversarial network," Proc. of Interspeech 2017, pp.3642-3646, 2017.
[3] C. Macartney and T. Weyde, "Improved speech enhancement with the Wave-U-Net," arXiv preprint arXiv:1811.11307, 2018.
[4] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," arXiv preprint arXiv: 1411.1784, 2014.
[5] C. Veaux, J. Yamagishi and S. King, "The voice bank corpus: Design, collection and data analysis of a large regional accent speech database," Proc. of Oriental COCOSDA, pp.1-4, 2013.
[6] J. Thiemann, N. Ito and E. Vincent, "The diverse environments multi-channel acoustic noise database: A database of multichannel environmental noise recordings," Journal of the Acoustical Society of America, vol.133, no.5, pp.3591, 2013.
[7] L. Luo, Y. Xiong, Y. Liu and X. Sun, "Adaptive gradient methods with dynamic bound of learning rate," Proc. of ICLR, 2019.
[8] P.862.2: Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs, ITU-T Std. P.862.2, 2007.

Baby Care System Design for Multi-Sensor Applications

Tsung-Hsun Wu
Computer Science and Information Engineering
National Cheng Kung University
Tainan, Taiwan, R.O.C.
P78054026@mail.ncku.edu.tw

Pei-Yin Chen
Computer Science and Information Engineering
National Cheng Kung University
Tainan, Taiwan, R.O.C.
pychen@mail.ncku.edu.tw

***Abstract*— Since a baby will be taken care of by the health care worker after the birth, and each health care worker is taking care of multiple babies, it is necessary to regularly understand the physical condition of each baby and make a record, in order to reduce the workload of the health care workers. Therefore, for this issue, the Microchip 16bits architecture dsPIC30F4011 is adopted as the control core in this study to construct a "low-cost, high-efficiency, intelligent baby care system." Using the control advantages of a microcontroller, with the relevant control and sensing circuits designed by the user, through measuring various physiological signals of the baby, and constructing a monitoring screen of the LabVIEW graphical language in the computer interface, the wireless transmission method is used for making a remote wireless connection with the care vehicle to facilitate the monitoring by the health care worker. In addition, the motor can be driven to drive the ambulance to swing back and forth to achieve the effect of comforting the baby, and the workload of the health care worker can be greatly reduced, so that it can be most efficiently applied to human beings.**

Keywords—baby care, multi-sensor, physiological signals

I. Introduction

The intelligent baby care system designed mainly uses the Microchip 16bis architecture chip dsPIC30F4011 [1] as the control core, and the whole system mainly performs the following:

- To obtain blood oxygen saturation (SaO_2) [2] through a Penetrating blood oxygen probe and a designed filter amplifier circuit [3].
- The heart rate can be obtained through the blood pressure waveform [4].
- The temperature sensing circuit measures the body temperature of the baby to determine whether the baby is abnormal in body temperature.
- The humidity sensing circuit detects the condition of the diaper to determine whether the diaper should be replaced.
- The sound capture circuit with a vibration sensor determines the baby's crying.

This system uses LabVIEW [5] graphical programming language to construct a monitoring screen on the PC side. The measured information is transmitted to the PC through XBee wireless transmission, allowing users monitoring various physiological signals. If the baby's blood oxygen saturation (SaO_2), heart rate, body temperature or diaper is abnormal, a warning message will be transmitted to the monitoring terminal. At this time, the health care worker can go to check the status when they find the message, except when the baby cries, the microcontroller will immediately drive the DC motor reversing, and drive the upper part of the car care to shake it back and forth to calm the baby's mood.

II. Design and Verification

Figure 1 is a system block diagram of the intelligent baby care system.

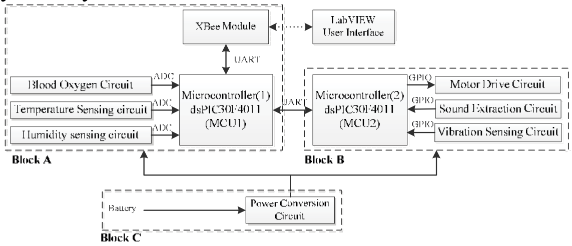


Figure 1. System block diagram

The overall function realization is mainly achieved through the program architecture diagram in Figure 2.

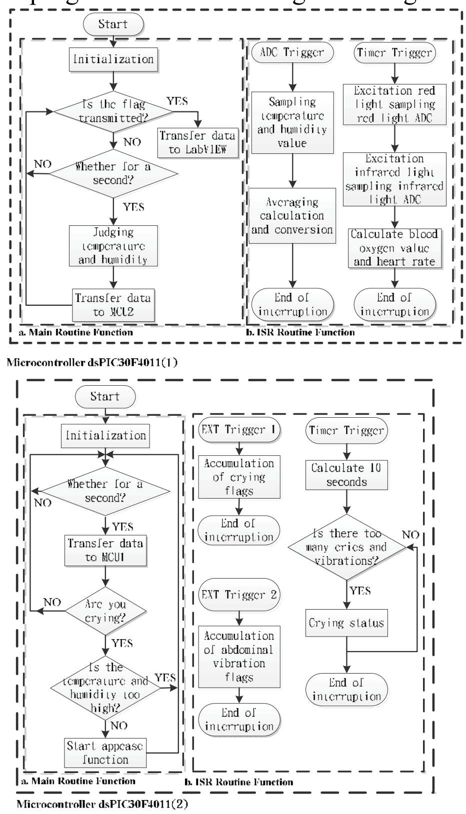


Figure 2. MCU firmware architecture diagram

The whole system is mainly divided into two parts. The first part is the MCU1 with blood oxygen, temperature and humidity sensing circuits, and the second part is MCU2 with motor drive, sound and vibration sensing circuits. The communication part of these two MCUs is mainly through the UART for data transmission.

The power required for the overall system is 3.3V, ±5V and 12V, so the power supply part uses 5 12V batteries, and the three voltage regulator ICs are 7805, 7905 and LM3940 respectively. Figure 3 is a block diagram of the power conversion circuit.

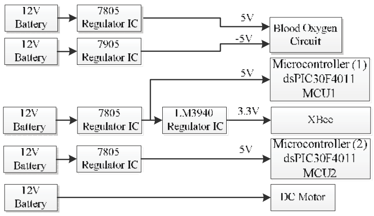


Figure 3. Power conversion circuit of block diagram

Blood oxygen measurement part of the study, red light and infrared light are first used as sensing light sources to sense changes in oxygenated hemoglobin$_2$ (HbO_2) and deoxygenated hemoglobin(Hb) in human arteries.

Measurements in this system are based on the Beer-Lambert Law [6], which probes into the variation in incident and exits light intensity. Furthermore, the relationship to the arterial blood oxygen saturation (SaO_2) can be derived.

The architecture diagram is shown in Figure 4. Through high-speed microcontroller and hardware circuit to drive the oxygen probe, the desired blood pressure waveform is obtained through a self-designed sensing and filtering conversion circuit, and the ADC is built in through the microcontroller. The corresponding value is obtained by conversion, and the blood oxygen concentration value is calculated.

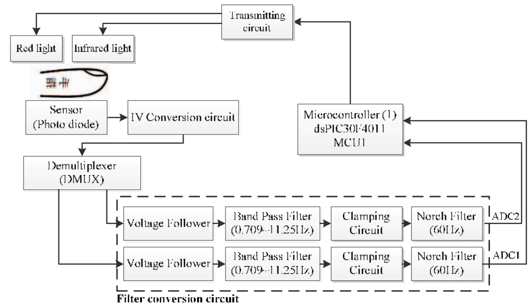


Figure 4. Blood Oxygen circuit architecture diagram

The Circuit test diagram is shown in Figure 5. After the signal passes through the band pass filter, the blood pressure wave AC signal (Figure A) is taken out. In order to operate the voltage in the ADC conversion range, the clamp circuit is used to boost the voltage (Figure B). Finally, the room-rejection filter is used to remove the room-temperature 60Hz noise to obtain the blood pressure waveform (Figure C).

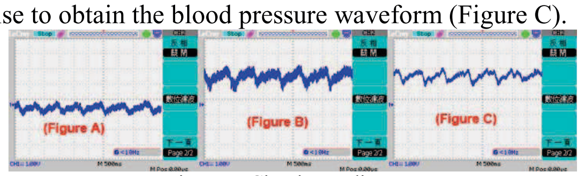


Figure 5. Circuit test diagram

III. Experimental Results and Discussion

Figure 6. LabVIEW User interface

The body structure of the car is mainly designed by using SolidWorks 3D drawing software, with the self-designed control and sensing circuit, and communicated and monitored by wireless transmission with the LabVIEW human-machine interface in Figure 6 on the PC side, to achieve baby care functions.

Figure 7. The overall system architecture diagram of the intelligent baby system.

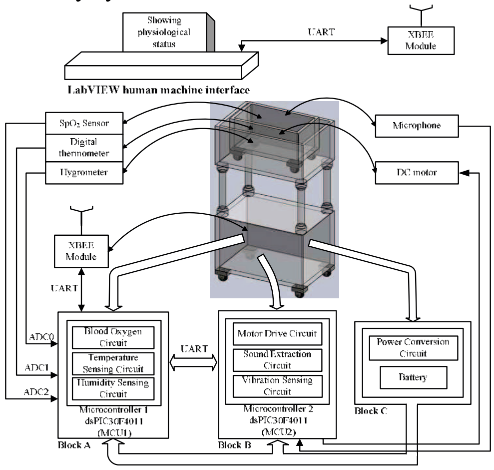


Figure 7. system architecture diagram

The actual installation diagram is shown in Figure 8.


Figure 8. Actual installation diagram

IV. Conclusions

In view of the fact that health care workers are often in short supply, when a baby's physiological signals can be automatically monitored through the computer, it can configure the human most effectively. In this project, a baby's abnormal state can be monitored through computer monitoring to automatically achieve the effect of appeasement through the automatic swing of the care car to avoid affecting other babies due to crying. In the future, in addition to the heartbeat value, blood oxygen saturation (SaO_2), and body temperature, a database can be built for data collection. Through a large amount of data analysis, the baby's health can be more accurately determined. Alternatively, reminders for feeding time or giving medicines can be set to facilitate health care workers with more efficient care.

References

[1] dsPIC30F4011, [online] Available: https://www.microchip.com/wwwproducts/en/dsPIC30F4011

[2] G. Zonios, "Pulse oximetry theory and calibration for low saturations," IEEE Trans. Biomed. Eng., vol. 51, pp. 818-823, 2004.

[3] Sergio Franco, Design with Operational Amplifiers and Analog Integrated Circuits, 3rd edition,McGraw-Hill College, 2001.

[4] K. Nakajima, T. Tamura and H. Miike, "Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique", Med. Eng. Phys., vol.18, pp.365-372, 1996.

[5] LabVIEW, [online] Available: http://taiwan.ni.com/labview-resource

[6] Yung-Hua Kao, Paul C.-P. Chao, Yueh Hung and Chin-Long Wey, "A New Reflective PPG LED-PD Sensor Module for Cuffless Blood Pressure Measurement at Wrist Artery," IEEE Sensors, 2017.

An All-Digital Clock Generator with Modified Dynamic Frequency Counting Loop and LFSR Dithering

Pao-Lung Chen,
National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

***Abstract*—This work presents a modified dynamic frequency counting loop and LFSR dithering for all-digital clock generator. In contrast to fixed clock cycles in conventional design, the modified dynamic frequency counting loop dependents on the states with variable clock cycles. We also applied the LFSR fractional dithering technique to enhance the resolution of multiphase digitally controlled oscillator. A test chip for the proposed all-digital clock generator was fabricated in a standard 0.18µm CMOS technology, and the core area was 0.112mm². The output frequency had a range of 113MHz~360MHz at 1.8V with RMS jitter 28ps at 359.78MHz/1.8V.**

I. INTRODUCTION

Many applications, such as video graphics card and telecommunication system, require a clock generator. Quartz oscillators frequently require conversion when operating at low frequency. Several methods exist for realizing clock generator: analog phase locked loop (PLL) [1] and all-digital phase locked loop (ADPLL) [2]. Both PLL and ADPLL compare the output clock and input reference clock at every rising cycle as shown in Fig. 1 (a). Our proposed method is the modified dynamic frequency counting loop (DFC) with LFSR dithering. It is a first-order system that uses a variable time period to estimate and tune the frequency of the DCO as shown in Fig. 1 (b).

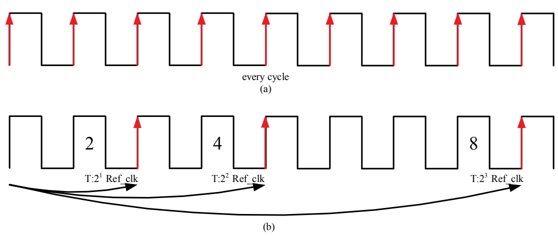


Fig. 1. Sampling period of conventional method and proposed method.

Our proposed LFSR dithering method is pure digital and easy to synthesis for different process technologies. In addition, the proposed method is low cost to enhance DCO resolution.

II. PROPOSED FREQUENCY MULTIPLIER WITH MODIFIED DFC AND LFSR DITHERING

The proposed all-digital clock generator with modified DFC control uint and linear feedback shift register (LFSR) dithering is shown in Fig. 2. The first stage is the DFC control unit design. In order to increase locking speed, a succession approximation register (SAR) is used to replace two ripple counters in [3]. The output stage is digitally controlled oscillator (DCO) with eight phase output.

**This work was supported by Ministry of Science and Technology, MOST 104-2221-E-327-032.*

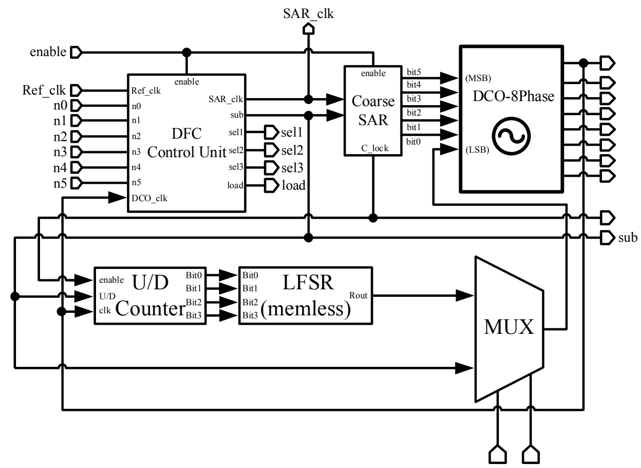


Fig. 2. Proposed all-digital clock generator with modified DFC Control Unit and LFSR dithering.

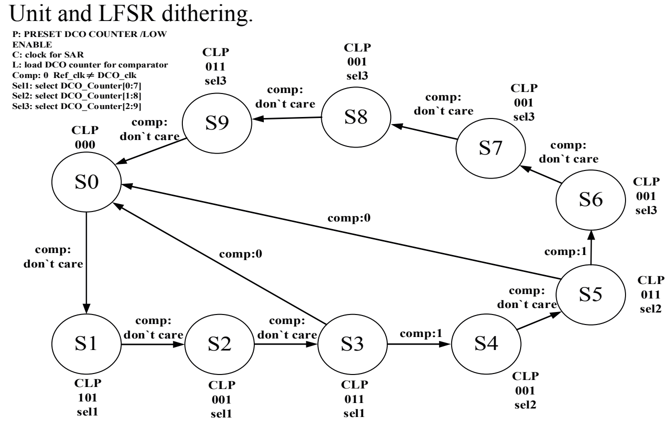


Fig. 3. State transition of modified DFC control unit..

The LFSR dithering is to enhance the resolution of digitally controlled oscillator (DCO). Figure 3 shows the state transition of DFC control unit. It consists of ten states (S0 ~ S9). The S0 is the initial state. If the signal of "comp" is 0, then the state machine will reset to the initial state. In addition, if the signal of "comp" is 1 for all states (S1 ~ S9), the state machine will also reset to the initial state. The modified DFC control unit with counter and comparator is shown in Fig. 4.

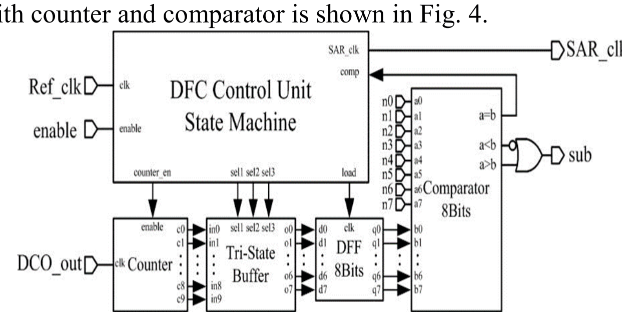


Fig. 4. Modified DFC control unit and counter for loop control.

III. IMPLEMENTATION OF PROPOSED LFSR DITHERING

A modification of LFSR has be made to ensure that even numbers of "1" and "0". A detection circuit for "0" must be added in a linear feedback shift register. The characteristic function of eight stages LFSR is given by Eq. (1)

$$P_8(X)= X_8 + X_6 + X_5 + X_4 + 1 \qquad (1).$$

An attempt is also made to modify the $P_8(x)$ with all possible states (2^8 =256). The purpose of NOR gates as shown in Fig. 5 is to insert the "all zero" state in between the "00…01" and the "10…00" states. The NOR gate in the feedback shift register ensures that modified LFSR has 256 rather than 255 states. The characteristic function of Fig. 5 is given by

$$P_{modified}(X)= P_8 \oplus (X_8+X_7+X_6+X_5+X_4+X_3+X_2)' \qquad (2).$$

The output sequence Q_1 in Fig. 5 is also called a De Bruijn sequence [4].

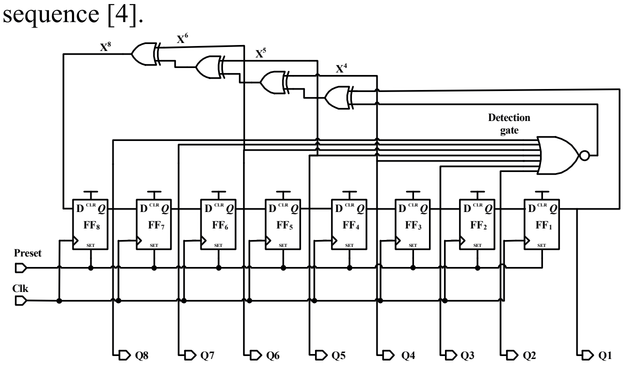


Fig. 5. Modified LFSR dithering.

IV. MEASUREMENTS OF PROPOSED ALL-DIGITAL CLOCK GENERATOR

The test chip is fabricated in the TSMC 0.18μm 1P6M CMOS process. The chip microphotograph is shown in Fig. 6. The core size is 363 x 308 μm^2. The measurement of control signals of DFC control unit is shown in Fig. 7. The upper part is chip measurement and bottom part is post layout simulation. The Load signal indicates that the sampling time is variable. The measurement of output clock at 359.78 MHz when input reference is 1MHz and multiplication factor is 360 is shown in Fig. 8. The spurious-free dynamic range (SFDR) is improved by 12.3 dB as compared with conventional method.

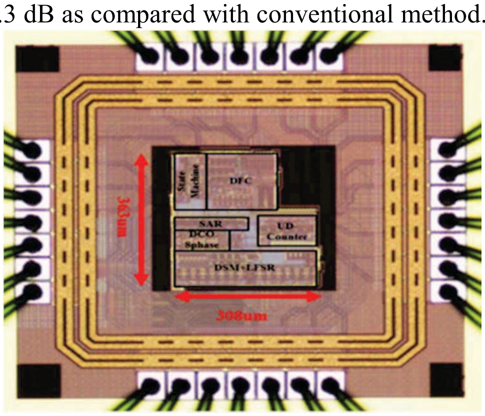


Fig. 6. Micophotograph of proposed all-digital clock generator.

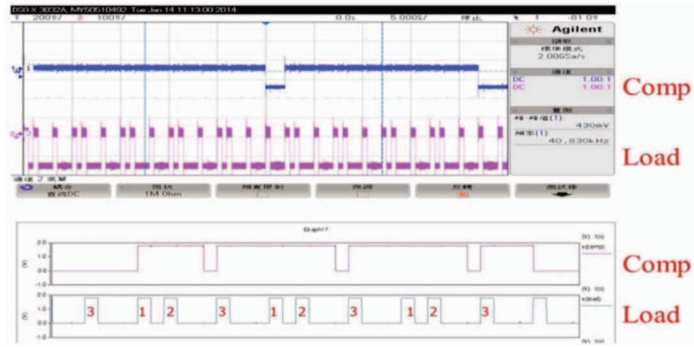


Fig. 7. Measured signals of Comp and Load.

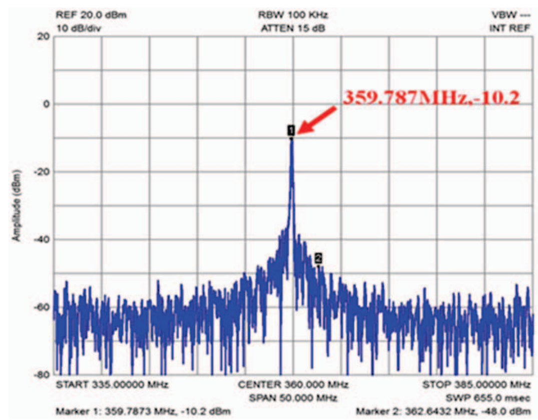


Fig. 8. Output clock measured at 359.78 MHz.

V. CONCLUSION

We propose an all-digital clock generator with modified dynamic frequency counting loop and LFSR dithering. A test chip was fabricated in was fabricated in a standard 0.18μm CMOS technology, and the core area was $0.112mm^2$. The measured output at 359.78 MHz with RMS jitter 28 ps. The spurious-free dynamic range (SFDR) is improved by 12.3 dB as compared with conventional method.

ACKNOWLEDGMENT

The authors would like to thank their colleagues in the VLSI Laboratory of National Kaohsiung University of Science and Technology for the many fruitful discussions.

REFERENCES

[1] J. G. Maneatis, J. Kim, I. McClatchie, J. Maxey and M. Shankarads, "Self-Biased High-Bandwidth Low-Jitter 1-to-4096 Multiplier Clock Generator PLL," IEEE J. Solid-State Circuits, Vol. 38, pp. 1795-1803, Nov. 2003.

[2] Ching-Che Chung, Wei-Siang Su, and Chi-Kuang Lo, "A 0.52/1 V Fast Lock-in ADPLL for Supporting Dynamic Voltage and Frequency Scaling , " IEEE Trans. on VLSI, Vol. 24, No. 1, pp. 408-412, Jan. 201

[3] Pao-Lung Chen, Ching-Che Chung, Jyh-Neng Yang, and Chen-Yi Lee, "A Clock Generator with Cascaded Dynamic Frequency Counting Loops for Wide Multiplication Range Applications," IEEE Journal of Solid-State Circuits, vol. 41, no. 6, pp. 1275-1285, Jun. 2006.

[4] P.-L. Chen, Da-Chen Lee and Wei-Chia Li "Flying-Adder Frequency Synthesizer with a Novel Counter-Based Randomization Method," IEICE Transactions on Electronics, Vol. E98-C, No.6, pp. 480-489, June 2015.

Performance Evaluation of Heterogeneous Cluster for Satellite Data Processing

Sethakarn Prongnuch* Suchada Sitjongsataporn† Theerayod Wiangtong‡
*The Electrical Engineering Graduate School, †Dept. of Electronic Engineering, Faculty of Engineering
Mahanakorn University of Technology, 140 Cheumsamphan Rd., Nongchok, Bangkok 10530 Thailand
E-mail: sethakarn.pr@ieee.org*and ssuchada@mut.ac.th†
‡Dept. of Electrical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang
1 Chalongkrung Rd., Ladkrabang, Bangkok 10520 Thailand, Email: theerayod.wi@kmitl.ac.th‡

***Abstract*—This paper presents the performance evaluation of heterogeneous cluster for satellite data processing. The proposed system consists of a master and eight slaves of the heterogeneous system connected to a start-network switch. The metric performance is used for satellite data processing algorithm. For the performance evaluation, the overall processing time is 102.85 seconds. Additionally, comparisons with the related work show better than 1.6 times of processing time compared with the previous work.**

I. Introduction

The heterogeneous cluster aims to address these issues by incorporating different types of processors and/or specific hardware in a single system [1]. In each cluster is the heterogeneous system more than one connected through a network system. The generally heterogeneous system architecture in the heterogeneous cluster as shown in Fig. 1 includes general-propose processor "GPP", coprocessor "COP", hardware accelerators "HW-ACC", local memory "Local MEM", and shared memory "SH. MEM". Some advantages of the heterogeneous cluster are as high-performance computing, energy efficiency, and scalability.

The principle design of heterogeneous system architecture (HSA) in [2] based on hardware-software computer architecture design is used. There are four layers including 1) an application layer is a software or program used for data processing that providing for users, 2) the application programming interface (API) layer is the software management between an application layer with the operating system layer such as OpenCL, OpenMP, MPI, etc., 3) the operating system (OS) layer is the software operating between an API layer with the hardware layer as known as a Linux open-source operating system, and 4) the hardware layer is the devices in the commercial market such as an embedded system board and the single board computer.

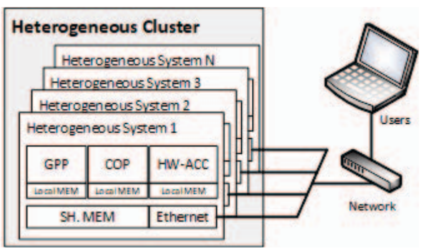


Fig. 1. The generally heterogeneous cluster.

The satellite data or satellite imagery is information about the earth in the space, gathered by man-made satellites in their orbits. The most common use of satellite data processing is earth observation, satellites deliver information about the surface and weather changes on the planet earth. An example of satellite data processing in the Thailand Meteorological Department (TMD) [3] that is used to supply weather forecasts for the entire country and publicize disaster warnings to fulfill the requirement from administration and management in natural disaster mitigation.

II. Proposed Heterogenous Cluster

The proposed heterogeneous cluster architecture design based on [6] composes the master heterogeneous system "MA-HS" and 8-slave heterogeneous system "SL-HS" are connected to a network switch which is star topology as shown in Fig.2(a). There are four layers described below.

A. Hardware Layer & OS Layer

The hardware layer employs the Microserver Parallella board [4] as shown in Fig. 2(b). This device has two different types of processors between a ZYNQ SoC (ARM and FPGA are inside) processor, which is the ARM processor called "a host processor" and a 16-cores Epiphany RISC coprocessor called "a coprocessor". A round of processors includes 1GB memory, Gigabit Ethernet, and 32GB μSD storage. The Parabuntu from Adapteva Inc. is an operating system.

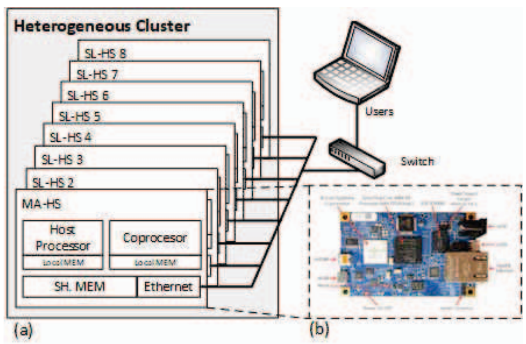


Fig. 2. The proposed heterogeneous cluster.

B. API Layer & Application Layer

APIs use COPRTHR SDK and Epiphany SDK that provides a collection of the library (e.g., OpenCL, MPI, Epiphany$-$C/C++, etc.) for implementing in the host processor, and, a coprocessor. The application using to evaluate the performance our heterogeneous cluster is the satellite data processing algorithm in the next section.

III. Satellite Data Processing Algorithm

In this paper, the satellite data from the TMD using to quantitative precipitation estimation rainfall over Thailand by Himawari-8 satellite observation every 10 minutes by cloud service.

The rainfall rate RR [5] in mm/hr. is the relation between rain and temperature, which is defined as:

$$RR = \alpha \times 10^{\beta \times temp_k^{\gamma}} , \quad (1)$$

where α, β and γ are the coefficients which depend on different climatic and locations. $temp_k$ is temperature in Kelvin from satellite data.

The normalization of temperature N_{temp_c} is defined as:

$$N_{temp_c} = \begin{cases} \text{zero} & ; \ temp_k - \delta > \epsilon \\ temp_k^5 & ; \ temp_k - \delta < \epsilon \end{cases} \quad (2)$$

where $\delta = 273$, ϵ depends on color bar monitoring.

The proposed satellite data processing (SDP) algorithm is introduced in Algorithm 1. Line 1 shows an initial satellite data and input the network common data form as satellite data called ".nc file" to host processor; Line 2 shows the conversion between .nc file to the common data form language called ".cdl file" and send to the coprocessor; Line 3 shows searching process for finding $temp_k$ in .cdl file; Line 4-6 show the decision of rain rate and send back to host processor when finished; Line 7 shows the conversion process .cdl file back to the .nc file; and Line 8 shows the results of data processing which is .nc file.

Algorithm 1 Satellite data processing (SDP) algorithm

Host processor
1: **Input** satellite data .nc file
2: Convert .nc file to .cdl file
Coprocessor
3: Search $temp_k$ in .cdl file
4: **If** N_{temp_c} = zero **do** $RR \Longleftarrow N_{temp_c}$
5: **else** $RR \Longleftarrow \alpha \times 10^{\beta \times temp_k^{\gamma}}$
6: **end if**
Host processor
7: Convert .cdl file to .nc file
8: Output satellite data processing to .nc file
9: End

IV. The Experimental Results

An experiment of the heterogeneous cluster proposed and performance are measured by SDP algorithm. An algorithm proposing is implemented programming by Epiphany$-$C/C++, complied by E-GCC/G++ on a heterogeneous cluster. Fig. 3 shows the output from SDP algorithm satellite imagery over Thailand on 10/09/2016, where

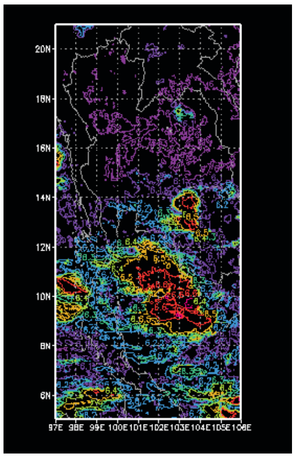


Fig. 3. The satellite imagery on 10/09/2016 processed by SDP algorithm.

TABLE I
Performance Evaluation

System	Processing time (seconds)		
	Hostprocessor	Coprocessor	Overall
The proposed system	95.13	7.72	102.85
The system in [6]	164.16	-	164.16

$\alpha = 1.1183$, $\beta = -0.036382$, $\gamma = 0.5$, and $\epsilon = -25$ which is the high temperature displayed in red tone color and low temperature shown in blue tone color. While, shade of color is translated with (1) followed the GrADS default rainbow sequence. The performance evaluation results are shown in Table I, and compared with [6]. The proposed system shows the processing time of the host processor. The overall processing time is better than 1.73 times and 1.60 times of processing time compared with previous work in [6].

V. Conclusions

In this paper, the performance evaluation of heterogeneous cluster is proposed for SDP algorithm. Proposed system consists of a master and 8-slave of the heterogeneous system are connected to a star-network switch. For the performance evaluation results, the overall processing time is 102.85 seconds. Additionally, comparisons with related work show better than 1.60 times of processing time compared with previous work.

References

[1] S. L.Harris and D. M. Harris, "Computer Architecture ARM Edition", San Francisco, CA, USA: Morgan Kaufmann, 2016.
[2] S. Prongnuch and T. Wiangtong, "Heterogeneous Computing Platform for data processing", *IEEE ISPACS*, Phuket, pp. 1-4, 2016.
[3] Thailand Meteorological Department., "Vision and Missions", [Online]. Available: https://www.tmd.go.th/en/aboutus/vision.php [Accessed: 8-Aug-2019].
[4] Adapteva Inc.,"Parallella-1.x Reference Manual", [Online]. Available: https://www.parallella.org/docs/parallella-manual.pdf [Accessed: 8-Aug-2019].
[5] K. P.N. Sakolnakhon and S. Nuntakamolwaree, "The Estimation Rainfall using Infrared (IR) band of Himawari-8 Satellite over Thailand", *Engineering: Naresuan University*, Thailand, Vol. 39, pp. 236-248, 2016.
[6] S. Prongnuch, J. Omkhet, T. Phikulthong and P. Malapim, "The Mobile High-performance Computing Design for Satellite Data Processin", *The 4th RSU National Research Conference on Science and Technology, Social Sciences and Humanities*, Rangsit University, pp. 298-307, 2019.

Softsign Function Hardware Implementation Using Piecewise Linear Approximation

Chih-Hsiang Chang, En-Hui Zhang, Shih-Hsu Huang
Department of Electronic Engineering
Chung-Yuan Christian University
Taoyuan, Taiwan
{g10676034, g10876003, shhuang}@cycu.edu.tw

Abstract—**Softsign is a widely used activation function in recurrent neural networks. However, no special attention has been paid to the hardware implementation of Softsign function. In this paper, we propose a hardware architecture, which is based on the piecewise linear (PWL) approximation, for the hardware implementation of Softsign function. The main advantage of the proposed hardware architecture is that we develop a low-power high-accuracy approximate multiplier for the PWL approximation. Implementation results show that the proposed hardware architecture can save 55.1% power consumption with only 4.2% increase in mean absolute error.**
Keywords—Activation Function, Hardware Approximation, Hardware Architecture, Approximate Multiplier, Low Power.

I. INTRODUCTION

Neural networks (NNs) have been successfully applied to many application fields. In NNs, nonlinear factors are added by activation functions, making them easier to solve complex problems. However, activation functions [1-6] have multiplications, divisions, and exponential mathematical operations, which consume a lot of computing resources on hardware. It is necessary to adopt approximation methods to implement activation functions in hardware for low power.

Widely-used activation functions include ReLU, Sigmoid, Hyperbolic Tangent and Softsign functions. Sigmoid, Hyperbolic Tangent and Softsign functions are widely used in recurrent neural networks (RNNs). Several research efforts [1-6] have been paid to the hardware approximation of Sigmoid and Hyperbolic Tangent functions. However, to the best of our knowledge, no special attention has been devoted to the hardware approximation of Softsign function.

In this paper, we propose a hardware architecture for the Softsign function. The proposed hardware architecture is based on the concept of piecewise linear (PWL) approximation method. If compared with the hardware architecture that uses a conventional multiplier (for the PWL approximation), the proposed hardware architecture can save 55.1% power consumption with only 4.2% mean absolute error.

II. PRIOR WORKS

Softsign function is defined as follows: $f(x) = \frac{x}{1+|x|}$

In the proposed hardware architecture, we use PWL approximation method to implement Softsign function. In PWL approximation, a nonlinear function is split into many linear segments. Each linear segment can be represented by a linear function ax+b, where a and b are two constant parameters for this linear function. Note that different linear segments have different constant parameters.

The data format of the proposed hardware architecture is below. We use 16 bits to represent each data. The values of constant parameters (i.e., the value of a and the value of b in a linear function) are within the range (0,1). Thus, for each constant parameter, we can use 16 bits to represent its fraction part. The input value (i.e., the value of x) is within the range (-8,8). Thus, we need to use 4 bits to represent the integer part of an input value. As a result, for the input value, we can use 12 bits to represent its fraction part. Fig. 1 displays the data format of a PWL multiplication. Since the multiplication result (i.e., the value of ax) is also represented by 16 bits, its integer part has 4 bits and its fraction part has 12 bits.

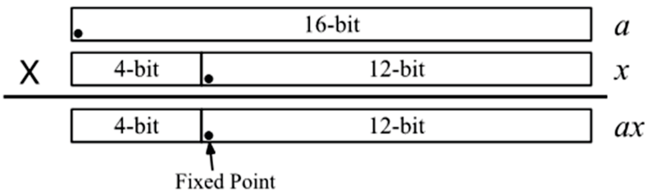


Fig. 1. The data format of 16-bit PWL multiplication.

A conventional 16-bit × 16-bit multiplier produces a 32-bit result. However, as shown in Fig. 1, the multiplication result is represented by the 16-bit data format. Intuitively, we can store the multiplication result in the 16-bit format by just ignoring the least significant 16 bits of the 32-bit result. The drawback of this intuitive method is that it consumes both large circuit area and large power consumption. Thus, we are motivated to develop a 16-bit × 16-bit approximate multiplier for directly producing a 16-bit result with low power, small area and high accuracy.

The specification for utilizing the proposed approximate multiplier architecture is as below. Both the two inputs are n bits, where n can be any positive integer number. The output (i.e., the multiplication result) should also be n bits. Suppose that the integer part of one input has p bits and the integer part of the other input has q bits. Then, obviously, the integer part of the output (i.e., the multiplication result) is p+q bits. The proposed approximate multiplier architecture can work if and only if p+q ≤ n (i.e., the number of bits of the integer part of the multiplication result should be less than or equal to the total number of bits of the multiplication result).

In the proposed Softsign function hardware architecture, we adopt 16-bit × 16-bit approximate multiplier to produce a 16-bit result as shown in Fig. 1. In fact, the algorithm of the proposed approximate multiplier architecture can be easily generalized to n-bit × n-bit, where n can be any positive integer. In the following, for the convenience of explanation, we use 4-bit × 4-bit approximate multiplier (to produce 4-bit result) as an example for illustration. We assume both the two inputs of this 4-bit × 4-bit approximate multiplier are within the range

(0,1). Thus, for each input, we can use 4 bits to represent its fraction part (i.e., 0 bit to represent its integer part).

Take the following two inputs for example: In1 $= \frac{9}{16} = 0.1001_2$ and In2 $= \frac{6}{16} = 0.0110_2$. We may also say In1[3] = 1, In1[2] = 0, In1[1] = 0, In1[0] = 1, In2[3] = 0, In2[2] = 1, In2[1] = 1 and In2[0] = 0. Here we use the notation $>> k$ to denote shift right k bits. The approximation algorithm is below. First, we calculate four partial products: R_3, R_2, R_1 and R_0 (note that each partial product is a 4-bit value).

$$R\_3=In2[3]\cdot(In1\gg1+In1[0])=0\cdot(0.0100+0.0001)$$
$$R\_2=In2[2]\cdot(In1\gg2+In1[1])=1\cdot(0.0010+0.0000)$$
$$R\_1=In2[1]\cdot(In1\gg3+In1[2])=1\cdot(0.0001+0.0000)$$
$$R\_0=In2[0]\cdot In1[3]=0\cdot 0.0000$$

Next, we add these four partial products together. We can derive that the multiplication result is $\frac{3}{16}$.

$$R_3+R_2+R_1+R_0=0.0000+0.0010+0.0001+0.0000=0.0011=\frac{3}{16}$$

Based on the above discussions, we find that conventional MBFF can reduce more power but LC-MBFF has a larger solution space for optimization. Therefore, we are motivated to develop a two-stage design flow to minimize the power consumption under timing constraints.

In the proposed Softsign function hardware architecture, we generalize the approximate multiplication algorithm to implement 16-bit × 16-bit multiplier. The overall Softsign function hardware architecture is displayed in Fig. 2. Because Softsign function is an odd function and the proposed approximate multiplier is an unsigned multiplier, in PWL hardware approximation we only deal with the absolute value part. We implement a selective sign-converter, in which the sign-bit of the input is used as the control. If the input value if negative, before PWL approximation, we translate it to its absolute value. Then, after PWL approximation, we need to translate the result to a negative value if the input is a negative value. We use ROM-A and ROM-B to store constant parameters of all the linear functions for PWL approximation. According to the input value, we use multiplexers (MUX) to select constant parameters. Then, multiplying a by x and plus b is the result, where the multiplication is performed by the proposed approximate multiplier.

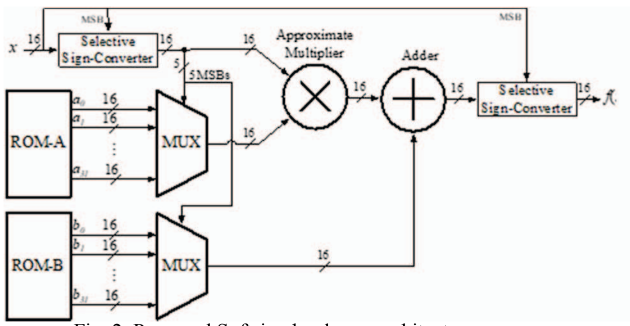


Fig. 2. Proposed Softsign hardware architecture.

Lastly, we discuss how to determine constant parameters for PWL approximation. Here we use the algorithm proposed in [5] to determine the best solution. We put all possible values into the algorithm [5] for exhaustive simulation (exhaustive analysis) and then find the best solution.

III. Proposed Approach

The proposed Softsign function hardware architecture has been implemented by using TSMC 40 nm cell library. We use the PWL approximate with 32 linear segments. The clock frequency is set to be 500 MHz. In addition to use the proposed 16-bit × 16-bit approximate multiplier, we also implement a conventional 16-bit × 16-bit multiplier (with ignoring the least significant 16 bits of the 32-bit result) for comparison.

Table I tabulates the comparisons on circuit area and power consumption. Compared with the hardware using a conventional multiplier, our approach can reduce 30.2% circuit area and 55.1% power consumption. Table II tabulates the comparisons on the maximum error and the average error (mean absolute error). As shown in Table II, the increase of error caused by our approach is small. Our increase in the maximum error is only 4.3%, and our increase in the average error is only 4.2%.

Table I. Comparisons on area and power.

	Area (μm^2)	Power (μW)
Conventional	1939.3688	479.7
Ours	1353.5424	215.2

Table II. Comparisons on maximum error and average error.

	Maximum Error	Average Error
Conventional	0.005615	0.0003675
Ours	0.005859	0.0003830

IV. Conclusions

In this paper, we use PWL approximation to implement Softsign function in hardware. The main advantage of the proposed hardware architecture is that we develop a low-power high-accuracy approximate multiplier for performing the PWL approximation. Implementation results show that the proposed approach can save 30.2% circuit area and reduce 55.1% power consumption with 4.3% increase on the maximum error and 4.2% increase on the average error.

V. Acknowledgements

This work was supported in part by Ministry of Science and Technology, Taiwan, under grant number MOST 108-2218-E-033-002.

VI. Reference

[1] K. Leboeuf, et al., “High speed VLSI Implementation of The Hyperbolic Tangent Sigmoid Function”, Proc. of ICCIT, 2008.

[2] P.K. Meher, “An Optimized Lookup-Table for The Evaluation of Sigmoid Function for Artificial Neural Networks”, Proc. of IEEE VLSI-SOC, 2010.

[3] A. Armato, et al., “Low-Error Digital Hardware Implementation of Artificial Neuron Activation Functions And Their Derivative”, Microprocessors and Microsystems, 2011.

[4] B. Zamanlooy and M. Mirhassani, “Efficient VLSI Implementation of Neural Networks with Hyperbolic Tangent Activation Function”, IEEE Trans. on VLSI Systems, , 2014.

[5] V. T. Nguyen, et al., “An Efficient Hardware Implementation of Activation Functions Using Stochastic Computing for Deep Neural Networks”, Proc. of IEEE MCSoC, 2018.

[6] C.H. Chang, et al., “Hardware Implementation for Multiple Activation Functions”, Proc. of IEEE ICCE-TW, 2019.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Evaluation of Wireless Body Area Network Utilizing Super Orthogonal Convolutional Code

Kento Takabayashi
Department of Information and Communication Engineering
Okayama Prefectural University
Soja, Japan
takabayashi.kento.xp@gmail.com,
https://orcid.org/0000-0001-5189-9253

Hirokazu Tanaka
Graduate School of Information Sciences
Hiroshima City University
Hiroshima, Japan
hi.tanaka@m.ieice.org

Katsumi Sakakibara
Department of Information and Communication Engineering
Okayama Prefectural University
Soja, Japan
sakaki@c.oka-pu.ac.jp,
https://orcid.org/0000-0001-5935-1794

***Abstract*—This paper provides an evaluation of IEEE 802.15.6 ultra-wideband (UWB) physical layer (PHY) based wireless body area network (WBAN) using a super orthogonal convolutional code (SOCC) as an error correcting code. Numerical results show that a low coding rate SOCC has better payload error probability performance in a constant process gain case. In addition, small delay is achieved with no retransmission in the case of a low coding rate SOCC.**

Keywords—WBAN, IEEE 802.15.6, UWB, Error Controlling, SOCC

I. Introduction

The use of the Internet of Things (IoT) in the medical and healthcare fields has received much attention [1]. Wireless Body Area Network (WBAN) is a very important technology to realize IoT systems in this field [2,3]. WBAN consist of a collection of low-power, miniaturized, invasive or non-invasive lightweight sensors with wireless communication capabilities that operate near the human body. IEEE 802.15.6 is known as one of WBAN's international standards for medical and healthcare applications [4]. The IEEE 802.15.6 defines three physical layers (PHY): narrowband (NB), ultra-wideband (UWB), and human body communications (HBC). This research focuses on a UWB PHY which offers high data rate transmission, low energy consumption, high multi-pass resolution, good coexistence with other wireless communication systems, and so on [5,6].

This paper proposes to apply an error correcting code called super orthogonal convolutional code (SOCC) to WBAN using UWB PHY of IEEE 802.15.6. SOCC can design very low coding rate error correction codes, and it has very high error correction capability thanks to its very low coding rate [7,8]. Thus, the combination of SOCC and UWB is very compatible. However, the use of SOCC is not assumed in the UWB PHY defined in IEEE 802.15.6. This research theoretically evaluates the performance when a coding rate of SOCC is changed under the condition of constant process gain.

II. UWB PHY

The UWB PHY frame format is formed by the synchronization header (SHR), the physical layer header (PHR), and the physical layer service data unit. (PSDU), respectively, as shown in Fig. 1 [4]. This research mainly focuses on the PHR and the payload (PSDU). The UWB symbol structure is shown in Fig. 1. The UWB symbol can be transmitted using single pulse or multiple pulses as shown in Fig. 1. In the on-off keying (OOK) combined with 2-ary waveform coding as modulation and an energy detection (ED) receiver case, signal-to-noise ratio (SNR_{DV}) at the decision variable of ED receiver is calculated as follows [6]:

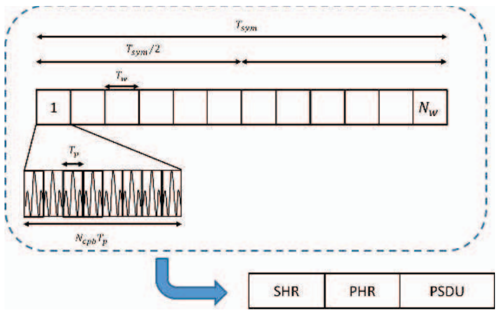


Figure 1. UWB PHY frame format and symbol structure.

$$SNR_{DV} = \frac{N_{cpb}\frac{E_s}{N_0}}{2 + TW\frac{N_0}{E_s}}. \tag{1}$$

Here, N_{cpb} is the number of pulses per one bit, T is the integration time per pulse, W is the signal bandwidth, E_s is the signal energy per pulse and N_0 is the energy of thermal noise. Hence, bit error probability (BEP) can be calculated by using (1) and the $Q(\cdot)$ function under the additive white Gaussian noise (AWGN) channel as follows [6]:

$$BEP = Q\left(\sqrt{SNR_{DV}}\right). \tag{2}$$

III. Super Othogonal Convolutional Code

SOCC is a kind of orthogonal convolutional code of very low coding rate [7,8]. The encoder of SOCC is illustrated in Fig. 2. The encoder is composed of a shift register of length K (constraint length) and a block orthogonal (Walsh-Hadamard) encoder. In the K bit shift register, the inner (K-2) bit memories are mapped into one Walsh sequence of length 2^{K-2}, while the two outer bits are added to each binary number of the selected Walsh sequence by exclusive OR. The coding rate is given by $R_c = \frac{1}{2^{K-2}}$. Then, the decoder uses the Viterbi algorithm. The BEP when encoded by SOCC is as follows [8]:

$$BEP_{SOCC} \leq \sum_{d=d_{free}}^{\infty} c_d P_d \tag{3}$$

$$P_d = \sum_{t=0}^{d} \binom{d-1-t}{t} (1-BEP)^t (BEP)^d. \tag{4}$$

Here, c_d is the total number of information bit errors produced by incorrect Hamming weight paths, and d_{free} is free distance of the code. These in the SOCC are referred to in [7].

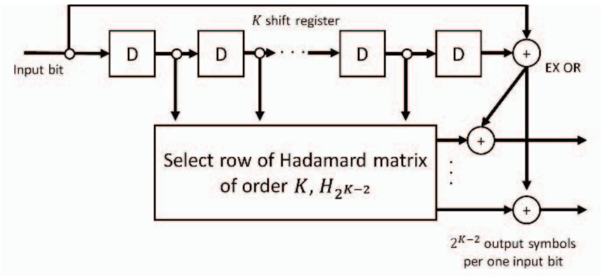


Figure 2. SOCC encoder.

IV. Numerical Results

The PSDU error probability (PSDUEP) and the delay time (T_{delay}) are evaluated when the process gain is fixed as $G = N_{cpb}/R_c = 32$. Each evaluation index is calculated as follows:

$$\mathrm{PSDUEP} = 1 - (1 - BEP_{SOCC})^{L_{PSDU}} \tag{5}$$

$$T_{delay} = T_{slot}\overline{N_{tr}} \tag{6}$$

$$PEP = 1 - (1 - BEP_{SOCC})^{L_{PHR}}(1 - BEP_{SOCC})^{L_{PSDU}} \tag{7}$$

$$\overline{N_{tr}} = \sum_{i=1}^{q-1} iPEP^{i-1}(1 - PEP) + qPEP^{q-1}\,. \tag{8}$$

Here, L_{PHR} and L_{PSDU} are the length of PHR and PSDU respectively, T_{slot} is a slot duration in the medium access control (MAC) layer, $\overline{N_{tr}}$ and q are the numbers of average and the maximum transmissions respectively. Figs. 3 and 4 are numerical results of PSDUEP and T_{delay}. The SOCC has high PSDUEP performances as compared with the cases where the error correction code is not used, and the BCH code of IEEE 802.15.6 is used. It can also be seen that lower R_c has better PSDUEP performance. In addition, the SOCC with lower R_c achieves less than T_{delay} = 10 [ms] with no retransmission under low E_s/N_0 conditions.

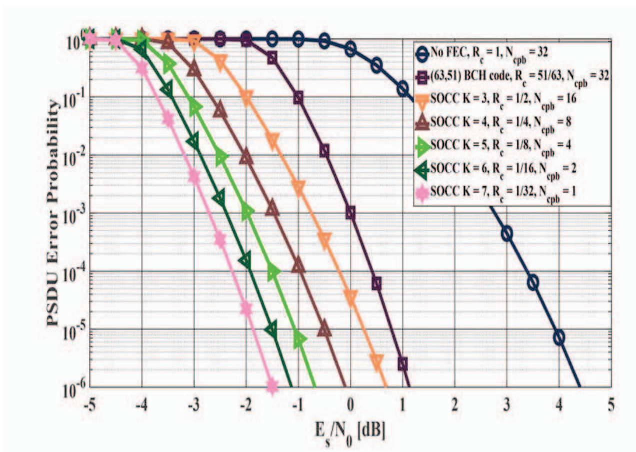


Figure 3. PSDUEP performance as a function of E_s/N_0 in the case $G = N_{cpb}/R_c = 32$, $L_{PHR} = 40$ bits and $L_{PSDU} = 255$ octets.

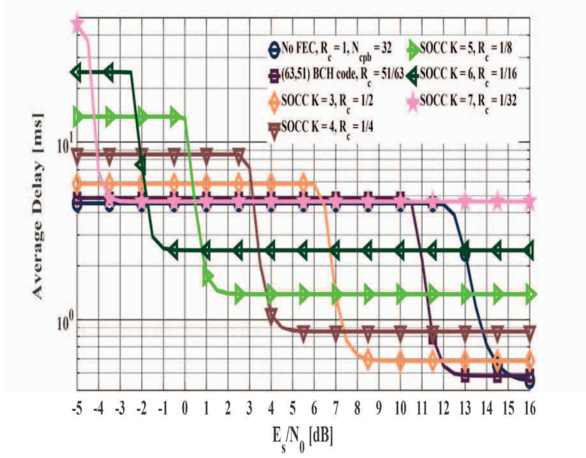


Figure 4. T_{delay} performance as a function of E_s/N_0 in the case $N_{cpb} = 1$, $L_{PHR} = 40$ bits and $L_{PSDU} = 255$ octets.

V. Conclusion

This paper provided the evaluation of PEP and T_{delay} when SOCC was used in UWB PHY of IEEE 802.15.6. Numerical results showed the effectiveness of SOCC. As for the future work, energy efficiency or other process gain cases should be evaluated. Cross-layer optimization that assumes an IEEE 802.15.6 MAC layer also needs to be performed.

References

[1] C. R. Costa, L. E. Anido-Rifón and M. J. Fernández-Iglesias, "An Open Architecture to Support Social and Health Services in a Smart TV Environment," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 2, pp. 549-560, March 2017.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] R. Cavallari, F. Martelli, R. Rosini, C. Buratti and R. Verdone, "A Survey on Wireless Body Area Networks: Technologies and Design Challenges," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1635-1657, Third Quarter 2014.

[3] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith and A. Jamalipour, "Wireless Body Area Networks: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1658-1686, Third Quarter 2014.

[4] Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs) used in or 12 around a body, IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks- Specific requirements: Part 15.6, 2012.

[5] I. Dotlic and R. Kohno, "Low Complexity Chirp Pulsed Ultra-Wideband System with Near-Optimum Multipath Performance," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 208-218, January 2011.

[6] H. Karvonen, J. Iinatti, M. Hämäläinen, "A cross-layer energy efficiency optimization model for WBAN using IR-UWB transceivers," *Telecommunication Systems*, vol.58, no.2, pp. 165-177, Feb. 2015.

[7] R. McEliece, S. Dolinar, F. Pollara, and H. Van Tilborg, "Some easily analyzable convolutional codes," *TDA Progress Report 42-99*, Jet Propulsion Laboratory, Pasadena, Calif. , pp. 105-114, Nov. 1989.

[8] T. Matsumoto, R. Kohno, "Performance of Super-Orthogonal Convolutional Coding for Ultra-Wideband Systems in Multipath and Multiuser Channels" *Wireless Personal Communications*, vol.40, no.3, pp. 355-370, Feb. 2007.

Probability Distribution Analysis of Backoff Time with Frozen Backoff in CSMA/CA

Koki Samoto, Ikuo Oka, Shingo Ata
Graduate School of Engineering,
Osaka City University, 3-3-138 Sugimoto,
Sumiyoshi-ku, Osaka, 558-8585 Japan
Email{samoto@c., oka@, ata@}info.eng.osaka-cu.ac.jp

***Abstract*— Recently, in wireless communications a transmission rate degradation due to the increase of traffic becomes a serious problem. To deal with this problem, the number of users is an important parameter. To the estimation of the number of users, a probability distribution of backoff is applicable with the maximum likelihood approach. In this paper, an analysis method is proposed to derive the probability distribution of backoff considering frozen backoff in time axis. The analysis and simulation results are compared to show the validity the proposed method.**

***Index Terms*— CSMA/CA, DCF, IEEE802.11, Backoff time**

I. Introduction

Recently, with the development of wireless communications, the demand for wireless communication devices such as smartphones and tablets is increasing. Most of these devices use wireless LAN and a packet transmission is controlled by CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) as MAC (Medium Access Control) protocols.

In the wireless communications, users communicate with other users by a relay access point(AP). When access is concentrated to the same AP due to the increase of users, the communication quality will be degraded by a packet drop. The best AP selection methods in multiple AP's are proposed in [1]-[3].

To deal with the problem, the number of users is an important parameter. In [4], the estimation of number of users is based on an extended Kalman filter with measured number of collisions. In [5], the number of users is estimated by the maximum likelihood estimation using probability distribution of backoff time. However, an estimation error increases as the number of users.

II. CSMA/CA

In CSMA/CA, when a user has a data frame to transmit, it has to wait until channel becomes idle for a period of time called IFS (Inter Frame Space). Then, it uses a random number generator to set a backoff time. In IEEE802.11, backoff time follows as

$$b = r \cdot s, \tag{1}$$

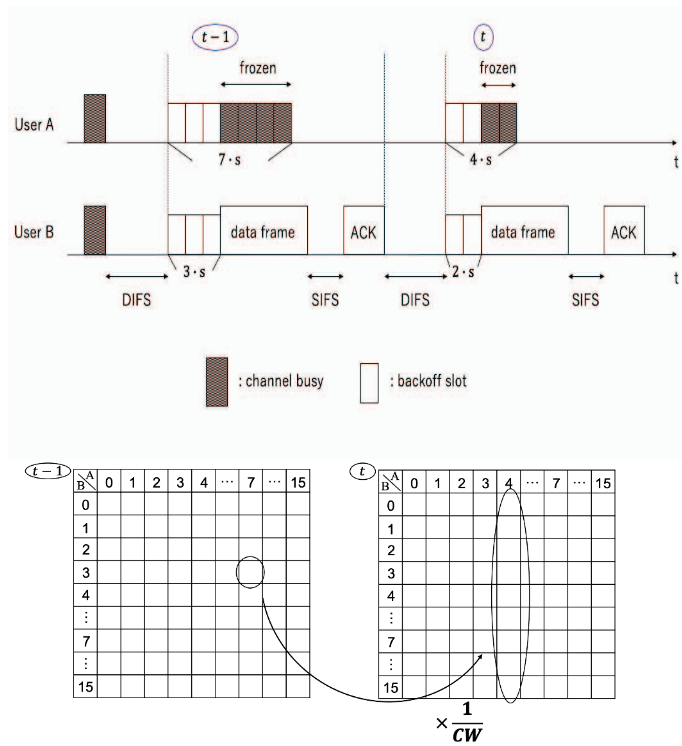


Fig. 1. Time axis analysis of $N = 2$

where r is a random integer generated from a uniform distribution of $[0, CW - 1]$, and s is a slot time. CW(Contention Window) is an integer between CW_{min} and CW_{max}, and determined by

$$CW = CW_{min} \times 2^i, \tag{2}$$

where i is a number of collision. The determined backoff time decreases for each slot time in the case of channel idle, and when it reaches 0, the user transmits a data frame. In the case of channel busy, the backoff time is frozen and carried over to the next transmission opportunity.

III. Backoff analysis

Fig.1 shows the time axis analysis when the number of users is 2. A window size is set to 15 which is employed

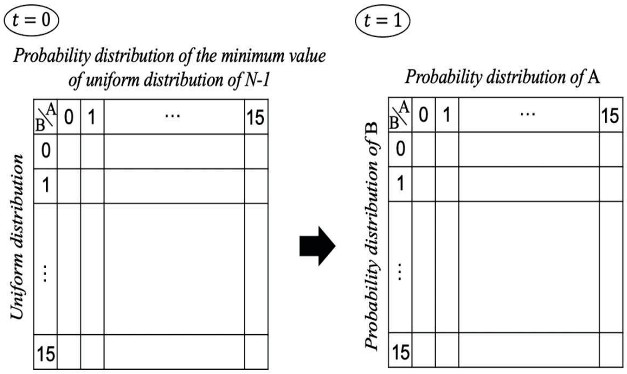


Fig. 2. Time axis analysis of $N \geq 3$

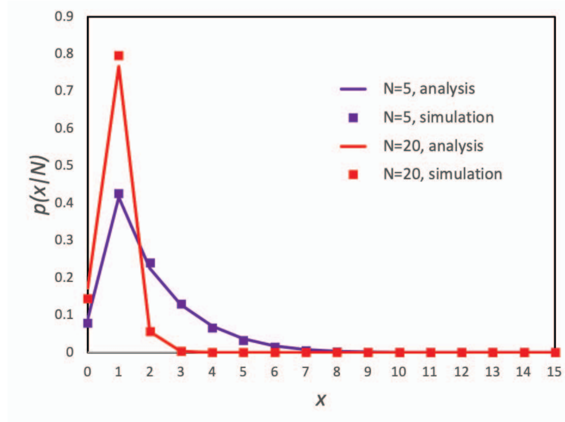


Fig. 3. The probability distribution of observed backoff

in IEEE802.11a, and the backoff stage is not considered. Users A and B are attempting to transmit a data frame to the AP. At time $t-1$, it is assumed that the backoff time's are set to $7{\cdot}s$ for user A and $3{\cdot}s$ for user B, where s denotes a slot time. At time t, user A is set to $4 \cdot s$, which has been frozen, and user B is reset to $x \cdot s(x : [0, 1, \cdots CW-1])$. Similarly, for all possible backoff slot combinations, the relation between time $t-1$ and time t is expressed by

$$P^{(t)}_{\alpha,\beta} = \begin{cases} K & (\alpha = \beta = 0) \\ K + (\sum\limits_{i=\alpha}^{CW-1} P^{(t-1)}_{i,i-\alpha}) \times \frac{1}{CW} & \\ \qquad +(\sum\limits_{j=\beta}^{CW-1} P^{(t-1)}_{j-\beta,j}) \times \frac{1}{CW} & (otherwise), \end{cases} \tag{3}$$

where $P^{(t)}_{\alpha,\beta}$ is the occurrence probability of the combination of backoff slots of users A and B at time t, and

$$K = (\sum_{i=0}^{CW-1} P^{(t-1)}_{i,i}) \times \frac{1}{CW^2}, \tag{4}$$

Based on (3), numerical calculations are executed until the steady state is obtained.

Fig.2 shows the time axis analysis when the number of users is 3 or more. Let B be the number of backoff slots in one user, and A be the minimum number of backoff slots in the other $N-1$ users. We assume the statistical independence among probability distributions. At time t, let $P_t(x \mid N)$ be the probability that the number of backoff slots in one target user becomes x when the number of user is N, and let $P'_t(x \mid N-1)$ be the probability that the minimum number of backoff slots in other $N-1$ users becomes x. Numerical calculations executed until $P_t(x \mid N) \times P'_t(x \mid N-1)$ becomes steady state, and the probability distribution of backoff time is derived.

When the number of users N is 2, the probability distribution of observed backoff is completely consistent with the simulation results. Fig.3 shows the probability distribution of observed backoff when the number of users N is 5 and 20. From Fig.3, the peak of probability is found at $x = 1$, and it is noted that a tail of the probability is a large x decreases as the number of users, N, increases. The analysis and simulation results are compared to show the validity the proposed method.

IV. Conclusion

In this paper, the analysis model has been presented to derive the probability distribution of backoff time with frozen backoff in time axis. For 2 users, the Markov state transition diagram for all possible combinations of backoff times gives the relation of distribution at time $t-1$ and time t. For 3 users or more, a time axis analysis has been extended by assuming statistical independence of probability distribution in the target user and other users. The analysis and simulation results have been compared to show the validity the proposed method.

This work was supported in parts by the Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science, No. 16KO6362.

References

[1] Fengyuan Xu, Xiaojun Zhu, Chiu C. Tan, Qun Li, Guanhua Yan, and Jie Wu, "SmartAssoc: Decentralized Access Point Selection Algorithm to Improve Throughput," *IEEE transactions on Parallel and distributed systems,* 24(12), 2482-2491

[2] Akihiro Fujiwara, Yasuhiro Sagara, and Masahiko Nakamura, "Access point selection algorithms for maximizing throughputs in wireless LAN environment," *International Conference on Parallel and Distributed Systems,* 2007.

[3] Uferah Shafi, Muhammad Zeeshan, Naveed Iqbal, Nadia Kalsoom, and Rafia Mumtaz, "An Optimal Distributed Algorithm for Best AP Selection and Load Balancing in WiFi," pp.65-69, 2001.

[4] Giuseppe Bianchi and Ilenia Tinnirello, "Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 network," *Proc. IEEE International Conference on Computer Communications 2003,* vol.1-3, pp.844-852, 2003.

[5] Norihiro Matsumoto, Ikuo Oka and Shingo Ata, "Number of Users Estimated by Statistics of Random Backoff Time in WiFi Networks," *Proc. RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing,* pp.267-270, 2018.

Phase Correction for Automatic Modulation Classification Using Iterative Closest Point

Wataru Machida, Kei Ichijo, Yosuke Sugiura, and Tetsuya Shimamura
Saitama University, Saitama, Japan
E-mail: {machida, ichijo, sugiura, shima}@sie.ics.saitama-u.ac.jp

***Abstract*—In this paper, we propose an accurate modulation classifier without the knowledge of noise variance for OFDM system. In order to estimate the amount of phase rotation caused by flat fading, we investigate to adopt iterative closest point, which is a kind of template matching technique. Combining the least-squares based phase estimation, the classification performance of the proposed method can be improved significantly.**

***Index Terms*—Modulation Classification, Least Square, Iterative Closest Point, and Channel Characteristics**

I. INTRODUCTION

Automatic modulation classification is an essential technique to detect the modulation scheme of received signals [1]. For modulation classification, most of classifiers require the prior knowledge of channel state information (CSI) such as noise variance. Unfortunately, the classification performance is seriously degraded when the CSI is unknown or estimated incorrectly [2] [3]. In this paper, we propose a new robust modulation classification method which requires no CSI.

The proposed method compensates the phase of the received signal corrupted by fading channel. The least squares (LS) method and iterative closest point (ICP) algorithm are used for estimating the amount of the phase shift.

II. PROPOSED METHOD

A. Robust Modulation Classification

The proposed method is based on [4]. The procedure of the modulation classification is as follows. At first, we use the constellation folding and then estimate one signal centroid S_M by the Automatic Constellation Grid Segmentation (ACGS) in the IQ complex plane. The centroid can be calculated from the mean of received samples r_n after folding. For BPSK, we assume that the centroid is spaced over the positive real axis. For QPSK, 16QAM and 64QAM, we also assume that the centroid is spaced in the first quadrant with common values of both the real part and the imaginary part. Next, we calculate two likelihood functions which are related with distance and phase, respectively. They are defined as

$$D(\mathbf{r}|H_{\mathcal{M}}) = \frac{1}{N}\sum_{n=1}^{N}\mathcal{I}\left\{|\hat{r}_n - S_{\mathcal{M}}| < \frac{R_{\mathcal{M}}}{I_{\mathcal{M}}}\right\}, \quad (1)$$

$$P(\mathbf{r}|H_{\mathcal{M}}) = \frac{1}{N}\sum_{n=1}^{N}\mathcal{I}\{||\tan^{-1}(\hat{r}_n)| - \tan^{-1}(S_{\mathcal{M}})| < \frac{R_{\mathcal{M}}}{I_{\mathcal{M}}}\}, \quad (2)$$

respectively, where $H_{\mathcal{M}}$ is the modulation schemes, M is the modulation order, $\mathcal{I}\{\cdot\}$ is an indicator function which returns 1 if the input is true and 0 otherwise. $\frac{R_{\mathcal{M}}}{I_{\mathcal{M}}}$ is the test radius determined by the threshold R_M. The classification result is decided according to the maximum likelihood criterion, which is defined as

$$\hat{H}_{\mathcal{M}} = \arg\max_{H_{\mathcal{M}}}(D(\mathbf{r}|H_{\mathcal{M}}) + 0.5P(\mathbf{r}|H_{\mathcal{M}})). \quad (3)$$

In actual situations, however, the performance of this method is degraded by the affect of fading. Therefore, we propose a new phase correction method that works well in flat fading environments.

B. LS Method

We firstly introduce a simple method based on the LS method. The procedure is as follows.

1) Rotate received signal points x_n by $\Delta\theta$ as

$$\begin{pmatrix} Re(\hat{x}_n) \\ Im(\hat{x}_n) \end{pmatrix} = \begin{pmatrix} \cos\Delta\theta - \sin\Delta\theta \\ \sin\Delta\theta + \cos\Delta\theta \end{pmatrix} \begin{pmatrix} Re(x_n) \\ Im(x_n) \end{pmatrix} \quad (4)$$

2) Find an optimal rotation angle $\hat{\theta}$ that minimizes the phase shift as

$$\hat{\theta} = \arg\max_{\Delta\theta}\sum_{n=1}^{N}\frac{1}{\sqrt{2}}|Im(\hat{x}_n) - Re(\hat{x}_n)|^2, \quad (5)$$

3) Correct the phase using the obtained rotation matrix.

The correction accuracy increases as the number of data points obtained increases. On the other hand, the matching performance drops significantly when the number of data points is inadequate and the error distribution does not follow the Gaussian distribution. Thus, for further performance improvement, we combine the ICP algorithm.

C. ICP Algorithm

The ICP algorithm is usually used for matching two sets of point cloud data. It sequentially calculates the amount of displacement between two point groups. We recognize signal sparce representations in communication as point cloud data so that the ICP algorithm can be applied to phase correction. It is known that the ICP algorithm can be implemented simply using singular value decomposition if the correspondence between point cloud data is known [5].

The ICP algorithm in this method is performed in the following steps.

1) Create $2 \times N$ coordinate matrices from the received signal point group and the constellation diagram, respectively, where N is the data length of the point group.
2) Calculate W by multiplying those matrices and decompose W by singular value decomposition.
3) Multiplying the left-singular matrix U and the right-singular matrix V, obtain a rotation matrix $R = UV^T$.
4) After correcting the point group of the received signal using R, calculate the norm error of each point between the point groups and repeat steps 1-4 until the error converges.

In order to cope with the problem that the accuracy depends on the initial correspondence between point groups, which is a drawback of the ICP algorithm, we use the phase correction result of the LS method as a good initial value.

III. Experiments And Results

We verify the phase correction performance of two methods; LS method and LS method + ICP algorithm, in a flat fading environment. Figures 1 and 2 show the results of rotating the received signal point to which only white noise is added by 30°, calculating the rotation angle corrected by each correction method, and evaluating the error with different signal scores. The modulation schemes used here are 64QAM. Fig.3 and 4 show the probability of the correct classification at several SNRs when each correction is performed in a flat fading, where four types of modulation schemes: BPSK, QPSK, 16QAM, and 64QAM are used. From these results, we can see that as compared with the method using only the least squares method, the method combining the LS method with the ICP algorithm does not deteriorate the accuracy of the phase correction even if the number of signal points decreases. The averaged calculation times for one trial of 500 trials are 7.05[s] in LS method and 7.30[s] in LS + ICP method. Comparing with LS method, the calculation time of the proposed method is increased by only 3.5%.

IV. Conclusion

We proposed a new modulation classification technique with phase correction. The experimental results revealed that the proposed method can mitigate the influence of phase fluctuation due to flat fading. Future work is to investigate the effect of the frequency selective fading for our method.

References

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends" IET Communications, vol. 1, no. 2, pp. 137-156, 2007.

[2] Z. Zhu and A. K. Nandi, "Approximate centroid estimation with constellation grid segmentation for blind M-QAM classification" Proc. IEEE Military Communications Conference, pp. 46-51, 2013.

[3] Z. Zhu and A. K. Nandi, "Blind digital modulation classification using minimum distance centroid estimator and non parametric likelihood function" IEEE Transactions on Wireless Communications, vol. 13, no. 8, pp. 4483-4494, 2014.

[4] K. Ichijo, Y. Sugiura, and T. Shimamura, "Automatic Modulation Classification Using Non-Parametric Distance and Phase Likelihood Functions" Proc. NCSP, 2018.

[5] K. Kanatani, "Analysis of 3-D rotation fitting," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.16, no5, pp.543-549, 1994.

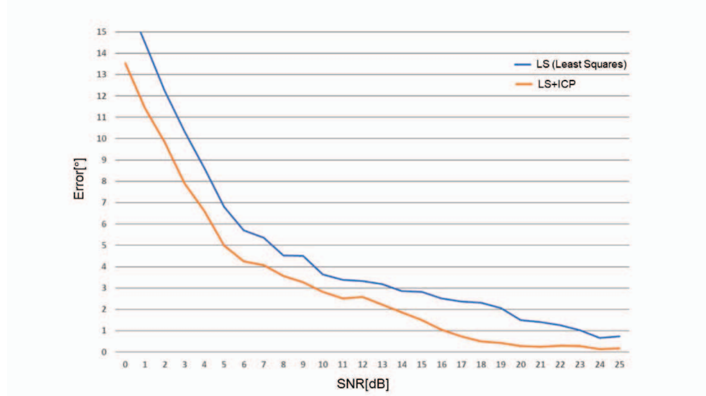


Fig. 1. Phase correction error for two methods with signal length N=64.

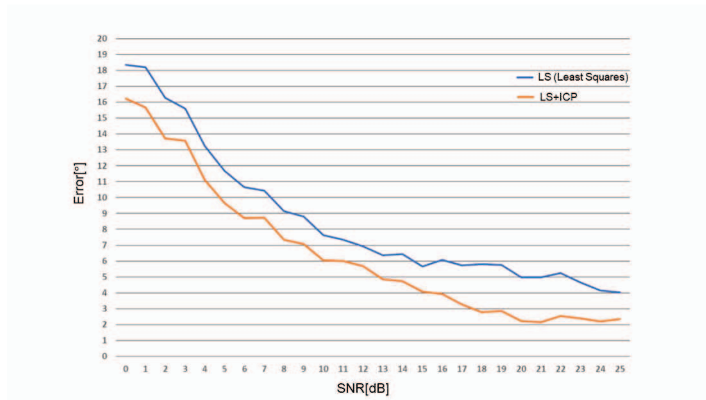


Fig. 2. Phase correction error for two methods with signal length N=256.

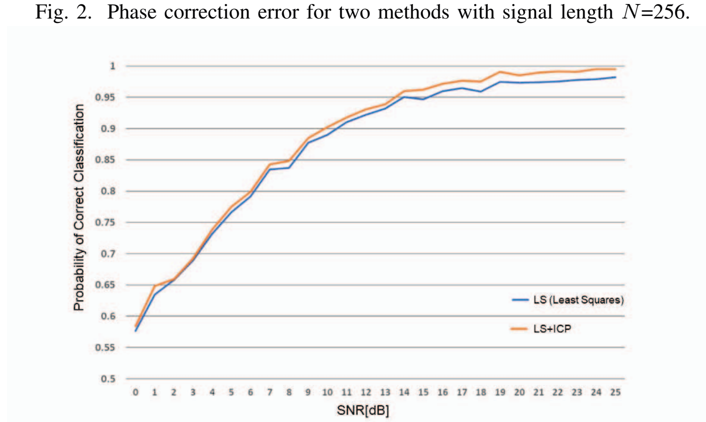


Fig. 3. Classification probability on flat fading channels by phase correction for two methods with signal length N=256.

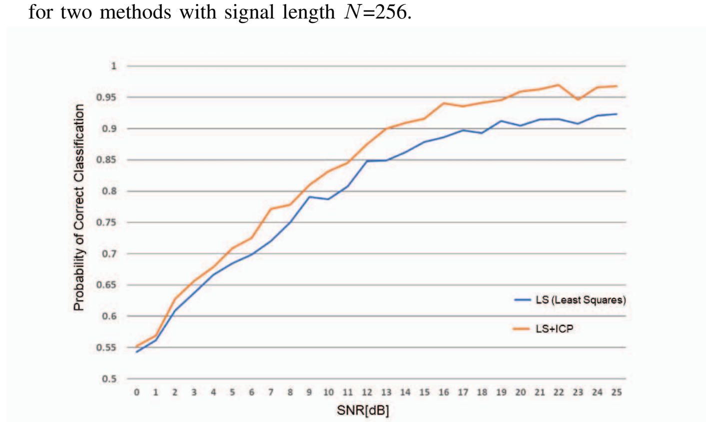


Fig. 4. Classification probability on flat fading channels by phase correction for two methods with signal length N=64.

Inter-cell Coordinated Transmission Power Control for IDMA-Based Random Access

Masayuki Kawata†, Tsubasa Miyata†, Yoshihisa Kishiyama‡, and Kenichi Higuchi†
†Graduate School of Science and Technology, Tokyo University of Science
‡5G Laboratory, NTT DOCOMO, INC.

Abstract—**This paper proposes an inter-cell coordinated transmission power control (TPC) scheme for interleave division multiple access (IDMA)-based random access that supports massive machine-type communications (mMTC) in the fifth generation mobile communication system. In IDMA-based random access, attempts to detect separately multiple packets are made using a successive interference canceller (SIC) at the base station receiver to achieve multi-packet reception. Therefore, the TPC in IDMA-based random access should achieve uniform transmission quality within a cell to actualize mMTC, while taking into account its impact on the signal separation capability during the SIC process and inter-cell interference levels. The proposed scheme is a frequency block-dependent fractional TPC where the overall random access bandwidth is divided into two frequency blocks each of which is allocated to either cell-inner terminals or cell-edge terminals. By coordinating the frequency block usage between neighboring cells, the proposed TPC achieves inter-cell interference statics appropriate for attaining uniform transmission quality within a cell, while the frequency block-dependent fractional TPC achieves appropriate received power distribution among terminals from the viewpoint of the SIC process. Computer simulation results show in quantitative terms the effectiveness of the proposed scheme.**

I. Introduction

In the fifth-generation (5G) mobile communication system [1, 2], massive machine-type communications (mMTC) will be supported to actualize the Internet-of-Things (IoT). To support the mMTC uplink, a grant-free and contention-based multiple access scheme is essential to reduce the control signaling overhead and transmission latency. To suppress the packet loss due to collision and to achieve multi-packet reception [3], non-orthogonal multiple access (NOMA) with interference cancellation [4] at the base station (BS) receiver is essential. In this paper, we focus on interleave division multiple access (IDMA) [5, 6] based random access. In IDMA, multiple packets transmitted simultaneously from different terminals are decomposed by multiuser detection such as using a successive interference canceller (SIC) at the BS receiver in which terminal-specific channel interleavers are utilized. The IDMA signal does not increase the peak power and its transmission performance is robust against error in time/frequency synchronization among terminals. These are desirable properties in the mMTC uplink.

This paper investigates a transmission power control (TPC) scheme appropriate for IDMA-based random access. TPC in IDMA-based random access should achieve uniform transmission quality within a cell to actualize mMTC. Furthermore, TPC should take into account its impact on the signal separation capability in the SIC process and inter-cell interference (ICI) levels in cellular scenarios. Variation in the received signal power levels among simultaneously transmitted packets should be within an appropriate range considering the signal separation aspect in the SIC. An excessively high transmission power at cell-edge terminals results in an increase in ICI and degrades the overall system performance.

To address these problems, we propose an inter-cell coordinated TPC scheme for IDMA-based random access. The proposed scheme is a frequency block-dependent fractional TPC where the overall random access bandwidth is divided into two frequency blocks each of which is allocated to cell-inner terminals or cell-edge terminals. Fractional TPC [7], which partially compensates the path loss between the transmitter and receiver, is widely used in LTE and LTE-Advanced. By coordinating the frequency block usage between neighboring cells, the proposed TPC achieves ICI statics appropriate for attaining uniform transmission quality within a cell, while the frequency block-dependent fractional TPC achieves appropriate received power distribution among terminals from the viewpoint of the SIC process. The effectiveness of the proposed scheme is shown quantitatively based on computer simulation.

The remainder of the paper is organized as follows. First, Section II describes the proposed TPC scheme. Section III presents numerical results based on computer simulations. Finally, Section IV concludes the paper.

II. Proposed TPC Scheme

In the proposed TPC scheme, the overall frequency band prepared for random access at each cell is divided into two frequency blocks. Fig. 1 shows the frequency usage in the proposed TPC scheme. For notational simplicity, we consider a two-cell scenario with target cell 1 and neighboring cell 2. However, the proposed scheme can be extended to general multiple neighboring cell cases similar to fractional frequency reuse (FFR) [8] in ICI coordination.

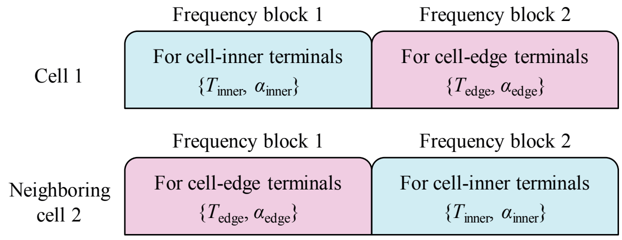


Figure 1. Frequency block usage in proposed scheme.

At each cell, one frequency block is allocated to terminals that are located in the vicinity of the serving BS (hereafter cell-inner terminals). The other frequency block is allocated to terminals near the cell border (hereafter cell-edge terminals). This frequency block usage is coordinated between cells 1 and 2 so that the frequency block for cell-inner terminals in cell 1 corresponds to that for cell-edge terminals in cell 2 and vice versa.

The transmission power density of terminal k, $P_k^{(\mathrm{TX})}$, in decibel-milliwatts per Hertz (dBm/Hz) is determined using the following formula based on fractional TPC.

$$P_k^{(\mathrm{TX})} = \min(T + P_{\mathrm{noise}} + \alpha \cdot L_k, P_{\max})\,(\mathrm{dBm/Hz})\,. \qquad (1)$$

Here T is the target signal-to-noise ratio (SNR) and P_{noise} is the receiver noise power density both in decibel notation. Term $P_{\max}$ is the maximum transmission power density of a user terminal. Term L_k is the average path loss between terminal k and the serving BS. Parameter α ($0 < \alpha \le 1$) is the compensation factor for the average path loss in fractional TPC. When α is one, the received average SNR for all users is maintained at T. However, in this case, the transmission power of the cell-edge terminal may be significantly high, which results in severe ICI. Using α lower than one, which partially compensates for the average path loss, achieves a tradeoff between the ICI level and the signal reception quality of the cell-edge terminals. In the proposed scheme, set (T, α) can be differently configured between the cell-inner and cell-edge

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

terminal groups, which are denoted as (T_{inner}, α_{inner}) and (T_{edge}, α_{edge}), respectively.

The expected advantageous properties of the proposed scheme are given hereafter. First is the ICI coordination effect. With fractional TPC, the cell-edge terminal uses high transmission power, which causes high ICI and is weak against it. Since the frequency block for cell-edge terminals corresponds to that for cell-inner terminals in a neighboring cell, the strong ICI generated by the cell-edge terminals is mitigated by the cell-inner terminals that are robust against the ICI. Meanwhile, the transmission quality of the cell-edge terminals is protected since the cell-inner terminals in the neighboring cell do not cause severe ICI. Overall, the proposed scheme reduces the variation in transmission quality depending on the terminal location within the cell. The second property is the appropriate variation in the received signal power levels among simultaneously transmitted packets within a frequency block from the viewpoint of the SIC process since the terminals are separated into two groups based on the path loss.

III. Numerical Results

We assume discrete Fourier transform (DFT)-spread orthogonal frequency division multiplexing (OFDM)-based single-carrier transmission [9]. The number of information bits per packet is 104. The overall bandwidth for random access is 2.5 MHz and that is equally divided into two frequency blocks in the proposed scheme. The packet length is 0.5 ms. As channel coding, we use a combination of the turbo and repetition codes. The code rate for the turbo code is 1/3, which is used in the 3GPP [9]. The number of repetitions in the repetition coding is set to six and three in the conventional scheme without using frequency blocks and the proposed scheme. QPSK data modulation is assumed. A random interleaver is used as a channel interleaver. As the channel mode, we assume distance-dependent average path loss with a decay factor of 3.76 and instantaneous block Rayleigh fading with an exponentially decayed six-path model where the rms delay spread is set to 1.3 μs. The $P_{\max}$ value is 24 dBm and P_{noise} is −169 dBm/Hz. Four-branch receiver antenna diversity followed by the SIC in [10] are employed. Iterative channel estimation using a Zadoff-Chu sequence-based preamble and decision feedback data symbols is assumed. We assume the two-BS line model shown in Fig. 2. Packet arrival follows as a Poisson distribution and the average number of packet-transmitting terminals per cell is set to 20. The locations of the terminals are uniformly distributed within a cell where the probability distribution function at distance d km from the serving BS is $2d$ mimicking the two-dimensional cell configuration. The threshold distance to the serving BS for the cell-inner and cell-edge terminal groupings is set to $1/\sqrt{2} \approx 0.707$ km.

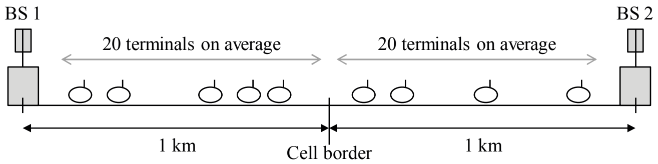


Figure 2. Two-BS line model.

Figs. 3 and 4 show the average signal-to-inter-cell interference plus noise ratio (SINR) and average packet error rate (PER), respectively, as a function of the terminal distance to the serving BS. In the proposed scheme, two cases are tested: one assumes the same (T, α) for two frequency blocks and the other assumes different (T, α) for two frequency blocks as shown in the figure. The proposed scheme using different (T, α) between the frequency blocks lowers the worst SINR and reduces the SINR variation within a cell. As a result, the proposed scheme achieves more uniform PER performance within a cell irrespective of the terminal distance compared to the conventional scheme.

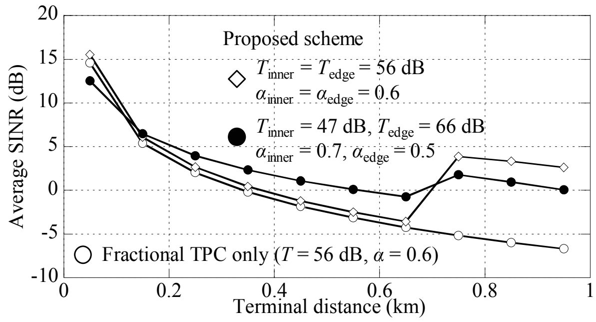


Figure 3. Average SINR as a function of terminal distance.

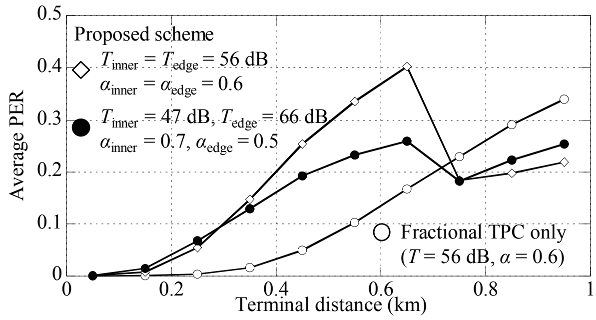


Figure 4. Average PER as a function of terminal distance.

IV. Conclusion

The proposed TPC scheme successfully lowers the worst case PER and achieves a more uniform PER distribution within a cell compared to that for the conventional scheme. This contributes to reducing the outage probability of the mMTC service in practice. Extension and evaluation of the proposed scheme in a general multicell scenario is left for future study.

References

[1] ITU-R, “IMT Vision − Framework and overall objectives of the future development of IMT for 2020 and beyond,” Recommendation M.2083-0, Sept. 2015.

[2] 3GPP TR38.913 (V0.4.0), “Study on scenarios and requirements for next generation access technologies (Release 14),” June 2016.

[3] J. Metzner, “On improving utilization in ALOHA networks,” IEEE Trans. Commun., vol. COM-24, no. 4, pp. 447-448, April 1976.

[4] K. Higuchi and A. Benjebbour, “Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access,” IEICE Trans. Commun. vol. E98-B, no. 3, pp. 403-414, March 2015.

[5] L. Ping, L. Liu, K. Wu, and W. K. Leung, “Interleave division multiple-access,” IEEE Trans. Wireless Commun., vol. 5, no. 4, pp. 938-947, April 2006.

[6] Y. Hu, C. Xu, and L. Ping, “NOMA and IDMA in random access systems,” in Proc. IEEE VTC2018-Spring, Porto, Portugal, June 2018.

[7] W. Xiao, R. Ratasuk, A. Ghosh, R. Love, Y. Sun, and R. Nory, “Uplink power control, interference coordination and resource allocation for 3GPP E-UTRA,” in Proc. IEEE VTC2006-Fall, pp. 1-5, Montréal, Canada, Sept. 2006.

[8] T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, “Analytical evaluation of fractional frequency reuse for OFDMA cellular downlink,” IEEE Trans. Wireless Commun., vol. 10, no. 12, pp. 4294–4305, Dec. 2011.

[9] 3GPP, TS 36.300, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description.

[10] M. Kawata, K. Tateishi, and K. Higuchi, “Investigation on structure of interference canceller for IDMA-based random access,” in Proc. IEEE VTC2018-Fall, Chicago, U.S.A., 27-30 Aug. 2018.

Low Latency HARQ Method Using Early Retransmission Before Channel Decoding Based on Superposition Coding

Koka Miura[†], Yoshihisa Kishiyama[‡], and Kenichi Higuchi[†]
[†]Graduate School of Science and Technology, Tokyo University of Science
[‡]5G Laboratory, NTT DOCOMO, INC.

Abstract—**This paper proposes a superposition coding-based early retransmission method for a previously reported low latency hybrid automatic repeat request (HARQ) method using early retransmission before channel decoding. The low latency HARQ method mitigates the increased transmission latency resulting mainly from the delay time required for channel decoding in HARQ by requesting early retransmission before the channel-decoding process is completed. However, early-retransmission decision may include error, which results in unnecessary retransmission and consequently throughput loss. The proposed method mitigates this throughput loss by multiplexing the early-retransmission packet and initial packet for the next transmission attempt within the same channel using superposition coding. The receiver applies an inter-packet interference canceller to cancel the interference between the superposition coded packets. Simulation results show that the proposed method further improves the achievable tradeoff between the transmission latency and throughput.**

I. INTRODUCTION

This paper investigates a low latency hybrid automatic repeat request (HARQ) method [1] aiming at application to ultra-reliable low latency communications (URLLC) in the fifth-generation mobile communication system [2]. Although HARQ efficiently achieves low error-rate transmission, this protocol suffers from increased transmission latency resulting mainly from the delay time required for channel decoding. To address this problem, members of our research group recently reported a low latency HARQ method that uses channel state information (CSI) prior to channel decoding [3]. This method mitigates the increased transmission latency by requesting early retransmission before the channel decoding process is completed based on the CSI obtained prior to channel decoding as shown in Fig. 1. As for the CSI metric for deciding the early-retransmission request, [3] reported that a mutual information (MI)-based metric calculated from the signal-to-noise ratio (SNR) of the received packet achieves reasonable decision accuracy.

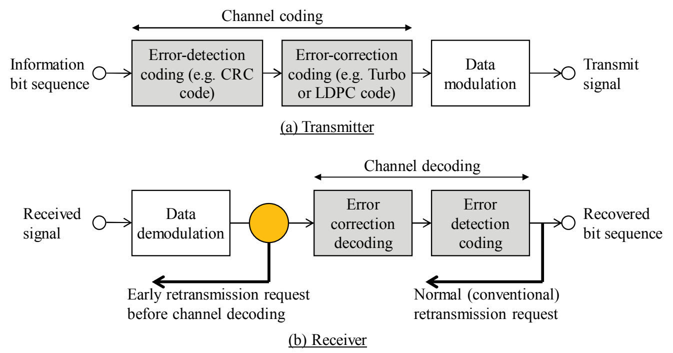


Figure 1. Low latency HARQ method using early retransmission.

However, a perfect decision on conducting early retransmission, and thus perfect estimation concerning whether or not bit error exists after channel decoding, based on the CSI before channel decoding is not possible in principle. When early retransmission is requested despite that the initial packet is eventually decoded successfully, this unnecessary packet retransmission results in throughput loss.

To address this potential throughput loss, we propose applying superposition coding [4-6] to the previously reported low latency HARQ method in [3]. The proposed method mitigates the throughput loss due to the unnecessary early retransmission by multiplexing the early-retransmission packet and initial packet for the subsequent transmission attempt within the same channel (time slot) using superposition coding. The receiver applies an inter-packet interference canceller to cancel the interference to the initial packet for the subsequent transmission attempt from the superimposed early-retransmission packet. The point is that when the early retransmission is unnecessary because the corresponding initial packet is eventually decoded successfully, the interference canceller works perfectly since the successfully recovered information-bit sequence of the early-retransmission packet can be utilized. This contributes to enhancing the performance obtained using the proposed method. Simulation results show quantitatively that the proposed method further improves the achievable tradeoff between the transmission latency and throughput.

The remainder of the paper is organized as follows. First, Section II describes the proposed HARQ method. Section III presents numerical results based on computer simulations. Finally, Section IV concludes the paper.

II. PROPOSED HARQ METHOD

In the proposed method, when early retransmission is requested, the early-retransmission packet and initial packet for the subsequent transmission attempt are multiplexed within the same channel using superposition coding to avoid potential throughput loss due to the unnecessary early retransmission. Fig. 2 illustrates the proposed superposition coding method assuming QPSK data modulation. For simplicity, the maximum number of retransmissions is assumed to be one in the paper.

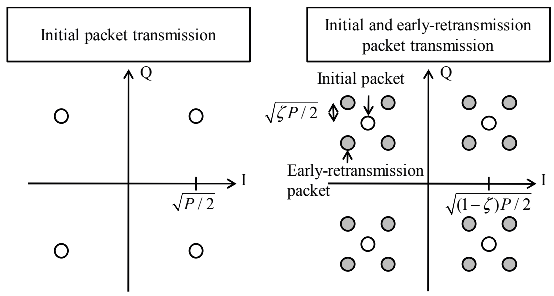


Figure 2. Superposition coding between the initial and early-retransmission packets.

We consider a case where the early-retransmission packet signal, $x_{\mathrm{ret}}[t]$, and the initial packet signal for the subsequent transmission attempt, $x_{\mathrm{init}}[t]$, are transmitted simultaneously using the same channel based on superposition coding at the t-th time instance. Packet signals $x_{\mathrm{ret}}[t]$ and $x_{\mathrm{init}}[t]$ are generated based on independent channel coding, interleaving, and symbol mapping. The superimposed transmission signal, $s[t]$, is represented as

$$\begin{aligned} s[t] &= \sqrt{P_{\mathrm{init}}}\,x_{\mathrm{init}}[t] + \sqrt{P_{\mathrm{ret}}}\,x_{\mathrm{ret}}[t]. \\ P_{\mathrm{init}} &= (1-\zeta)P, P_{\mathrm{ret}} = \zeta P \end{aligned} \quad (1)$$

Here, P is the total transmission power per symbol. Terms P_{init} and P_{ret} are transmission powers allocated to $x_{\mathrm{init}}[t]$ and $x_{\mathrm{ret}}[t]$, respectively. Parameter ζ represents the ratio of the transmission power allocation to $x_{\mathrm{ret}}[t]$.

In the proposed method using superposition coding, inter-packet interference occurs between packets multiplexed within the same channel. The receiver applies an inter-packet interference canceller to cancel the interference imparted to the initial packet for the subsequent transmission attempt from the superimposed early-retransmission packet. Fig. 3 is a block diagram of the interference canceller. The interference canceller first decodes the early-retransmission packet from the received signal. Then, the signal component of the early-retransmission packet is subtracted from the received signal using the decoding results of the early-retransmission packet. Finally, the initial packet is decoded using the received signal after subtracting the signal component of the early-retransmission packet.

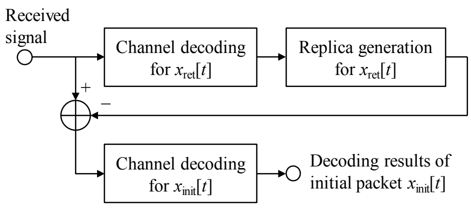


Figure 3. Block diagram of the inter-packet interference canceller.

It is worth emphasizing that when early retransmission is unnecessary because the corresponding initial packet is eventually decoded successfully, which is the main problem on which we focus, perfect interference cancellation is achieved since the successfully recovered information-bit sequence of the early-retransmission packet can be utilized for generation of the interference replica. In this case, the performance loss due to the unnecessary early retransmission in the proposed method is merely a $(1-\zeta)$-times power loss to the initial packet. Therefore, it is expected that the throughput loss can be mitigated using the proposed method compared to using the conventional method where the unnecessary early retransmission directly results in bandwidth loss.

III. Numerical Results

Orthogonal frequency division multiplexing (OFDM) with a 4.5-MHz transmission bandwidth is assumed as the basic signal transmission scheme. QPSK data modulation is used. One packet comprises 14 OFDM symbols corresponding to a 1-ms packet duration. A rate-1/3 turbo code with the constraint length of four is used as the error-correction channel coding. As the channel model, six-path block Rayleigh fading with a 1-μs delay spread is assumed. The maximum Doppler frequency is set to 5.55 Hz. Maximal ratio combining with two-branch receiver antenna diversity is applied. As for the packet combining between the initial and retransmission packets, chase combining is assumed. The transmission interval for the retransmission packet is assumed to be 2 ms for early retransmission in the proposed method. When retransmission is requested after channel decoding in the same way as in the conventional HARQ, the transmission interval for the retransmission packet is set to 8 ms. The transmission delay time is assumed to be 4 ms when the initial packet is successfully decoded, 6 ms when the early-retransmission packet is successfully decoded when the initial packet failed to decode, and 12 ms when the conventional retransmission packet after channel decoding is correctly decoded when the initial packet failed to decode. In the proposed method, ζ is set to 0.2. The conventional low latency HARQ method in [3] and the case where early retransmission is not performed are tested for comparison.

Figs. 4 and 5 show the average throughput and transmission delay, respectively, as a function of the SNR. The channel code rate is changed from 1/3 to 3/4 depending on the SNR so that the throughput is maximized for the respective HARQ methods to simulate adaptive modulation and channel coding (AMC). The achievable throughput for the proposed method is higher than that for the conventional method and is approximately the same as that without early retransmission. Meanwhile, the proposed method achieves the shortest transmission delay time. This is because the proposed method mitigates the throughput loss using the superposition coding-based packet multiplexing when unnecessary early retransmission is requested. This allows the proposed method to use early retransmission more aggressively, which contributes to reducing further the transmission delay time.

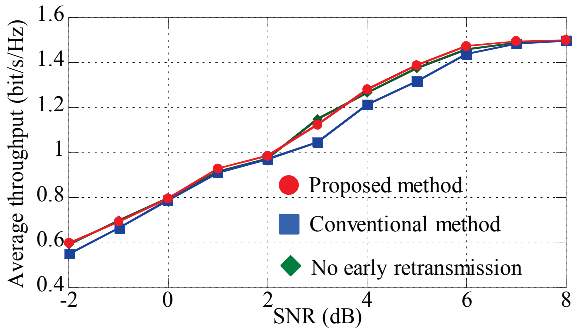


Figure 4. Average throughput as a function of SNR.

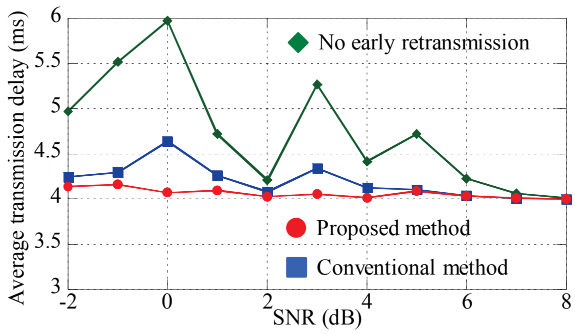


Figure 5. Average transmission delay as a function of SNR.

IV. Conclusion

The proposed low latency HARQ method using early retransmission before channel decoding based on superposition coding further improves the achievable tradeoff between the transmission latency and throughput compared to the conventional low latency HARQ method.

References

[1] D. N. Rowitch and L. B. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo (RCPT) codes," in IEEE Trans. Commun., vol. 48, no. 6, pp. 948-959, June 2000.
[2] ITU-R, "IMT Vision − Framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation M.2083-0, Sept. 2015.
[3] Y. Imamura, D. Muramatsu, Y. Kishiyama, and K. Higuchi, "Low latency hybrid ARQ method using channel state information before channel decoding," in Proc. APCC2017, Perth, Australia, Dec. 2017.
[4] F. Takahashi and K. Higuchi, "HARQ for predetermined-rate multicast channel," in Proc. IEEE VTC2010-Spring, Taipei, Taiwan, May 2010.
[5] R. Zhang and L. Hanzo, "Superpositon-coding aided multiplexed hybrid HARQ scheme for improved link-layer transmission efficiency," in Proc. IEEE ICC2009, Dresden, Germany, June 2009.
[6] Y. Hasegawa and K. Higuchi, "Bi-directional signal detection and decoding for hybrid ARQ using superposition coding," in Proc. IEEE VTC2011-Fall, San Francisco, U.S.A., Sept. 2011.

Machine learning model with technical analysis for stock price prediction: Empirical study of Semiconductor Company in Taiwan

Po-Chao Lan
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
s981724@gmail.com

Wei-Ling Kung
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
winniekung0524@gmail.com

Yao-Lun Ou
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
a775869@gmail.com

Chun-Yueh Lin
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
ljy898@gmail.com

Wen-Cheng Hu
Department of Artificial Intelligence
CTBC Business School
Tainan, Taiwan
wencheng@ctbc.edu.tw

Yi-Hsien Wang
Department of Banking and Finance
Chinese Culture University
Taipei, Taiwan
holland@mail2000.com.tw

***Abstract*—The stock market was affected by different variables, such as the overall economic situation, political events, Sino-US relations and corporate operations. Therefore, if you want to get returns in the stock market, predicting the time series of financial markets in advance is the most important thing for analysts and investors. However, predicting the direction of the stock market need to access information from existing markets and past historical data. Under such complicated work and costs, it is always the most difficult and important issue to achieve accurate forecasting and reduce forecasting costs. In this paper, the back-propagation neural network is used as a research tool to analyze the historical data of Taiwan Semiconductor Manufacturing Company (hereinafter referred to as TSMC) during the sample period from 2014 to 2018. In this study, the standardized technical analysis indicators and the related variables of TSMC are taken as input variables, and the closing price of the next day is taken as the output variable to predict the closing price for TSMC of the next day. The empirical results confirm that this method does improve the forecast of stock price of TSMC.**

Keywords—neural network 、Time Series 、forecast 、technical analysis

I. INTRODUCTION

Reference [10] suggested that stock price forecasting has always been the most important thing in financial markets, which has become one of the concerns from investors and shareholders, and stock price forecasts can be divided into three categories. Including fundamental analysis, technical analysis and technology method analysis. Most investors will invest based on fundamental analysis, and fundamental analysis is divided into three categories. First one is analysis of the macroeconomic aspects. Second one is forecasting the value of the enterprise for the current status of the industry. Last one is the enterprise operations and internal value.

However, in addition to analyzing the stock market in the past, more and more studies are conducting stock market forecasts through technical methods. Among them, machine learning are widely used in recent years and gradually become a research trend [2] [5] [6]. Therefore, this paper uses the back-propagation neural network in machine learning to predict the TSMC with the highest market value and weight in the Taiwan Weighted Stock Index. Main source of profit of TSMC is foundry, and Samsung of South Korea is its main competitor. The competitive relationship between TSMC and Samsung will have an impact on the stock price of TSMC. More importantly, Reference [8] indicated that average domestic production of Samsung of South Korea was 28,401 from 2014 to 2018, and its operations will directly affect the overall economy of Korea. Therefore, this study takes the exchange rate of the Taiwan dollar to the Korean won and the Taiwan weighted stock price index as input variables to explore. The empirical results also confirm that the exchange rate and the market are brought into the input variables, which helps to reduce the prediction error to improve the accuracy.

This paper hopes to increase the possibility of investment profit of investors in the Taiwan market by predicting the trend of stock price of TSMC. This paper is divided into four paragraphs. The first part is the introduction, which explains the research background, motivation and research purposes. The second part is the research method, the third part is the empirical result, and the last part is the research conclusion of this paper.

II. METHODOLOGY

A. NEURAL NETWORK

Neural networks are widely used in the prediction of time series. For example, Babu and Reddy designed a new hybrid ARIMA-ANN model to predict time series data. The empirical results show that the proposed hybrid model has higher prediction accuracy for multi-step-advance time series prediction[3].

In the stock price prediction, the neural network will change the combination of variables and the structure because of the different property and characteristics of the target, in order to make the prediction result more accurate. There are many studies that use these indicators as input functions, such as Chang, Fan and Liu, which input them from the turning point (trough or peak) of historical data to the back-propagation neural network (BPNN) used for training of models. It is predicted that the stock will be bought or sold. Reference [4] proposes that the ability of learning and promotion of neural network is well suited to problem areas such as stock market forecasting in nonlinear data trends. In addition, the ANN is able to adapt to the data pattern and the relationship between input layer and output layer, resulting in better prediction accuracy than traditional methods.

The most commonly used model in the neural network model is the representative back-propagation neural network model. This model belongs to the supervised learning network, and the back-propagation neural network model is used during learning, it will be fed back the wrong signal, and the weight of the processing unit will be corrected by the error signal. Because of the back-propagation neural network model, its characteristics are fast recall, high learning precision, and the ability to process nonlinear relational data. In practice, back-propagation neural network are often used for classification, evaluation, and prediction. Therefore, this study uses an back-propagation neural network to establish a Taiwan stock market forecasting model. Tripathy, N predicts changes in the daily price of Indian stock market for the eight-year period from January 1, 2008 to April 8, 2016 by using a back-propagation neural network. The study shows that the predicted values are very close to the actual values.

III. Empirical results

In 2014, the best predictive combination was predicted in the 5-day prediction group, the hidden layer neurons were combined into 2 layers (14, 6) with an RMSE of 0.0971. The best predictive combination in the 10-day prediction group, the hidden layer neurons were combined into 2 layers (14, 6) with an RMSE of 0.0983. The best predictive combination in the 20-day prediction group, the hidden layer neurons were combined into 2 layers (14, 11) with an RMSE of 0.1121.

In 2015, the best predictive combination in the predicted 5-day prediction group, the hidden layer neurons were combined into 2 layers (1, 11) with an RMSE of 0.0610. The best predictive combination in the 10-day prediction group, the hidden layer neurons were combined into 2 layers (1, 6) with an RMSE of 0.0620. The best predictive combination in the 20-day prediction group, the hidden layer neurons were combined into a layer of 9 neuron combinations with an RMSE of 0.0664.

In 2016, the best predictive combination in the predicted 5-day prediction group, the hidden layer neurons were combined into a 14-layer neuron combination with an RMSE of 0.0690. The best predictive combination in the 10-day prediction group, the hidden layer neurons were combined into a single layer of 11 neuron combinations with an RMSE of 0.0686. In the best predictive combination in the 20-day prediction group, the hidden layer neurons were combined into a 14-layer neuron combination with an RMSE of 0.0524.

In 2017, the best predictive combination in the predicted 5-day prediction group, the hidden layer neurons were combined into a 13-layer neuron combination with an RMSE of 0.0682. In the 10-day prediction group, the best combination of predictions, the hidden layer neurons were combined into a layer of 13 neuron combinations with an RMSE of 0.0694. The best predictive combination in the 20-day prediction group, the hidden layer neurons were combined into 2 layers (12, 14) with an RMSE of 0.0741.

In 2018, the best predictive combination was predicted in the 5-day prediction group. The hidden layer neurons were combined into 2 layers (14, 3) with an RMSE of 0.1089. The best predictive combination in the 10-day prediction group, the hidden layer neurons were combined into 2 layers (8, 9) with an RMSE of 0.1093. The best predictive combination in the 20-day prediction group, the hidden layer neurons were combined into 2 layers (14, 11) with an RMSE of 0.1216.

IV. Conclusion and Suggestions

According to the empirical results, in the 5-day of the predictive data sets when predicting 2017, the hidden layer neurons set which is comprised of 1 layer and 13 neurons is the best. In the 10-day of the predictive data sets when predicting 2015, the hidden layer neurons set which is comprised of 2 layers (1, 6) is the best. And in the 20-day of the predictive data sets when predicting 2016, the hidden layer neurons set which is comprised of 1 layer and 14 neurons is the best. Prediction of the overall neural network, the best set of prediction neural network in the 20-day is comprised of 1 layer and 14 neurons of 2016, RMSE reached to 0.0524.

This study constructed the neural network which mixed the technical analysis indicators mostly for short-term data, and the empirical results show that the long-term forecast performance is relatively low. Therefore, subsequent studies can screen long-term data and bring in input variable to test predictions for each combination to improve the long-term predictive ability.

References

[1] A. Bhatia, H. Hagras, and J. J. Lepley, "Machine Learning Approach to Extracting Emotions Information from Open Source Data for Relative Forecasting of Stock Prices," In *2018 10th Computer Science and Electronic Engineering (CEEC)*, 2018, pp. 142-147. IEEE.

[2] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical Analysis and Sentiment Embeddings for Market Trend Prediction," *Expert Systems with Applications*, 2019.

[3] C. N. Babu, and B. E. Reddy, "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data," *Applied Soft Computing*, vol.*23*, 2014, pp.27-38.

[4] C. S. Vui, G. K. Soon, C. K. On, R. Alfred, and P. Anthony, "A review of stock market prediction with Artificial neural network (ANN). In 2013 IEEE International Conference on Control System," Computing and Engineering, 2013, pp. 477-482. IEEE.

[5] D. S. Sisodia, and S. Jadhav, "Machine Learning Models for Forecasting of Individual Stocks Price Patterns," In *Handbook of Research on Pattern Engineering System Development for Big Data Analytics*, 2018, pp.111-129. IGI Global.

[6] D. Shah, H. Isah, and F. Zulkernine, "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques," International Journal of Financial Studies, vol.7, 2019, pp.26.

[7] M. Sedighi, H. Jahangirnia, M. Gharakhani, and S. Farahani Fard, "A Novel Hybrid Model for Stock Price Forecasting Based on Metaheuristics and Support Vector Machine," *Data*, vol.*4*, 2019, pp.75.

[8] N. Tripathy, "Predicting Stock Market Price Using Neural Network Model," International Journal of Strategic Decision Sciences (IJSDS), vol.9, 2018, pp.84-94.

[9] P. C. Chang, C. Y. Fan, and C. H. Liu, "Integrating a piecewise linear representation method and a neural network model for stock trading points prediction. IEEE Transactions on Systems, Man, and Cybernetics," Part C (Applications and Reviews), vol.39, 2008, pp.80-92.

[10] S. P. Das, and S. Padhy, "A novel hybrid model using teaching–learning-based optimization and a support vector machine for commodity futures index forecasting," *International Journal of Machine Learning and Cybernetics*, vol.*9*, 2018, pp.97-111.

A Spur-Suppression Technique for Frequency Synthesizer With Pulse-Width to Current Conversion

Po-Yu Hsieh, Shao-Yu Shu, and Ching-Yuan Yang
Department of Electrical Engineering, National Chung Hsing University
145 Xingda Rd., South Dist., Taichung City 40254, Taiwan (R.O.C)

***Abstract*—A pulse-width to current conversion circuit is adopted to suppress reference frequency spurious tone of frequency synthesizer. Propose pulse-width to current conversion technique convert pulse-width of input clock signal to modulated output current. This characteristic of pulse-width sensing applies to replace conventional charge-pump of phase-lock loop. Continuous and linear output current improve the reference frequency spur, which results from non-ideal effect of charge-pump. Proposed frequency synthesizer is fabricated with TSMC 0.18µm CMOS process and operation frequency is between 5.12 GHz and 5.28 GHz under the channel width of 10 MHz. The total power consumption is 23 mW without output buffer. The measured output spurious tone is -48.21 dBm and phase noise performance is -102 dBc/Hz at 1 MHz offset.**

Keywords*—*spur suppression, frequency synthesizer, charge-pump (CP), phase-locked loop (PLL), pulse-width to current converter (PWCC).

I. Introduction

With the popularity of mobile device, the demand of wireless communication systems increases gradually and it usually requires a clock signal with high spectral purity to up-convert or down-convert signal. To assure the accuracy of signal, local oscillator (LO) play an important role in many communication system. Therefore, a low phase noise and low spurious tone are necessary for frequency synthesizer.

Reference Frequency spur result from the non-ideal effect of charge-pump (CP) and phase-frequency detector (PFD). The signal skew of up and down clock from phase-frequency detector and the mismatch of charge-pump output current result in spurious tone on spectral. And hence, a period ripple which usually equals to reference frequency on control voltage of voltage control oscillator (VCO) is generated. To suppress the reference frequency spur, several method have be presented. In this work, we present a different architecture of charge-pump to reduce current mismatch and clock skew. The pulse-width to current converter (PWCC) generates a modulated current varies with pulse-width to charge/discharge the loop filter of phase-lock loop. Phase-frequency detector provides the input clock signal to PWCC. If the phase-lock loop has larger phase error, the PWCC will provide larger output current to loop filter, and vice versa. The conventional charge pump control output current with switch on and off, and the current changes rapidly. For this reason the conventional charge pump would be sensitive to clock skew of up and down signal. The PWCC provide a constant and smooth charging current for loop filter to suppress the magnitude of the reference frequency spur. Architecture of PWCC and circuit implementation will be present in Section II. Section III shows the experimental results of PWCC-PLL and Section IV concludes this work.

II. Circuit Description

The architecture of proposed PWCC-PLL is shown as Fig.1. Phase-frequency detector detects phase error of frequency synthesizer and generates the signal CLK_{up}/CLK_{dn}. A pair of PWCC circuit converts the pulse-width of signal CLK_{up}/CLK_{dn} into control voltage ($V_n \pm$ and $V_p \pm$). Next, Mirror block (MIR) converts the $V_n \pm$ and $V_p \pm$ to moderate current level for proper system parameter. PWCC provides the smoother current to loop filter. On the hand, we can say that PWCC provides an additive pole for loop filter to filter out the spurious tone.

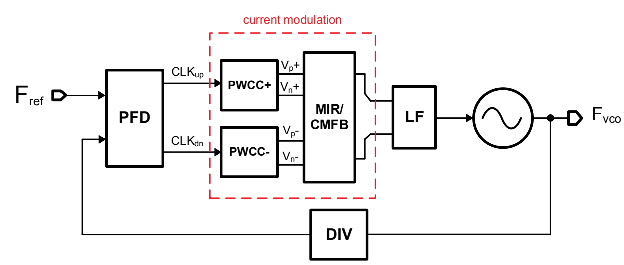
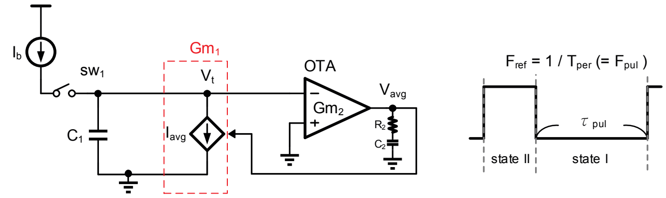


Fig. 1. Architecture of PWCC-PLL frequency synthesizer.

To verify the design of PWCC, the architecture of PWCC can be shown in Fig.2(a). There is a switch (SW_1) to control the current I_b. The current I_b charges and the other current source I_{avg} discharges the capacitor C_1. The OTA_1 controls the current I_{avg} and there is a small RC filter (R_2 and C_2) on feedback node V_{avg}. If we have the input clock signal as Fig.2(b), switch SW_1 will turn on in τ_1 (state I). At steady state, the carrier of I_b and I_{avg} will reach equilibrium in each clock cycle T_{per}. We can find the close loop characteristic function of PWCC.

$$H(s)|_{PWCC} = \frac{V_{avg}}{\tau_1}(s) = \frac{I_b Gm_1 R_2}{2\pi C_1} \cdot \frac{s+\frac{1}{R_2 C_2}}{s^2+\frac{Gm_1 Gm_2 R_2}{C_1}+\frac{Gm_1 Gm_2}{C_1 C_2}} \quad (1)$$

τ_1 is input pulse-width (phase error) and capacitance of RC filter is enlarged by transconductance Gm_1 and Gm_2. There are a zero and two pole for extra signal filtering.



(a) (b)

Fig. 2. (a) Architecture of PWCC. (b) Input clock period.

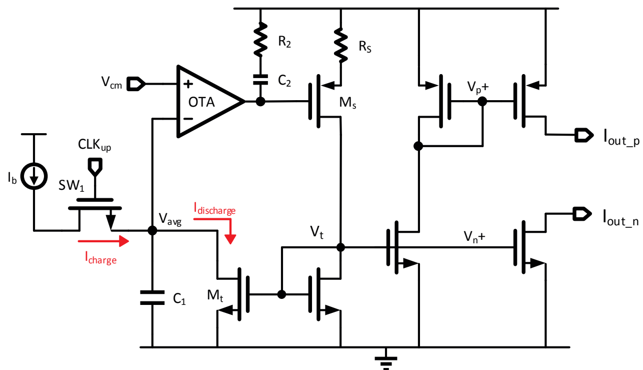


Fig. 3. Circuit implementation of PWCC

The circuit realization can be shown as Fig.3 [1]. I_{out_n} and I_{out_p} are output current and degenerative resistor R_S improves the linearity of transconductance Gm_1. Characteristic of feedback system generates an average current I_{avg} ($I_{discharge}$ on Fig. 3) which equals to $I_b \times \tau_1$ in each reference cycle and reaches better area efficiency of extra RC filter then traditional 3-order loop filter of phase-lock loop.

III. Experimential Results

This work was fabricated in TSMC 0.18-µm CMOS technology. Fig. 4 shows the chip micrograph of PWCC-PLL and active area is 0.85×0.8 mm^2 without ESD. The range of VCO output frequency is 4.8 to 5.32 which covers our divider ratio. When system locked at 5.19 GHz, the output phase noise measurement is -102.12 dBc/Hz at 1 MHz shown in Fig. 5. The output spectrum is shown as Fig. 6 with the reference frequency spur suppression of -48.21 dBm at 10 MHz offset. Table I. gives a summary of experimental results.

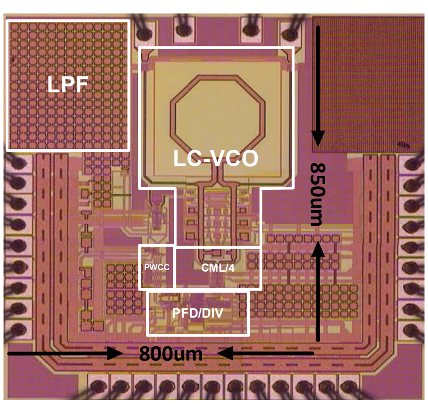


Fig. 4. Micrograph of PWCC-PLL

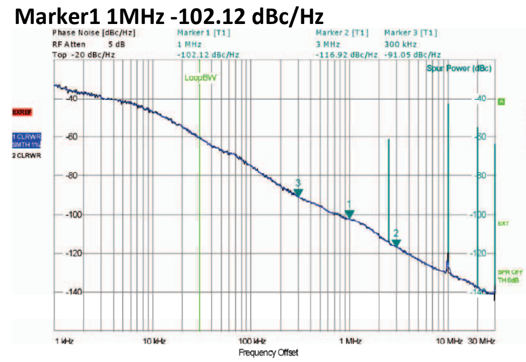


Fig. 5. Phase noise measurement of PWCC-PLL

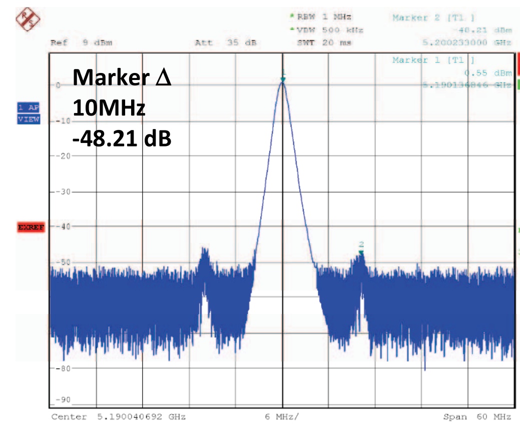


Fig. 6. Output spectrum of PWCC-PLL

TABLE I. Performance Summary

Parameter	Implementation
Technology	TSMC 0.18-um CMOS
Ref. Freq. (MHz)	10
Output Freq. (GHz)	5.12-5.28
Bandwidth (kHz)	250
K_{VCO} (MHz/V)	200
Phase Noise (dBc/Hz)	-102.12 @1 MHz
Ref. Spur (dBm)	-48
Power (mW)	23

IV. Conclusion

The propose PWCC-PLL improve the reference frequency spur of synthesizer. The current modulation technique senses the variation of input clock pulse-width and converts it into a continuous output current. The linear control current improve the non-ideal effect of switching control charge-pump. Additionally, transconductance Gm_1 and Gm_2 of PWCC enlarges the capacitors C_1 and C_2 to improve the area efficiency of PWCC. This work presents an optional solation for spur suppression of frequency synthesizer.

References

[1] A. Mortara and E. A. Vittoz, "A 12-transistors PFM demodulator for analog neural networks communication" *IEEE Trans. Neural Networks*, vol. 6, no. 5, 1995.

Action Conditioned Response Prediction with Uncertainty for Automated Vehicles

Hayoung Kim, Gihoon Kim, Jongwon Park, Kyushik Min, Dongchan Kim and Kunsoo Huh
Department of Automotive Engineering
Hanyang University
Seoul, Korea
{hayoung.kim, rlgns4861, pjw2091, mks0813, jookker, khuh2}@hanyang.ac.kr

***Abstract*—Interaction-aware prediction is a critical component for realistic path planning that prevents automated vehicles from overly cautious driving. It requires to consider internal states of other driver such as driving style and intention, which the automated vehicle cannot directly measure. This paper proposes a probabilistic driver model for response prediction given the planned future actions of automated vehicle. The drivers internal states are considered in an unsupervised manner. The prediction model utilizes mixture density network to estimate future acceleration and yaw-rate profile of interacting vehicles. The proposed method is evaluated by using real-world trajectory data.**

***Index Terms*—action conditioned prediction, mixture density network, response prediction, autonomous vehicle**

I. Introduction

In most motion planning studies [1]–[3] for autonomous vehicles, the robot first makes predictions of surrounding vehicles and then plans the trajectory based on the predictions. These approaches tend to make the vehicle overly cautious because they consider surrounding vehicles not as interacting objects but as obstacles to avoid.

For naturalistic driving, An automated vehicle is required to have the ability to interact with surrounding vehicles while driving. When the automated vehicle tries to merge into dense traffic, the success of the attempt depends on the concession of the traffic participant. In other words, the other human driver's action can influence to the autonomated vehicle's action, and vice versa shown as Fig. 1.

In this context, it is important to predict responses of human taking into account uncertainty arising from the interaction or driver's internal states (e.g. intentions, preference and driving style). From this point of view, the main contributions of this paper are:

- Developing a model to predict the response of the human driver given the planned action of the automated vehicle.
- Considering internal states of the human driver in an unsupervised manner.
- Predicting both the response and its uncertainty.

This work was supported by the Industrial Strategic Technology Development Program(10079730, Development and Evaluation of Automated Driving Systems for Motorway and City Road) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).

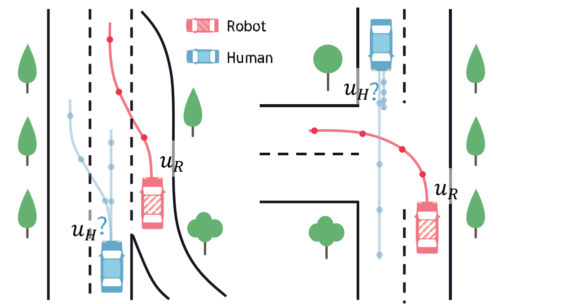


Fig. 1: Various situations that require inter-vehicle interaction

II. Probabilistic Human Driver Model

A. Problem Formulation

The goal of response prediction problem is learning a posterior distribution of human responses, u_H, under planned action of robot, u_R, initial states, x^0, and internal states of human, z. This distribution can be written as $P(u_H^{0:N_p}|u_R^{0:N_p}, x^0, z)$ where $u_H^k = [a_R^k, \dot{\psi}_R^k]^T, u_R^k = [a_H^k, \dot{\psi}_H^k]^T$ and $x^0 = [x_{env}^0, x_R^0, x_H^0]^T$. a is acceleration and $\dot{\psi}$ is yawrate. In this representation, superscript means time horizon and subscript means the agent. Since the trajectory appears through the internal states of driver, it is assumed that the internal states can be estimated from past history of action $u_H^{-N_h:-1}, u_R^{-N_h:-1}$ where N_h is length of history. In short, inputs and expected outputs of the response prediction problem are as follows:

- Inputs: $u_H^{-N_h:-1}$, $u_R^{-N_h:-1}$, x_R^0, x_H^0, x_{env}^0, $u_R^{0:N_p}$.
- Output: $P(u_H^{0:N_p}|u_R^{0:N_p}, x^0, z)$.

B. Model

In order to consider the sequence, we adopt a Sequence-to-Sequence Variational Auto-encoder [4] in our framework. While this model generally consists of encoder and decoder, proposed model has additional state embedding part considering initial states for robot and human. The encoder takes past histories of robot and human as inputs, and generates latent vector of size N_z. History sequence and reversed history sequence are fed into encoder. As a results, the encoder compresses the historical information into latent vector, z, which follows $\mathcal{N} \sim [\mu, \sigma]$. To summarize this part, latent vector, z, is defined as random vector conditioned on action history.

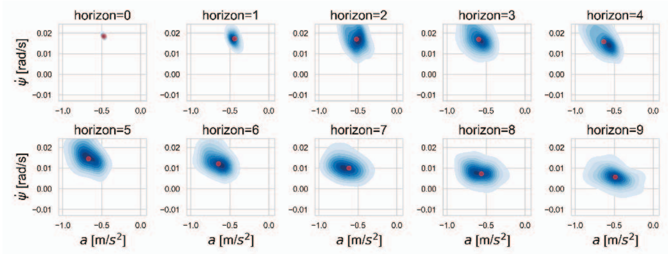


Fig. 2: An example of predicted distribution of human responses. The red dots represent ground truth of human responses for each future time step.

The decoder predicts u_H^k corresponding to u_R^k for every horizon step k. The distribution of u_H^k is modeled by the Gaussian Mixture Model (GMM) with M gaussian distributions $p(u_H^k|u_R^{0:k}, x^0, z) = \sum_{i=1}^{M} \pi_i^k \mathcal{N}(\mu_i^k, \Sigma_i^k)$, where π_i represents the mixture weights of the GMM. Fig. 2 shows an example of predicted response distribution in two seconds.

C. Loss function

The objective of training for this model is to minimize the three kind of loss functions: the Prediction loss, L_p, the Kullback-Leibler Divergence Loss, L_{KL}, and the Weight regularization loss for W_y:

$$L_p = -\frac{1}{N_p} \sum_{k=0}^{N_p} log(\sum_{i=1}^{M} \pi_i \mathcal{N}(u_H^k|\mu_i^k, \Sigma_i^k)) \tag{1}$$

$$L_{KL} = -\frac{1}{2N_z}(1 + \sigma - \mu^2 - \exp(\sigma)) \tag{2}$$

$$L_{l2} = \sum_{i,j} {(w_y)_{ij}}^2 \tag{3}$$

where $(w_y)_{ij}$ is the (i, j) element of matrix W_y.

The total loss function is a weighted sum of losses mentioned above:

$$L_{total} = Lp + w_{KL}L_{KL} + w_{l2}L_{l2} \tag{4}$$

III. Experimental Results

A. Dataset

We use the public Next-Generation Simulation (NGSIM) datasets collected from US Highway 101 [5]. In order to make interacting robot-human pairs from the dataset, we manually extract the vehicles that attempt to change the lane. The vehicle trying to change lane is assumed as robot and the most interacting vehicle in the situation is assumed as human driver.

B. Evaluation Metric and Baselines

In order to evaluate the future response prediction problem, predicted control inputs are compared. we evaluate mean and maximum absolute error for acceleration, a, and yawrate, $\dot{\psi}$, at a future point in two seconds.

The following models are selected for comparison:

TABLE I: Quantitative prediction results

	a absolute error $[m/s^2]$		$\dot{\psi}$ absolute error $[rad/s]$	
	mean	max	mean	max
`proposed`	0.32	**2.81**	**0.008**	**0.085**
`LSTM-ED`	0.54	5.12	0.04	0.11
`CTRV`	**0.31**	5.78	0.04	0.21

- Proposed: Proposed model with $M = 10$ Gaussian mixtures and size of latent vector $N_z = 80$.
- LSTM-ED: A LSTM encoder-decoder model directly predicting human inputs using initial states and planned robot inputs.
- CTRA: A constant turn rate and acceleration model.

C. Results

Table I shows the prediction errors of proposed method and baseline models. For acceleration error, CTRA model has best performance in terms of mean absolute error. This is because the number of data with large acceleration change within 2 seconds is relatively small considering the natural characteristic of vehicles driving on highway. However, CTRA model has poor performance in terms of maximum acceleration error. It seems that CTRA model cause large acceleration errors in the situation where deceleration is required through interaction. For yawrate error, the proposed methods perform better than others in terms of both mean and maximum error. Compared with LSTM-ED, the performance of the proposed method, which uses latent vector, performed better in all aspects.

IV. Conclusion

In this paper, the probabilistic human driver model is proposed to predict the control input response of human driver to the planned control inputs of automated vehicle (robot). The proposed model estimates the latent vector representing drivers internal states first and then predicts the future control input response of human driver with uncertainty by using the latent vector. The uncertainty is modeled as a Gaussian mixture model with mixture density network. The simulation results demonstrate that the proposed method predicts the future response of human driver accurately in terms of acceleration and yawrate errors.

References

[1] H. Kim, J. Cho, D. Kim, and K. Huh, "Intervention minimized semi-autonomous control using decoupled model predictive control," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 618–623.

[2] D. Kim, H. Kim, and K. Huh, "Trajectory planning for autonomous highway driving using the adaptive potential field," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 1069–1074.

[3] D. Ferguson, T. M. Howard, and M. Likhachev, "Motion planning in urban environments," *Journal of Field Robotics*, vol. 25, no. 11-12, pp. 939–960, 2008.

[4] D. Ha and D. Eck, "A neural representation of sketch drawings," *arXiv preprint arXiv:1704.03477*, 2017.

[5] J. Colyar and J. Halkias, "Us highway 101 dataset," *Federal Highway Administration (FHWA), Tech. Rep. FHWA-HRT-07-030*, 2007.

Image Classification by Multilayer Feature Extraction Based on Nuclear Norm Minimization

Tomoya Hirakawa[†], Shohei Kubota, and Yoshimitsu Kuroki[††]
National Institute of Technology(KOSEN), Kurume College, Fukuoka, Japan
[†]a3519th@kurume.kosen-ac.jp, [††]kuroki@kurume-nct.ac.jp

***Abstract*—This paper describes a novel multilayer image classification method based on a nuclear norm of dictionary atoms minimization; in each layer, the method designs dictionary atoms for training dataset by solving a nuclear norm regularized convex optimization problem. The problem promotes non-orthogonality of dictionary atoms and makes coefficient vectors non-negative; these contribute to the improvement of classification accuracies. Experiments result shows that our method is superior to the conventional KLT (Karhunen-Loéve Transform) based multilayer classification method in classification accuracy.**

***Index Terms*—image classification, nuclear norm, dictionary learning**

I. Introduction

CNNs (Convolutional Neural Networks) have great progress in not only image recognition but also various fields [1]; in image classification tasks, a multilayer structure which consists of linear convolutional operations and nonlinear activate functions such as the Rectified Linear Unit (ReLU) are attributable to high classification accuracy. Kuo et. al proposes a concept of the REctified-Correlations On a Sphere (RECOS) transform [2] to interpret the CNNs structure by projection and rectification of vectors, and the multilayer Saak (subspace approximation with augmented kernels) transform [3], which retrieves negative information eliminated by the ReLU function. Instead of the ReLU function, this transform adapts the Sign-to-Position (S/P) conversion which expresses a real number as a two dimensional positive vector: if the number is positive, it appears in the first element, else its magnitude appears in the second element. Additionally, the Saak transform designs atoms in each layer by the KLT (Karhunen-Loéve Transform), and the atoms cross orthogonally each other. We focus on designing dictionary atoms and formulate nuclear norm regularized convex optimization problems in each layer. By this approach, the nuclear-norm regularization of a dictionary promotes non-orthogonality of the dictionary atoms. We expect this scheme contributes to higher classification accuracy.

II. Preliminaries

In this paper, we formulate a nuclear norm regularized convex problem; a nuclear norm of a matrix $W \in \mathbb{R}^{K\times L}$ is expressed as $\|W\|_* = \sum_{i=1}^{N}\sigma_i$, where $\sigma_1, \sigma_2, \cdots, \sigma_N$ are singular values of the Singular Value Decomposition (SVD) of $W = U\mathrm{diag}(\sigma_1, ..., \sigma_N)^t V$. Both matrices $U \in \mathbb{R}^{K\times N}$ and $V \in \mathbb{R}^{L\times N}$ are orthogonal. Then, this norm regularizes the rank of the matrix W, and we can obtain its proximal operator in a easy way: the SVD of W and the element-wise shrinkage for each σ_i $(i = 1, ..., N)$.

An indicator function $\imath_A(\boldsymbol{a})$ expresses whether an input vector $\boldsymbol{a}$ is in a set A: if $\boldsymbol{a} \in A$, $\imath_A(\boldsymbol{a}) = 0$, else $\imath_A(\boldsymbol{a}) = \infty$. The proximal operator of an indicator function is a convex projection if the set A is convex. The non-negativity indicator function defined as

$$\imath_{[0,\infty)}(X_{i,j}) = \begin{cases} 0 & (X_{i,j} \geq 0) \\ \infty & (X_{i,j} \leq 0) \end{cases}, \tag{1}$$

which is applied to each layer's optimization problem instead of the ReLU function. The ℓ_2-norm-ball indicator function is used to normalize each dictionary atoms and is defined as

$$\imath_{\|\cdot\|_2 \leq r}(\boldsymbol{a}) = \begin{cases} 0 & (\|\boldsymbol{a}\|_2 \leq r) \\ \infty & (\|\boldsymbol{a}\|_2 > r) \end{cases}. \tag{2}$$

We use ADMM (Alternating Direction Method of Multipliers) [4] to solve the proposed problem; ADMM can solve a convex problem formulated as

$$\arg\min_{\boldsymbol{x},\boldsymbol{z}} f(\boldsymbol{x}) + g(\boldsymbol{z}) \quad \text{s.t.} \quad \boldsymbol{z} = A\boldsymbol{x}, \tag{3}$$

where $\boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{z} \in \mathbb{R}^M$, $A \in \mathbb{R}^{N\times M}$, and both $f : \mathbb{R}^N \to \mathbb{R}$ and $g : \mathbb{R}^M \to \mathbb{R}$ are convex functions. ADMM introduces an augmented Lagrangian function:

$$L_{aug} = f(\boldsymbol{x}) + g(\boldsymbol{z}) + <\boldsymbol{y}, A\boldsymbol{x} - \boldsymbol{z}> + \frac{\rho}{2}\|L\boldsymbol{x} - \boldsymbol{z}\|_2^2, \tag{4}$$

where $\boldsymbol{y}$ is the Lagrange Multiplier vector, and solve $\min L_{aug}$ with respect to each variable alternatively.

The conventional single layer Saak transform firstly divides an input image into non-overlapping patches and obtains dictionary atoms by conducting the KLT to the patches. Secondly, the input patches are projected onto designed atoms, and the layer obtains coefficient vectors. At last, the S/P conversion is applied to the coefficient vectors, and the output is transferred to the next layer as its input. Figure 1 illustrates the first stage Saak transform.

III. Proposed Method

The Saak Transform obtains only orthogonal dictionary atoms, and we consider this is too strong constraint in feature extraction. The proposed method designs non-orthogonal dictionary atoms flexibly by regularizing a nuclear norm of a dictionary atom matrix to improve the classification accuracy.

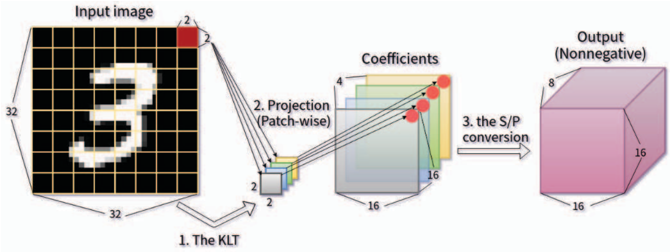


Fig. 1. The first stage Saak transform: the output in the figure is the input of the second stage Saak transform.

The proposed method formulates each layer's optimization problem as

$$\arg\min_{X,W} \|Y - WX\|_F^2 + \imath_{[0,\infty)}(X) + \imath_{\|\cdot\|_2 \leq 1}(W) + \lambda\|W\|_*, \tag{5}$$

where $Y \in \mathbb{R}^{M\times N}$, $W \in \mathbb{R}^{M\times K}$, and $X \in \mathbb{R}^{K\times N}$ are an input matrix, a dictionary atom matrix and a coefficient matrix; each column of the matrices expresses an input vector, a dictionary atom, and a coefficient vector respectively; $\|\cdot\|_F$ means the Frobenius norm, and the parameter λ controls the trade off between the data fidelity term and the regularization term. It is difficult to optimize the problem (5) for two variables X and W at the same time, then our method divides the problem into two sub-problems formulated as,

$$\arg\min_{X} \|Y - WX\|_F^2 + \imath_{[0,\infty)}(X), \text{ and} \tag{6}$$

$$\arg\min_{W} \|Y - WX\|_F^2 + \imath_{\|\cdot\|_2 \leq 1}(W) + \lambda\|W\|_*. \tag{7}$$

Both problem (6) and (7) are convex problems which can be solved by ADMM, and our method optimizes these two problems alternatively. Figure 2 illustrates the first stage transform of the proposed method.

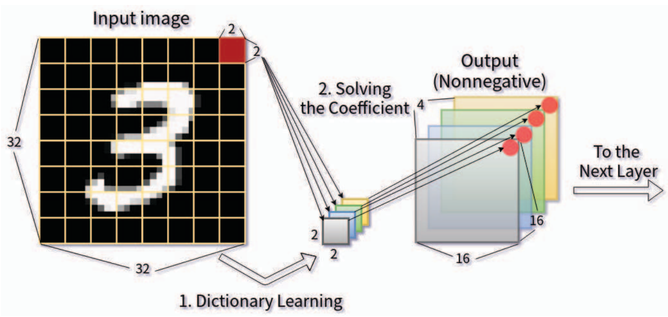


Fig. 2. The first stage of proposed method

IV. Experiments

To evaluate our method and the conventional one, we compare classification accuracy on MNIST handwritten image dataset [5] which consists of 10 classes and 32 × 32 arrays quantized to 8 bit gray scale. We use 600 images for each digit for the training, and 1,000 images for the test. Both methods employ the Support Vector Machine (SVM) as the classifier. The training stage has three steps; (i) design of the dictionary atoms by solving problem (6) and (7) alternatively, (ii) selection of coefficients of whole layers used in classification stage based on following f value defined as

$$f = \frac{\text{between-group variability (BGV)}}{\text{within-group variability (WGB)}}, \tag{8}$$

and (iii) training the SVM.

The test stage firstly obtains coefficients by solving the problem (6) for test images, and selects coefficients by the f value. Next, the SVM classifies the input using selected coefficient vectors. In all experiments, we set the five layer model, which have 3, 4, 7, 6, and 8 dictionary atoms in each layer respectively. The spatial size of patches is 2 × 2 pixels. The number of coefficients selected by the f value is set to 32, 64, 96, 128, 160, 192, 224, 256, 288, 320, 352, 384, 416, 448, 480, and 512.

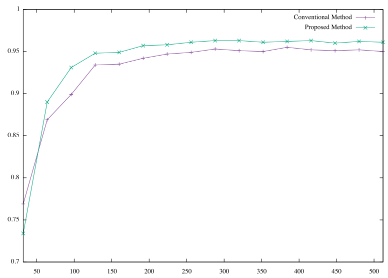


Fig. 3. Comparison of classification accuracy: the horizontal and vertical axes mean the number of selected coefficients and the the classification accuracy respectively.

Experimental result in Fig.3 shows our method can classify more correctly than the conventional method in all situations except for at the number of selected coefficients is 32.

V. Conclusion

We propose a novel multilayer image classification method which regularizes the nuclear norm of a dictionary atom matrix to promote non-orthogonality between the atoms. The experimental result shows our method improves the classification accuracy on MNIST dataset. Future tasks are validation on other datasets, comparison other regularization method, and the use of other classifiers.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 5 2015, doi: 10.1038/nature14539.

[2] C. . J. Kuo, "The cnn as a guided multilayer recos transform [lecture notes]," *IEEE Signal Processing Magazine*, vol. 34, no. 3, pp. 81–89, May 2017, doi: 10.1109/MSP.2017.2671158.

[3] C.-C. J. Kuo and Y. Chen, "On data-driven saak transform," *Journal of Visual Communication and Image Representation*, vol. 50, pp. 237 – 246, 2018, doi: https://doi.org/10.1016/j.jvcir.2017.11.023 .

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011, doi: http://dx.doi.org/10.1561/2200000016. [Online]. Available: http://dx.doi.org/10.1561/2200000016

[5] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist/

Embedded Implementation of Human Detection Using Only Color Features on the NVIDIA Xavier

Masahiko Tsuyama†, Takuro Oki†, Shingo Kobayashi†, Risako Aoki†,
Ryusuke Miyamoto‡, Hiroyuki Yomo§, and Shinsuke Hara*
†*Dept. of Computer Science, Graduate School of Science and Technology, Meiji University, Kawasaki, Japan*
‡*Dept. of Computer Science, School of Science and Technology, Meiji University, Kawasaki, Japan*
§*Dept. of Electrial and Electronic Engineering, Faculty of Engineering Science, Kansai University, Suita, Japan*
**Dept. of Electric and Information Engineering, Graduate School of Engineering, Osaka City University, Osaka, Japan*
Email: {tsuya, o_tkr, sin, aori, miya}@cs.meiji.ac.jp, yomo@kansai-u.ac.jp, hara@info.eng.osaka-cu.ac.jp

***Abstract*—The authors are developing a novel vital sensing system for real-time exercises using image-assisted routing (IAR), which requires the accurate detection of humans wearing sensor nodes in aerial images captured by a camera mounted on an UAV for sensor localization. To achieve real-time human detection for IAR, this paper proposes the embedded implementation of an accurate human detection scheme, which is constructed with informed filters using only color features and can achieve higher accuracy than deep learning for sports scenes.The experimental results for the actual aerial images of a sports scene revealed that our implementation on the NVIDIA Jetson Xavier board could process a 3840×2160 image in approximately 44.93 ms.**

***Index Terms*—Human detection, edge computing, informed-filters, NVIDIA Jetson Xavier**

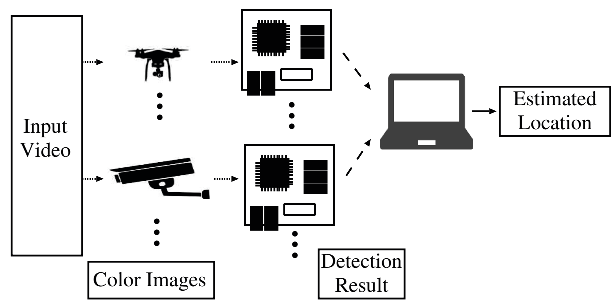


Fig. 1. Overview of system used in IAR.

I. Introduction

The authors are trying to build a novel vital sensing system that enables real-time sensing during exercises using multi-hop sensor networking [1]. To this end, a novel routing scheme based on vision-based sensor localization, namely, image-assisted routing (IAR), is proposed. This is an improvement over existing routing schemes using the RSSI or GPS, which may not be suitable for real-time sensing during exercises because the density and speed of the sensor nodes are excessively high in some cases. For sensor localization, IAR uses the aerial images captured by UAVs. Because the sensor nodes are worn by humans, the IAR estimates the humans' location in each frame to localize the sensors.

The IAR consists of visual human detection and tracking: the detection process finds humans in the captured images, and the tracking process resolves the identification problem over multiple images. Our previous study demonstrated that a simple tracking scheme can provide practical results for sports scenes [2]. Fig. 1 shows an overview of the localization system used in IAR. As can be seen, fixed cameras may occasionally be used in addition to cameras mounted on UAVs.

Therefore, one of the most significant problems to be solved is the accurate and fast computation of human detection, which is carried out on edge-computers to reduce the amount of data transmitted through wireless networks.

To solve this problem, this paper proposes the embedded implementation of a human detection scheme on the NVIDIA Jetson Xavier board using only color features selected by informed-filters during the training process. This approach has better accuracy compared with recently proposed schemes using deep learning to process sports scenes.

II. Implementation

Some of the authors have proposed the parallel implementation [3] of a human detection scheme [4], which uses a classifier comprising weak-classifiers trained with informed filters using only color features. The objective of the previous implementation was to measure the processing speed using a large-scale graphics processing unit (GPU) on a desktop personal computer (PC). Therefore, the previous implementation could not accept input images obtained at run-time.

To realize run-time human detection on the NVIDIA Xavier board, the GPU-based implementation was ported to a board and camera interface that accepts image sequences at run-time. Tab. I shows the specification of the Jetson AGX Xavier processor [5]. The implemented software executes the following operations:

1) Transfer of input image from the camera.
2) Exhaustive search based on sliding windows.
3) Classification of extracted sub-windows using the trained classifier.n
4) Region-merging for multiple results for the detection target by NMS [6].
5) Generating a list of positive samples for the input image.
6) Iterating from (1) to execute the next frame.

TABLE I
SPECIFICATION [5]

Board	Jetson AGX Xavier
GPU	512-core Volta GPU with Tensor Cores
CPU	8-core ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3
Memory	16GB 256-Bit LPDDR4x — 137GB/s
Storage	32GB eMMC 5.1

Fig. 2. Detection result.

III. Evaluation

This section discusses the evaluation of the human detector implemented on the Xavier board.

A. Training and detection

The classifier used in the experiment was constructed in the same way as our previous work [4]. The depth of the weak classifier was one, the number of weak classifiers was two hundred, and the final classifier was constructed using Adaboost.

Fig. 2 shows an example image of the detection results, where the red rectangles indicate positive detections. As can see, there was no false detection and no detection misses in this frame. Fig. 3 shows the DET curve obtained by the classifier trained with informed-filters using only color features. In this figure, the left and bottom have high accuracy, which indicates that the detection accuracy for the actual images is also quite high.

These results reveal that, in the actual scenes, the humans on a field could be accurately detected by the classifier, whose quantitative performance was assessed only with the 3D CG dataset [7]. Thus, it was found that the detection accuracy improved compared with recently proposed schemes based on deep learning.

B. Processing speed

To evaluate the processing speed of the proposed implementation, the time usage per image was measured with the Actual Image Dataset, which has a resolution of 3840×2160.

The time usage per image by the medians was 44.93 ms for one thousand frames. This indicates that, for the actual image, the processing speed is sufficient to operate the system in real-time.

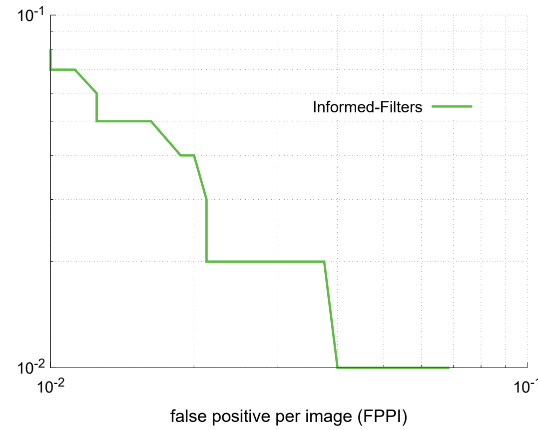


Fig. 3. Det Curve.

IV. Conclusion

This study evaluated the processing speed of human detection with an NVIDIA Xavier board by implementing a classifier trained with informed-filters using only color features. The experimental results showed that the implemented software could accurately detect the target humans captured from UAV. The processing speed was reached approximately 23 frames per second even though the image size was 3840×2160.

Acknowledgments

The research results have been partly achieved by "Research and development of Innovative Network Technologies to Create the Future", the Commissioned Research of National Institute of Information and Communications Technology (NICT), JAPAN.

References

[1] S. Hara, H. Yomo, R. Miyamoto, Y. Kawamoto, H. Okuhata, T. Kawabata, and H. Nakamura, "Challenges in Real-Time Vital Signs Monitoring for Persons during Exercises," *International Journal of Wireless Information Networks*, vol. 24, pp. 91–108, 2017.

[2] H. Yokokawa, T. Oki, and R. Miyamoto, "Feasibility study of a simple tracking scheme for multiple objects based on target motions," in *Proc. of International Workshop on Smart Info-Media Systems in Asia*, 2017, pp. 293–298.

[3] T. Oki and R. Miyamoto, "Efficient GPU implementation of informed-filters for fast computation," in *Image and Video Technology*, 2017, pp. 302–313.

[4] R. Miyamoto and T. Oki, "Soccer player detection with only color features selected using informed haar-like features," in *Proc. Advanced Concepts for Intelligent Vision Systems*, 2016, pp. 238–249.

[5] NVIDIA, "AI Platform for Autonomous Machines — NVIDIA Jetson AGX Xavier," https://www.nvidia.com/en-us/autonomous-machines/jetson-agx-xavier/, 2018.

[6] R. Miyamoto, S. Kobayashi, T. Oki, H. Yomo, and S. Hara, "Improved pairwise max suppression considering total number of targets," in *IEEE Proc. of International Conference on Systems, Man, and Cybernetics, SMC 2018*, 2018, pp. 2091–2095.

[7] R. Miyamoto, H. Yokokawa, T. Oki, H. Yomo, and S. Hara, "Human detection in top-view images using only color features," *The Journal of the Institute of Image Electronics Engineers of Japan*, vol. 46, no. 4, pp. 559–567, 2017.

2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)

Irreversible Privacy-Preserving Images Holding Spatial Information for HOG Feature Extraction

1st Masaki Kitayama
Tokyo Metropolitan University
Tokyo, Japan
kitayama-masaki@ed.tmu.ac.jp

2nd Hitoshi Kiya
Tokyo Metropolitan University
Tokyo, Japan
kiya@tmu.ac.jp

***Abstract*—In this paper, we propose a generation method of visually protected images and its application to privacy-preserving machine learning. Images generated by the proposed method hold the gradient direction information of the original images, but have no the information. Histogram-of-Oriented-Gradients (HOG) features are extracted from the protected images, and the features are applied to machine learning algorithms. In addition, the proposed generation method is an irreversible one, so there is no need to manage secret keys, unlike encryption methods. In an experiment, a face classification task is carried out under the use of a support vector machine algorithm with the HOG features to demonstrate the effectiveness of the proposed method.**

I. Introduction

In recent years, cloud computing has been rapidly spreading in many fields. However, cloud environments are generally semi-trusted, so there are some security concerns such as unauthorized use of data and privacy compromise. To solve the security concerns, machine learning with encrypted data has been researched [1]–[3].

In this paper, we propose a generation method of visually protected images (referred to as "protected images") which hold the spatial information of images. Moreover, we propose an extraction method of Histogram-of-Oriented-Gradients (HOG) [4] features from the protected images for machine learning. The generation of protected images is performed by generating random pixels under certain restrictions, and is irreversible. Therefore, the proposed method has no need to manage secret keys. Furthermore, since the protected images retains the spatial information of the original image, it can be applied to not only simple image recognition but also object detection. In an experiment, image recognition with a support vector machine algorithm is carried out to confirm the effectiveness of the proposed method.

II. Proposed Method

A. Overview of Proposed Method

Figure 1 shows a privacy-preserving image recognition system considered in this paper. In both training and testing phases, each user generates protected images in the user's local before sending the images to a cloud server.

978-1-7281-3038-5/19/$31.00 2019 IEEE

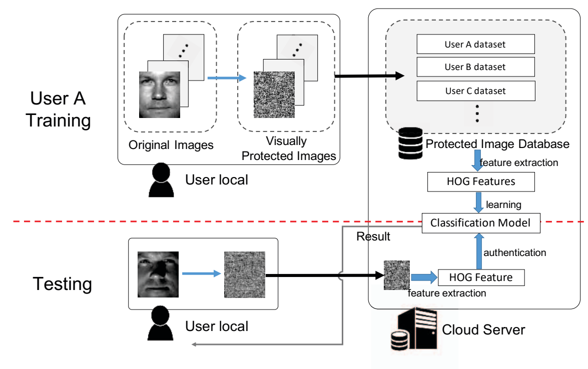


Fig. 1. Privacy-preserving image recognition system.

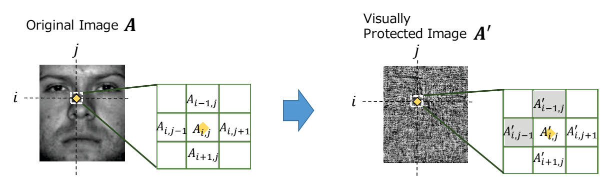


Fig. 2. Relation between $\boldsymbol{A}$ and $\boldsymbol{A}'_{j_s}$. Two pixels $A'_{i-1,j}$ and $A'_{i,j-1}$ were calculated prior to $A'_{i,j+1}$ and $A'_{i+1,j}$.

Then the cloud server carries out an image recognition algorithm with HOG features extracted from the protected images.

B. Generation of Protected Image

1) Restriction of The Protected Image: Now, let $\boldsymbol{A} \in \mathbb{R}^{I\times J}$ and $\boldsymbol{A}'_{j_s} \in \mathbb{R}^{I\times J}$ be an original image and the protected image, and let $A_{i,j}$ and $A'_{i,j}$ be pixel values at a position (i,j) of $\boldsymbol{A}$ and $\boldsymbol{A}'_{j_s}$, $i,j \in \mathbb{Z}$ respectively.

In Fig.2, we focus on a position (i,j) to explain the restriction between two images, $\boldsymbol{A}$ and $\boldsymbol{A}'_{j_s}$. At the position (i,j), the gradient direction $\theta_{i,j}$ is defined for $\boldsymbol{A}$ as

$$\theta_{i,j} = \tan^{-1}(y_{i,j}/x_{i,j}) \quad , \tag{1}$$

where $x_{i,j} = A_{i,j+1} - A_{i,j-1}$ and $y_{i,j} = A_{i+1,j} - A_{i-1,j}$.

Similarly as $\theta_{i,j}$, the gradient direction $\theta'_{i,j}$ is defined for image $\boldsymbol{A}'_{j_s}$ as

$$\theta'_{i,j} = \tan^{-1}(y'_{i,j}/x'_{i,j}) \quad , \tag{2}$$

where $x'_{i,j} = A'_{i,j+1} - A'_{i,j-1}$ and $y'_{i,j} = A'_{i+1,j} - A'_{i-1,j}$. If the relation

$$\theta'_{i,j} = \theta_{i,j} \quad , \tag{3}$$

is satisfied, $\boldsymbol{A}'_{j_s}$ has the same gradient direction as $\boldsymbol{A}$ at the position (i,j). In this paper, $\boldsymbol{A}'_{j_s}$ is designed under

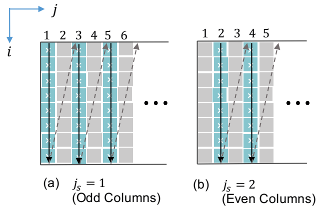


Fig. 3. Attention order of a position (i, j) in $\boldsymbol{A}'_{j_s}$

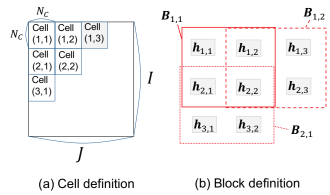


Fig. 4. Cell and block definition

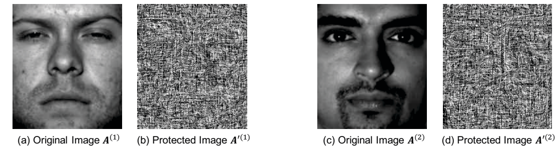


Fig. 5. Original images and their protected images.

Eq.(3).

2) Generation of $\boldsymbol{A}'_{j_s}$: In Fig.3, a position (i, j) with $\theta'_{i,j}$ in $\boldsymbol{A}'_{j_s}$ is illustrated. There are attention positions every two columns, where the initial value of the attention positions is decided by the parameter $j_s \in \{1, 2\}$, and the attention position moves in the order of the arrow. At a position (i, j), four pixel values: $A'_{i-1,j}, A'_{i,j-1}, A'_{i,j+1}$, and $A'_{i+1,j}$ have to be decided under the condition of Eq.(3), where two pixels $A'_{i,j+1}$ and $A'_{i+1,j}$ are the pixels randomly generated under the condition of Eq.(3). The remaining two pixels $A'_{i-1,j}$ and $A'_{i,j-1}$ were generated in the previous processing.

C. HOG Feature Extraction

Next, we propose a method of extracting HOG features from a protected image $\boldsymbol{A}'_{j_s}$ as follows.

step-1 **gradient direction map** : The gradient direction map $\boldsymbol{\theta}' \in \mathbb{R}^{I \times J}$ is calculated in accordance with Eq.(2).

step-2 **histogram voting map** : The histogram voting map $\boldsymbol{M}' \in \mathbb{R}^{I \times J}$ is generated by using j_s as

$$M'_{i,j} = \begin{cases} 0 & (j \in \{js + 2m \mid m \in \mathbb{Z}\}) \\ 1 & (\text{the others}) \end{cases}, \tag{4}$$

where $M'_{i,j}$ is a pixel value of $\boldsymbol{M}'$ at a position (i, j).

step-3 **histograms of gradient direction** : As shown in Fig.4 (a), maps $\boldsymbol{\theta}'$ and $\boldsymbol{M}'$ are commonly divided into small grids called "cells" with $N_C \times N_C$ pixels, and then, $\theta'_{i,j}$ is quantized and its histogram, $\boldsymbol{h}_{p,l}$ is made up per each cell, where (p, l) is a index of the histogram. The quantization level of $\theta'_{i,j}$ is b, and the votes are weighted by $M'_{i,j}$. In this paper, $N_C = 8$ and $b = 9$ are chosen as parameters.

step-4 **block normalization** : Let us define "blocks" as the concatenation of 2×2 cell histograms, allowing overlapping of the middle cells (see Fig.4 (b)). Each block is normalized by the L2 norm, and the HOG feature of $\boldsymbol{A}'_{j_s}$ is the vector produced by concatenating all blocks.

The difference between the proposed HOG feature extraction and the conventional one [4] is that the proposed one uses the histogram voting map $\boldsymbol{M}'$ instead of the gradient strength map $\boldsymbol{G}$ as the weight for voting $\boldsymbol{h}_{p,l}$ in *step-3*.

III. Experiment

A face recognition experiment was carried out using the Extended Yale Face Database B with a support vector machine (SVM) algorithm. This dataset contains 38 individuals and 64 frontal facial images with 168×192 pixels per each person. The images for each person were divided into 16 for training and 48 for testing. Fig.5 shows two samples from the dataset and the protected images.

TABLE I
Face recognition performances with SVM

feature set	EER
Set-1: conventional with non-protected HOG	0.0033
Set-2: proposed with protected HOG	0.0049
Set-3: conventional with non-protected Eigen Face	0.0742

TABLE I shows the experimental result. To evaluate the effectiveness, equal error rate (EER), which is the point at which false reject rate (FRR) is equal to false accept rate (FAR), was used. EER is acquired by changing the threshold of classiication score. *Set-1* is a set of HOG features extracted with the conventional method [4], and *Set-2* is a set extracted with the proposed one. *Set-3* is "Eigen Face" [5] features (150 dimension) which is an representative one for face recognition.

TABLE I indicates that the proposed method (*Set-2*) has a good performance than *Set-3*. The difference of EER values between *Set-1* and *Set-2* is caused due to no information on gradient strength.

IV. Conclusion

We proposed a generation method of visually protected images, which allows us to extract HOG features from the protected ones without any keys, while maintaining a reasonable performance.

References

[1] Mauro Barni, Giulia Droandi, and Riccardo Lazzeretti, "Privacy protection in biometric-based recognition systems: A marriage between cryptography and signal processing," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 66–76, 2015.

[2] Takahiro Maekawa, Ayana Kawamura, Takayuki Nakachi, and Hitoshi Kiya, "Privacy-preserving support vector machine computing using random unitary transformation," *arXiv preprint arXiv:1908.07915*, 2019.

[3] W. Sirichotedumrong, T. Maekawa, Y. Kinoshita, and H. Kiya, "Privacy-preserving deep neural networks with pixel-based image encryption considering data augmentation in the encrypted domain," in *2019 IEEE International Conference on Image Processing (ICIP)*, Sep. 2019, pp. 674–678.

[4] Navneet Dalal and Bill Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE International conference on Computer Vision and Pattern Recognition*, 2005, pp. 886–893.

[5] Jun Zhang, Yong Yan, and Martin Lades, "Face recognition: eigenface, elastic matching, and neural nets," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1423–1435, 1997.

978-1-7281-3038-5/19/$31.00 2019 IEEE

XOR-ed Based Friendly-Progressive Secret Sharing

Heri Prasetyo
Department of Informatics
Universitas Sebelas Maret
Surakarta, Indonesia
heri.prasetyo@staff.uns.ac.id

Joni Welman Simatupang
Study Program of Electrical Engineering
President University
Bekasi, Indonesia
joniwsmtp@president.ac.id

***Abstract*— This paper presents a simple approach for secret sharing under the bitwise eXclusive-OR (XOR) framework. The proposed method extends the effectiveness of Progressive Secret Sharing (PSS) into the Friendly Secret Sharing (FSS). The PSS produces a set of shared images in noisy-like form, while FSS generates shared images into more-visually-friendly appearance. The proposed method offers lossless ability in the recovery result of secret image. At the same time, it shows superiority compared to the former existing scheme in the secret sharing task.**

Keywords*— *friendly, lossless, progressive, secret sharing

I. INTRODUCTION

Several methods have been developed for secret sharing such as PSS [1], multiple secret sharing [2], lossless PSS [3], etc. Most of them yield promising results on secret sharing systems. This section presents the proposed XOR-ed based FPSS. It extends the usability of PSS [1] into FSS. The proposed method offers lossless ability on secret image reconstruction. Let I be a quantized secret image of size $M \times N$. This image is obtained after performing the scalar quantization with coefficient Q. Suppose that this quantized image is in RGB color space. Each pixel is denoted as $I(x, y, c)$, where $x = 1,2, \dots, M$ and $y = 1,2, \dots, N$ are spatial positions. The symbol c is color channel, i.e. $c = 1,2,3$.

The proposed method generates n shared images denoted as $\{S^1, S^2, \dots, S^n\}$, where S^i is the i-th shared image. The PSS [1] produces a set of shared images in noise-like form. The shared image can be easily recognized by investigating its content. The proposed method overcomes this problem by rendering the secret image into the cover image. Let C be a color cover image of size $M \times N$. This image size should be identical to that of I. Each pixel of C is denoted as $C(x, y, c)$.

The computation of shared image generation can be explained as follow. For each pixel on spatial position (x, y, c) with $x = 1,2, \dots, M$, $y = 1,2, \dots, N$, and $c = 1,2,3$, the proposed method firstly computes the masking coefficient R as follow:

$$R \leftarrow U_I(0, \left\lceil \frac{255}{Q} \right\rceil), \tag{1}$$

where $U_I(a, b)$ denotes the random number generator which uniformly produces an integer in range $[a, b]$. The symbols $\lceil \cdot \rceil$ and $\leftarrow$ are ceiling and assignment operators, respectively. The proposed method subsequently determines the indices of two selected shared images, i.e. r_1 and r_2, with constraint $1 \le r_1, r_2 \le n$ and $r_1 \ne r_2$. These two indices are randomly chosen. The first selected shared image, i.e. S^{r_1} is determined as:

$$S^{r_1}(x, y, c) \leftarrow C(x, y, c) \oplus I(x, y, c) \oplus R, \tag{2}$$

where $\oplus$ represents the bitwise-based XOR operation. Whereas, the second selected shared image, i.e. S^{r_2} is computed as:

$$S^{r_2}(x, y, c) \leftarrow C(x, y, c) \oplus R. \tag{3}$$

For the rest of shared images with $i = 1,2, \dots, n$ and $i \ne r_1, r_2$, we perform the following process;

$$S^i(x, y, c) \leftarrow C(x, y, c). \tag{4}$$

At the end of this process, one obtains a set of shared images $\{S^1, S^2, \dots, S^n\}$.

The secret image can be reconstructed by stacking several shared images as follow:

$$\hat{I}(x, y, c) \leftarrow S^{t_1}(x, y, c) \oplus S^{t_2}(x, y, c) \oplus \dots \oplus S^{t_T}(x, y, c), \tag{5}$$

where $\hat{I}(x, y, c)$ is the recovered secret image at spatial position (x, y, c) and $\{t_1, t_2, \dots, t_T\}$ is the index of stacked shared image. The symbol T denotes the number of stacked shared images with condition $T \le n$. The proposed method achieves FSS since the content of S^i is almost similar to C under some extends, while it gives PSS based on the facts that the quality of recovered secret image is improved by stacking more shared images. Thus, the proposed method can be categorized as friendly and progressive secret sharing.

II. ANALYSIS OF PROPOSED XOR-ED BASED FPSS

This section supports the correctness of the proposed method with the theoretical analysis. It considers the lossless ability of the proposed method on recovering secret image. For simplicity, we omit the spatial position of an image. Then, the formal analysis is given as follow.

Theorem 1: *The proposed XOR-ed Based FPSS is lossless if* $T \le n$ *and* $1 \le r_1, r_2 \le T$.

Proof: Stacking several shared images $\{S^{t_1}, S^{t_2}, \dots, S^{t_T}\}$ produces a recovered secret image as:

$$\hat{I} \leftarrow S^{t_1} \oplus S^{t_2} \oplus \dots \oplus S^{t_T}.$$

For $T \le n$ and $1 \le r_1, r_2 \le T$, the value $\hat{I}$ is then obtained as:

$$\hat{I} \leftarrow S^{t_1} \oplus S^{t_2} \oplus \dots \oplus S^{r_1} \oplus \dots \oplus S^{r_2} \oplus \dots \oplus S^{t_T}.$$

This form can be alternatively rewritten as follow:

$$\hat{I} \leftarrow S^{r_1} \oplus S^{r_2} \oplus \underbrace{S^{t_1} \oplus S^{t_2} \oplus \dots \oplus S^{t_T}}_{T-2},$$

$$\hat{I} \leftarrow S^{r_1} \oplus S^{r_2} \oplus \underbrace{C \oplus C \oplus \dots \oplus C}_{T-2}.$$

If $T - 2$ is even number, the XOR property [3] gives the following result:

$$\hat{I} \leftarrow S^{r_1} \oplus S^{r_2} \oplus 0 = S^{r_1} \oplus S^{r_2}. \tag{6}$$

The recovered secret image $\hat{I}$ can be simply obtained by performing XOR operation between S^{r_1} and S^{r_2}.

Since $S^{r_1} \leftarrow C \oplus I \oplus R$ and $S^{r_2} \leftarrow C \oplus R$, the form in (6) can be further rewritten as follow:

$$\hat{I} \leftarrow S^{r_1} \oplus S^{r_2} = C \oplus I \oplus R \oplus C \oplus R,$$

$$\hat{I} \leftarrow C \oplus C \oplus R \oplus R \oplus I.$$

The XOR property [3] simplifies the computation as:

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

$$\hat{I} \leftarrow 0 \oplus 0 \oplus I = I. \qquad (7)$$

Simplification in (7) indicates that the proposed method is lossless, i.e. $\hat{I} = I$. It completes a proof. ■

III. Experimental Results

We report some experiments in this section. Two color images (for cover and secret image) are used in this experiment as shown in Fig. 1. Herein, the quantization coefficient is simply set as $Q = 25$ indicating that each pixel of secret image is represented with four bits. The number of shared images is $n = 10$. Fig. 2 depicts two shared images obtained by the proposed method. As it can be seen, the proposed method effectively generates a set of shared images, in which the contents of shared image are almost identical to that of the cover image.

Fig. 3 displays the performance comparison between the proposed method and extended PSS [1] for FSS in terms of visual investigation on the quality of $\hat{I}$. The number of stacked shared images are set as $T = \{2, 4, 6, 8, 10\}$. It can be deducted from Fig. 3 that the proposed method achieves PSS and FSS criteria for a good secret sharing scheme. The proposed method yields lossless recovered secret image if all shared images are stacked by XOR operation. In addition, the proposed method is superior compared to that of the extended PSS [6] for FSS.

References

[1] H.-C. Chao and T.-Y. Fan, "XOR-based progressive visual secret sharing using generalized random grids," *Displays*, vol. 49, pp. 6-15, 2017.

[2] H. Prasetyo and J. M. Guo, "A note on multiple secret sharing using Chinese remainder theorem and exclusive-OR," *IEEE Access*, vol. 7, no. 1, pp. 37473-37497, 2019.

[3] H. Prasetyo and C. H. Hsia, "Lossless progressive secret sharing for grayscale and color images," *Multimed. Tools App.*, 2019. https://doi.org/10.1007/s11042-019-7710-5.

(a) (b)

Fig. 1. Two testing color images: (a) cover image, and (b) quantized secret image.

(a) (b)

Fig. 2. Two shared images generated by proposed method with $n = 10$: (a-d) $\{S^1, S^2, \dots, S^4\}$.

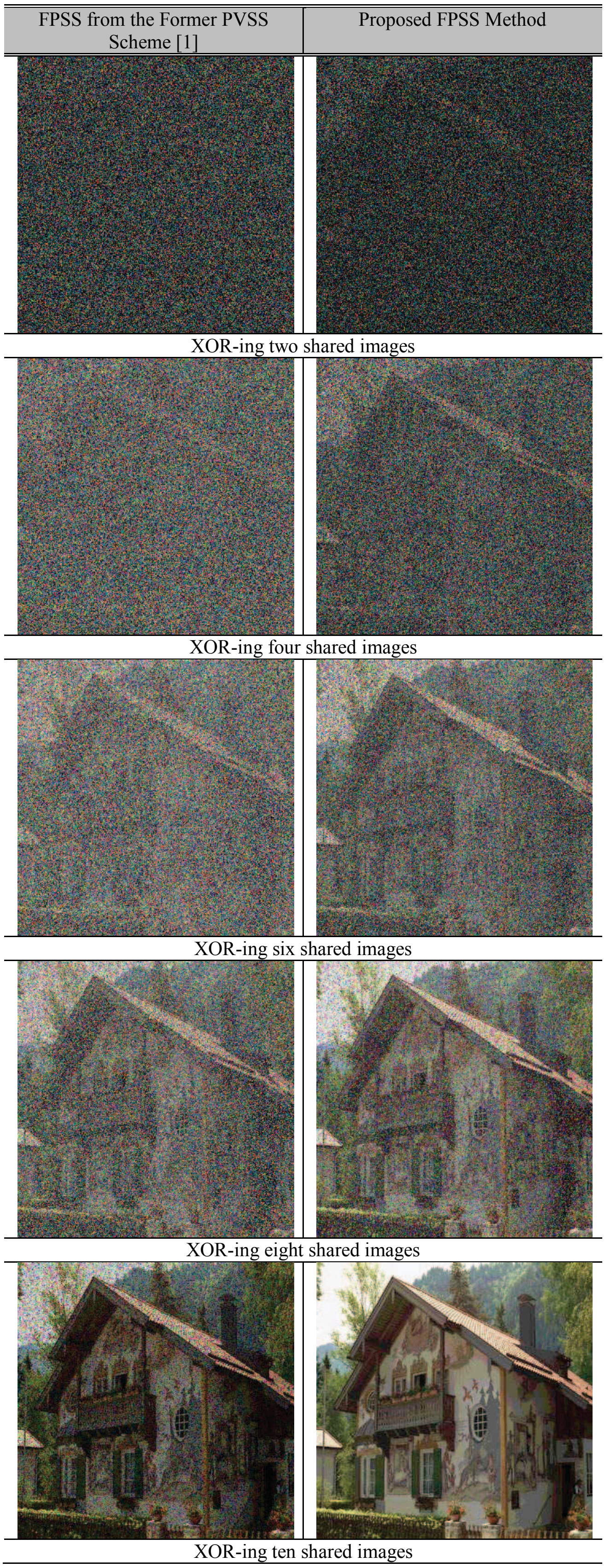


Fig. 3. The reconstruction process of secret image.

Influence of significant target on image quality assessment via EEG

Tianyan Wu, Lihuo He, Hongxia Cai, Wen Lu, Xinbo Gao
School of Electronic Engineering
Xidian University
Xian, China
lhhe@mail.xidian.edu.cn

***Abstract*—The significant target in the image shows how much people pay attention to certain objects in the image when viewing the image. These significant targets will affect human perception of image quality when quality change occurs. But in recent years, most of the image quality perception research based on elsctroencephalography(EEG) were not considered of these significant targets. Therefore, this paper obtains the EEG signals of the subjects through an image rendering experiment for the presence or absence of significant targets, and compares the signal differences. The experimental results show that there is no significant effect on the image quality perception in the perceptible range. If there is a large degree of distortion, whether there is significant target has an important impact on the image quality perception.**

***Index Terms*—Electroencephalography, Significant target, Image quality.**

I. Introduction

Image degradation often occurs during the acquisition process. Therefore, how to make more objective and accurate quality assessment of images has broad application prospects. The significant target in an image refers to the degree to which a person as an observer pays attention to certain objects of the image while viewing the image. These significant goals affect human perception of image quality as quality change occurs. In recent years, image quality assessment research based on electroencephalography (EEG) has become a hot topic because it more fully considers human visual characteristics. However, in most of the research, the significant target in the image is not separately considered.

Lea Lindemann *et al.* [1] used three images of *baboon*, *peppers*, and *flower* as stimulating materials when using EEG to study the quality perception of compressed images. The results of *baboon* and *flower* were similar, only the results of *peppers* is different from the others. Lea Lindemann *et al.* [2] also used the images of three different scenes *ramp*, *cafe*, and *girl* as the stimulation materials in the process of using event-related potential to study video artifact perception. The results showed that there was no significant difference in the EEG signals. Simon Scholler *et al.* [3] used the inconspicuous checkerboard image as a stimulus in order to avoid the influence of image content when they studied video quality.

The paper is structured as follows. The details of the experimental design and implementation is introduced in section 2. The section 3 analyzes the experimental results. Finally, the section 4 describes the conclusions and future work.

II. Experiment

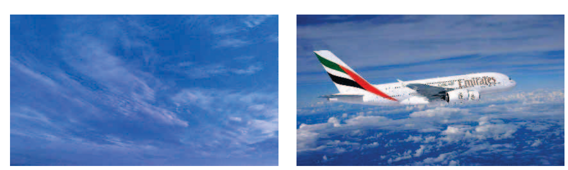
A. Participants

Eighteen subjects (11 males and 7 females, with an average age of 24 years) participated in the experiment. Subjects had normal vision or corrected vision. The participants confirmed the agreement before the experiment.

B. Stimuli

Two 768 $\times$ 432 pixel images *sky* and *airplane* (Fig.1) were used as stimulus. The image *sky* has no significant goal. In the other image *airplane*, the aircraft is a significant target. Six different jpeg compression levels were set for these two images, including one original image and five compressed images with quality parameters of 32, 20, 12, 7, and 4.

The experiment started with a 500ms fixed screen and a red gaze point was displayed in the center of the screen. This was followed by a picture rendering screen with a total duration of 2.5s, first rendering the original image, presenting the original image and then a distorted image at random time points between 1-1.5s. Random time allowed the participants to concentrate and avoid habitual judgment. Next was a judgment screen, the subject should judge whether there was distortion through the left and right keys of the arrow keys on the keyboard.



(a) (b)

Fig. 1. The two test images *sky* and *airplane*.

C. Procedure

Participants should judge whether the image is distorted by pressing the button after each image disappeared. During one experiment, the subject only viewed one of the two images. Each level of image is presented 60 times, and each image is presented in 6 levels, so the subject needed to observe

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

720 experimental stimuli. Experimental stimulation sequence is random.

D. Data acquisition

The subjects sat in front of the computer screen 48cm according to the international standard distance [4], the entire picture was visible. EEG signals were recorded from 64 scalp electrodes. In addition, horizontal and vertical EEG were recorded, and the EEG signal of the mastoid was recorded as a reference. The acquired data was high pass filtered at a frequency of 30 Hz to eliminate DC offset.

III. Result

Fig. 2 shows the EEG signals obtained at different distortion levels for the experimental image *sky* (The result of *airplane* is similar). As can be seen from the figure, as the degree of image distortion increases, the peak of the P300 is higher.

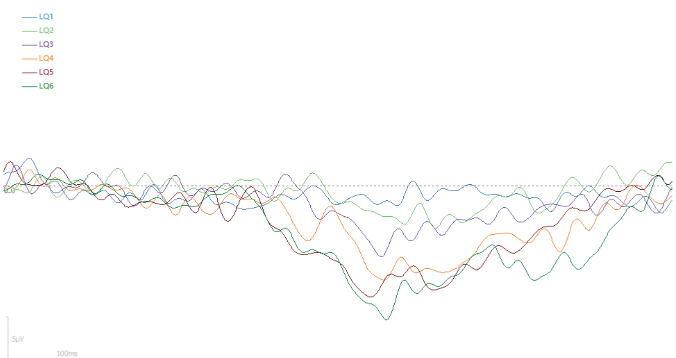
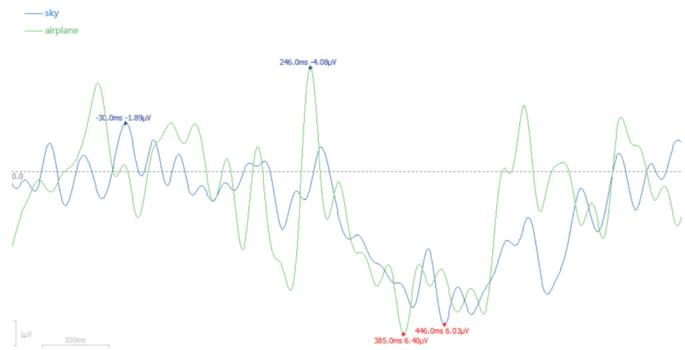


Fig. 2. The starts of the image distortion elicits P300. The higher the distortion level, the higher the peak of P300.

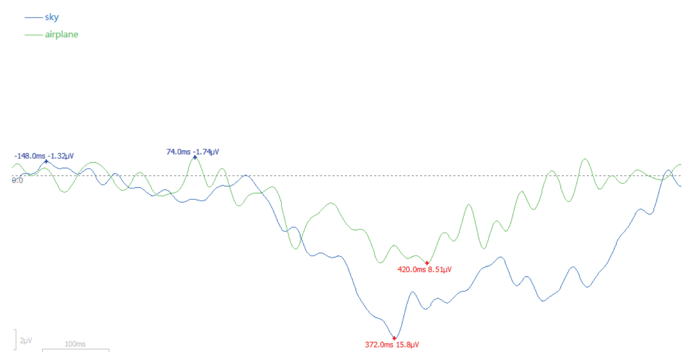
Fig. 3 compares the EEG signals of the two images and analyzes each distortion level. From the perspective of human visual perception, at a low level of distortion (LQ2 and LQ3), it is already apparent that the image is distorted. When the degree of distortion is already large(LQ4, LQ5 and LQ6), it is almost impossible to distinguish the image content. Therefore, for the corresponding EEG signals, we analyze that within the range of perceptible image content, the EEG waveforms obtained by the two images are not much different(Fig.3 (a)), so that whether there is significant target has little influence on the distortion judgment, which is in line with the results of the subjective rating. In the case of a large degree of distortion, the peak value of the EEG signal obtained by the two images has a significant difference(Fig.3(b)), and the peak of the EEG signal obtained by the image with a significant target is smaller than the peak of the EEG signal obtained by the image without the significant target. Therefore, significant targets at this time have an impact on image distortion judgment.

IV. Conclusion

In this paper, a comparative experiment was carried out to study the influence of the presence or absence of significant targets on image quality perception. The results showed that in the perceptible range, whether there was a significant target had little effect on the image quality perception; and in the case of a large degree of image distortion, whether there was the significant target affected the judgment of the image distortion. However, only the influence of jpeg distortion was studied in this paper. Future work considers the same comparative experimental study on other kinds of distortion.



(a)



(b)

Fig. 3. At a low level of distortion (LQ2 and LQ3), the EEG waveforms obtained by the two images are not much different (a). When the degree of distortion is already large(LQ4, LQ5 and LQ6), and the peak of the EEG signal obtained by the image with a significant target is smaller than the peak of the EEG signal obtained by the image without the significant target(b).

References

[1] Lindemann, and Magnor, "Assessing the quality of compressed images using EEG, IEEE International Conference on Image Processing," 2011, pp.3170-3173.

[2] Lindemann, Lea , S. Wenger , and M. A. Magnor, "Evaluation of video artifact perception using event-related potentials," Proc. ACM Applied Perception in Computer Graphics and Visualization, 2011, pp. 53-58.

[3] Scholler S , Bosse S , Treder M S , et al, "Toward a Direct Measure of Video Quality Perception Using EEG," IEEE Transactions on Image Processing, vol.21, no.5, pp.2619-2629, May 2012.

[4] Methodology for the Subjective Assessment of the Quality of Television Pictures, Rec. ITU-R BT.500-11, 2002.

[5] Lixiu Jia, Yan Tu, Lili Wang, Xuefei Zhong, Ying Wang, "Study of image quality using event-related potentials measurement," Journal of Electronic Imaging, vol.27, no.3, pp.33-46, June 2018.

[6] S. Arndt, J.-N. Antons, R. Schleicher, S. Mller, S. Scholler, G. Curio, "A Physiological Approach to Determine Video Quality," IEEE International Symposium on Multimedia, 2011, pp.518-523.

[7] Tcheslavski G V , Vasefi M , Gonen F F, "Response of a human visual system to continuous color variation: An EEG-based approach," Biomedical Signal Processing and Control, vol.43, pp.130-137, March 2018.

Designing High-Performance Green Filters Using Downsampling Techniques

David Shiung
Graduate Institute of Communications Engineering,
National Changhua University of Education (NCUE),
Changhua, 500, Taiwan
davids@cc.ncue.edu.tw

Wen-Long Chin
Department of Engineering Science,
National Cheng Kung University (NCKU), Tainan, 701, Taiwan

wlchin@mail.ncku.edu.tw

***Abstract*—Reduction of the power consumed by high-performance filters is the ultimate objective of all digital signal processing (DSP) practitioners. However, low power consumption and high-performance are often contradictory design goals. This paper uses a technique involves extending the applicability of the Noble identity in commuting downsamplers with "sparse transfer functions," which can be expressed as a function of z^M, where $M = 2, 3, 4, \ldots$. More precisely, the transfer functions are relaxed so that they can be expressed as a function of z^1. This commuting relation, termed the generalized Noble identity (GNI), substantially reduces the power consumption of a filter. The GNI is applied to elaborately designed lowpass filters (LPFs) and achieves a favorable compromise between contradictory design goals.**

***Index Terms*—green filter, downsampler, generalized Noble identity.**

I. Introduction

Small electric gadgets and office equipment, as well as the networks that connect them, consume a considerable amount of power. Information and communication technology (ICT) produces almost the same amount of greenhouse gas emissions as the aviation industry and the situation is expected to worsen, with billions of new devices added to the Internet by 2020 [1].

Very large scale integration (VLSI) is a process for integrating thousands of transistors onto one chip. It is also the most popular method of manufacturing electric ICT devices. Green VLSI technologies, or power-efficient VLSI technologies, play a key role in reducing worldwide greenhouse gas emissions.

High filtering performance often requires high computational complexity and large chip size. Designing a filter that meets various design targets is challenging, yet highly desirable. In this paper, special attention is given to remodeling high-performance filters of limited computational complexity and small die size for low power consumption. High throughput is not emphasized in this paper because many techniques such as retiming and pipelining can be easily performed toward this objective [2].

II. Proposed Techniques

We first study the factors that impact the power consumption of a VLSI circuitry and then propose the technique of GNI to reduce power consumption.

A. Power Consumption Analysis

In a digital complementary metal-oxide-semiconductor (CMOS) circuit, the power consumption P can be represented as [3]

$$P = CV^2 f, \tag{1}$$

where C is the capacitance of the logic gates, V is the operating voltage, and f is the operating frequency. We focus on reducing the operationg frequency of a digital filter to reduce its power consumption.

B. Noble Identity versus GNI

The Noble identity for downsampling is depicted in Fig. 1, where $H(z)$ denotes the transfer function of an lowpass filter (LPF). In Fig. 1, the transfer function $H(z^M)$ must be a sparse function of z^M, where $M = 2,3,4, \ldots$; otherwise, the identity cannot be applied [4][5]. This hampers its usefulness at reducing system power consumption. However, the constraint regarding "sparse transfer function" can be relaxed to "any transfer function."

For illustration purposes, consider a finite impulse response (FIR) LPF in cascade with a downsampler of decimation factor M. Assuming the transfer function of the LPF is $H(z) = b_0 + b_1 z^{-1} + \cdots + b_K z^{-K}$, with real coefficients, and the filter order is K (thus $b_K \neq 0$), the number of filter taps is N. In other words, $N = K + 1$, and $N = 1,2,3, \ldots$. In general, $H(z)$ is not a function of z^M and cannot be transformed using the Noble identity. However, $H(z)$ can be remodeled using the following steps to generalize the Noble identity.

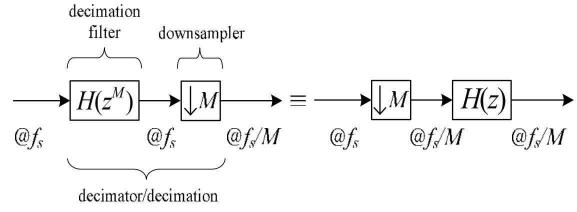


Fig. 1 Multirate Noble identity for downsampling.

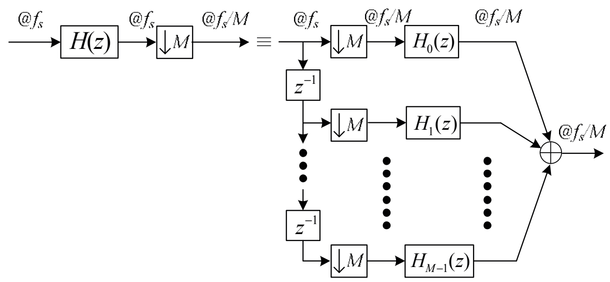


Fig. 2 GNI for downsampling.

Fig. 2 displays the GNI for downsampling and the sample rate at each stage. GNI can be subdivided into three steps.

Step 1: Add L zero-terms to $H(z)$ so that the number of terms in $H(z)$ is a multiple of M. L is an integer that satisfies $0 \leq L < M$. Thus, $H(z)$ contains a total of $K + L + 1$ terms with $(K + L + 1) = JM$, where J is a positive integer.

$H(z)$ can be rewritten as

$$H(z) = b_0 + b_1 z^{-1} + \cdots + b_K z^{-K} + b_{K+1} z^{-(K+1)} + \cdots + b_{K+L} z^{-(K+L)}. \quad (2)$$

The coefficients $b_{K+1}, \ldots, b_{K+L}$ are all zero.

Step 2: Remodel $H(z)$ as

$$H(z) = H_0(z^M) + z^{-1} H_1(z^M) + \cdots + z^{-(M-1)} H_{M-1}(z^M) \quad (3a)$$

where

$$H_i(z^M) = b_i + b_{M+i} z^{-M} + \cdots + b_{(J-1)M+i} z^{-(J-1)M},$$

with i=0,…,M-1. (3b)

Step 3: Apply the Noble identity to (3b) and obtain $H_i(z)$ for $i = 0, \ldots, M-1$. Next, replace $b_{K+1}, \ldots, b_{K+L}$ with zeros to simplify $H_i(z)$. This produces

$$H_i(z) = \begin{cases} b_i + b_{M+i} z^{-1} + \cdots + b_{(J-1)M+i} z^{-(J-1)}, 0 \leq i < L-1, \\ b_i + b_{M+i} z^{-1} + \cdots + b_{(J-2)M+i} z^{-(J-2)}, L \leq i < M. \end{cases}$$

(4)

III. Estimated Power Consumption of Three Different Architectures and Discussions

Assume that the power consumption of an adder operating at sample rate f_s is P_a and that of a multiplier is P_m. The power consumed by the multipliers and adders dominates the entire filter [2]. An estimate of the filter power consumption can therefore be quickly obtained by evaluating the number of multipliers and adders in the filter and the sample rate with which they operate [2][4].

A 5-tap decimation filter with symmetric filter coefficients is depicted in Fig. 3(a). The filter transfer function is $H(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} + b_1 z^{-3} + b_0 z^{-4}$, where the filter coefficients b_0, b_1, and b_2 are real numbers and the decimation factor is $M = 2$. Fig. 3(b) and 3(c) present the other two structures for implementing the decimator. We use the notation $P_{org,N,M}$ to denote the power consumption of an original N-tap decimation filter of decimation factor M. The power consumption in this paradigm is $P_{org,5,2} = 5P_m + 4P_a$.

Applying the GNI to Fig. 3(a) obtains the modified structure in Fig. 3(b). In this paradigm, the power consumption is $P_{gni,5,2} = \frac{5}{2} P_m + 2P_a$. A popular structure for reducing the filter power and hardware complexity is called the folded FIR [4]. The power consumption for this paradigm, denoted as $P_{fld,N,M}$, is $P_{fld,5,2} = 3P_m + 4P_a$. Clearly, the GNI-derived structure is the most power efficient when compared with the other two structures.

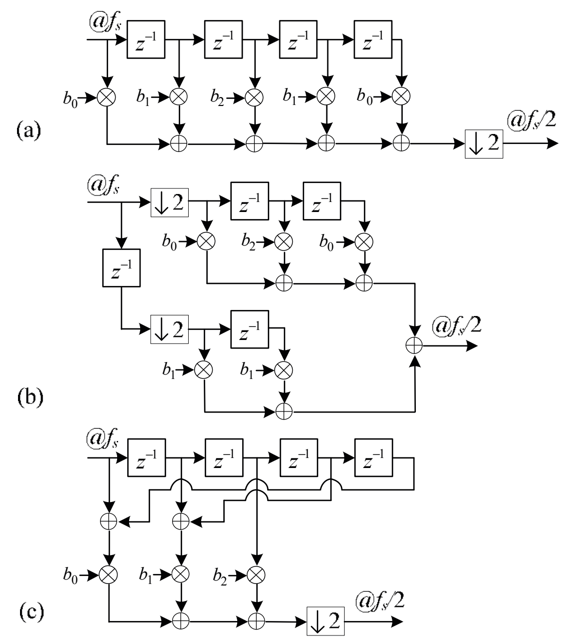


Fig. 3 Three different structures for implementing a 5-tap decimation filter with decimation factor $M = 2$: (a) original, (b) GNI-derived, and (c) folded FIR structures.

References

[1] K. Pretz, “Greener technologies and networks,” *IEEE The Institute,* vol. 40, issue 1, p. 6, Mar. 2016.

[2] K. K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation.* New York: Wiley, 1999.

[3] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective.* Boston: Pearson, 2010.

[4] R. G. Lyons, *Understanding Digital Signal Processing.* Boston: Person Education, 2011.

[5] J. O. Smith III, *Spectral Audio Signal Processing*, Dec. 2011. [Online]. Available: https://www.dsprelated.com/freebooks/sasp/Multirate_Noble_Identities.html

Proposal of Extracting Pulse Wave During Driving a Car Based on Frequency Domain BSS

Kakeru Ishikawa, Minoru Komatsu, Hiroki Matsumoto
$m1671005$, $komatsu$, $matsumoto@maebashi-it.ac.jp$
Maebashi Institute of Technology, 460–1 Kamisadori, Maebashi, Gumma, 371–0816 Japan

I. Introduction

Recently, monitor for the health condition of a driver during driving a car is actively studied [1]. Especially, pulse wave (PW) is studied to monitor the health condition of a driver during driving a car. Conventionally, the observed signals are assumed instantaneous mixture of PW and noise, and the separation precision with conventional method is lower [3]. Therefore, the observed signal may be convolutive mixtures. In this paper, in order to improve separation precision, we propose the method based on frequency domain Blind Source Separation (BSS) for convolutive mixture. We evaluate the advantage of the proposed method by computer simulation.

II. Proposed method

In this section, we describe the proposed method. First, observed signals are mapped to the frequency domain, and observed signals on frequency domain are separated PW with nonstationary noise (NSnoise) and stationary noise (Snoise) with NSnoise using frequency domain BSS. Next, NSnoise is removed from two separated signals. As a result, PW and Snoise are generated. We select PW by using characteristic of amplitude spectrum for two generated signals. We map frequency spectrum corresponding to amplitude spectrum of PW into time domain. From the above, we can extract PW accurately. We use block processing because it is difficult to process in real time when we process to use all data.

A. Blocking of data

We assume that an observed signal is convolutive mixture model of PW with NSnoise and Snoise with NSnoise. A column observed signal vector $\boldsymbol{x}_2(n)$ is defined by

$$\boldsymbol{x}_2(n) = \left[x_{(1)}(n), x_{(2)}(n)\right]^T. \tag{1}$$

The observed signal is expressed by

$$\boldsymbol{x}_2(n) = \sum_{m=0}^{M-1} \boldsymbol{a}_{2,2}(m)\boldsymbol{s}_2(n-m), \tag{2}$$

where, $\boldsymbol{a}_{2,2}(m)$ is the mixing matrix and $\boldsymbol{s}_2(n-m)$ is a source signal vector.

Blocking of data is required. Because, in the proposed method, real time processing is performed in frequency domain. In order to regenerate L data of PW for each block, we use $2L$ data of observed signals for each block because we perform block processing with overlap. Letting the block observed signal matrix constructed of column vector $\boldsymbol{x}_2(n)$ be $\tilde{\boldsymbol{x}}_{2,2L}^{(l)}$, $\tilde{\boldsymbol{x}}_{2,2L}^{(l)}$ is given by

$$\begin{aligned}\tilde{\boldsymbol{x}}_{2,2L}^{(l)} &= \left[\boldsymbol{x}_2(lL-\frac{L}{2}), \cdots, \boldsymbol{x}_2(lL+L-1+\frac{L}{2})\right] \quad (3)\\ &= \left[\tilde{\boldsymbol{x}}_{(1),2L}^{(l)}, \tilde{\boldsymbol{x}}_{(2),2L}^{(l)}\right]^T, \quad (4)\end{aligned}$$

where, row vectors $\tilde{\boldsymbol{x}}_{(i),2L}^{(l)}$ denote $\left[x_{(i)}(lL-\frac{L}{2}), \cdots, x_{(i)}(lL+L-1+\frac{L}{2})\right], (i=1,2)$. $2L$ is the block length and l is the block number.

B. Discrete Fourier Transform (DFT)

Letting the matrix constructed of row vector $\tilde{\boldsymbol{X}}_{(i),2L}^{(l)} = \left[\tilde{X}_{(i)}^{(l)}(0), \cdots, \tilde{X}_{(i)}^{(l)}(2L-1)\right] = DFT\left[\tilde{\boldsymbol{x}}_{(i),2L}^{(l)}\right]$, $(i=1,2)$ be $\tilde{\boldsymbol{X}}_{2,2L}^{(l)}$, $\tilde{\boldsymbol{X}}_{2,2L}^{(l)}$ is expressed by

$$\begin{aligned}\tilde{\boldsymbol{X}}_{2,2L}^{(l)} &= \left[\tilde{\boldsymbol{X}}_{(1),2L}^{(l)}, \tilde{\boldsymbol{X}}_{(2),2L}^{(l)}\right]^T \quad (5)\\ &= \left[\tilde{\boldsymbol{X}}_2^{(l)}(0), \cdots, \tilde{\boldsymbol{X}}_2^{(l)}(2L-1)\right]. \quad (6)\end{aligned}$$

where, column vectors $\tilde{\boldsymbol{X}}_2^{(l)}(k)$ is expressed by

$$\tilde{\boldsymbol{X}}_2^{(l)}(k) = \left[\tilde{X}_{(1)}^{(l)}(k), \tilde{X}_{(2)}^{(l)}(k)\right]^T, (0 \le k \le 2L-1), \tag{7}$$

k is a frequency.

C. Signal separation

Column separated signal vectors $\tilde{\boldsymbol{Y}}_2^{(l)}(k)$ for each frequency k is calculated by

$$\tilde{\boldsymbol{Y}}_2^{(l)}(k) = \boldsymbol{W}_{2,2}^{(l)}(k)\tilde{\boldsymbol{X}}_2^{(l)}(k), (0 \le k \le 2L-1). \tag{8}$$

Where, $\tilde{\boldsymbol{Y}}_2^{(l)}(k)$ denote $\left[\tilde{Y}_{(1)}^{(l)}(k), \tilde{Y}_{(2)}^{(l)}(k)\right]^T$. The separation matrix in frequency domain $\boldsymbol{W}_{2,2}^{(l)}(k)$ is expressed by

$$\boldsymbol{W}_{2,2}^{(l)}(k) = \begin{pmatrix} W_{11}^{(l)}(k) & W_{12}^{(l)}(k) \\ W_{21}^{(l)}(k) & W_{22}^{(l)}(k) \end{pmatrix}, (0 \le k \le 2L-1). \tag{9}$$

The observed signals are separated PW with NSnoise and Snoise with NSnoise. We use the method of Ref. [2] to update $\boldsymbol{W}_{2,2}^{(l)}(k)$, $\boldsymbol{W}_{2,2}^{(l)}(k)$ is as follows:

$$\begin{aligned}\boldsymbol{W}_{2,2}^{(l+1)}(k) = \boldsymbol{W}_{2,2}^{(l)}(k) + \eta[diag\left\langle \varphi\left(\tilde{\boldsymbol{Y}}_2(k)\right)\tilde{\boldsymbol{Y}}_2^H(k)\right\rangle \\ - \varphi\left(\tilde{\boldsymbol{Y}}_2(k)\right)\tilde{\boldsymbol{Y}}_2^H(k)]\boldsymbol{W}_{2,2}^{(l)}(k) \quad (10)\end{aligned}$$

$$(0 \le k \le 2L-1),$$

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

where, η is a step gain. PW is the Super-Gaussian distribution. Therefore, the nonlinear function $\varphi\left(\tilde{\boldsymbol{Y}}_2(n)\right)$ is expressed by

$$\varphi\left(\tilde{\boldsymbol{Y}}_2(k)\right) = \tanh\left(\tilde{\boldsymbol{Y}}_2^{(R)}(k)\right) + j\tanh\left(\tilde{\boldsymbol{Y}}_2^{(I)}(k)\right), \quad (11)$$

where, $\tilde{\boldsymbol{Y}}_2^{(R)}$ is real part of $\tilde{\boldsymbol{Y}}_2$ and $\tilde{\boldsymbol{Y}}_2^{(I)}$ is imaginary part of $\tilde{\boldsymbol{Y}}_2$.

D. NSnoise reduction

Letting the separated signal matrix constructed of column vectors $\tilde{\boldsymbol{Y}}_2^{(l)}(k)$ be $\tilde{\boldsymbol{Y}}_{2,2L}^{(l)}$, $\tilde{\boldsymbol{Y}}_{2,2L}^{(l)}$ is expressed by

$$\begin{aligned} \tilde{\boldsymbol{Y}}_{2,2L}^{(l)} &= \left[\tilde{\boldsymbol{Y}}_2^{(l)}(0), \cdots, \tilde{\boldsymbol{Y}}_2^{(l)}(2L-1)\right] &(12)\\ &= \left[\tilde{\boldsymbol{Y}}_{(1),2L}^{(l)}, \tilde{\boldsymbol{Y}}_{(2),2L}^{(l)}\right]^T. &(13)\end{aligned}$$

Where, row vectors $\tilde{\boldsymbol{Y}}_{(i),2L}^{(l)}$ is given by

$$\tilde{\boldsymbol{Y}}_{(i),2L}^{(l)} = \left[\tilde{Y}_{(i)}^{(l)}(0), \cdots, \tilde{Y}_{(i)}^{(l)}(2L-1)\right], (i=1,2), \quad (14)$$

where, i is the signal sequence number. Letting amplitude spectrum vector of $\tilde{\boldsymbol{Y}}_{(i),2L}^{(l)}$ be $\hat{\boldsymbol{Y}}_{(i),2L}^{(l)}$, $\hat{\boldsymbol{Y}}_{(i),2L}^{(l)}$ is given by

$$\hat{\boldsymbol{Y}}_{(i),2L}^{(l)} = \left[\hat{Y}_{(i)}^{(l)}(0), \cdots, \hat{Y}_{(i)}^{(l)}(2L-1)\right], (i=1,2). \quad (15)$$

Here, if the data length is sufficiently long, the amplitude of NSnoise is small. Therefore, in order to reduce NSnoise, the following processing is performed for $\hat{Y}_{(i)}^{(l)}(k)$,

$$\bar{Y}_{(i)}^{(l)}(k) = \begin{cases} \hat{Y}_{(i)}^{(l)}(k) & (\hat{Y}_{(i)}^{(l)}(k) > th) \\ 0 & (\text{otherwise}) \end{cases} \quad (16)$$
$$(0 \le k \le 2L-1\ ,\ i=1,2).$$

The amplitude spectrum row vectors $\bar{\boldsymbol{Y}}_{(i),2L}^{(l)}$ without NSnoise is given by

$$\bar{\boldsymbol{Y}}_{(i),2L}^{(l)} = \left[\bar{Y}_{(i)}^{(l)}(0), \cdots, \bar{Y}_{(i)}^{(l)}(2L-1)\right], (i=1,2). \quad (17)$$

E. Signal selection

Either $\bar{\boldsymbol{Y}}_{(1),2L}^{(l)}$ or $\bar{\boldsymbol{Y}}_{(2),2L}^{(l)}$ is PW, and the other is the Snoise. Therefore, we have to select PW. For the selection, we use the characteristic that the amplitude spectrum of PW has regularity. First, the local maximum value is detected for each amplitude spectrums. Next, amplitude spectrum with regularity at the local maximum value is PW.

F. Inverse Discrete Fourier Transform(IDFT)

Letting frequency spectrum corresponding to amplitude spectrum $\bar{\boldsymbol{Y}}_{(i),2L}^{(l)}$ of PW be $\boldsymbol{Y}_{2L}^{(l)}$, the row regenerated signal vector $\boldsymbol{y}_{2L}^{(l)} = IDFT[\boldsymbol{Y}_{2L}^{(l)}]$ is given by

$$\boldsymbol{y}_{2L}^{(l)} = \left[\boldsymbol{y}(lL - \frac{L}{2}), \cdots, \boldsymbol{y}(lL + L - 1 + \frac{L}{2})\right]. \quad (18)$$

The proposed algorithm is a block processing with overlap. Therefore, L central portions of $\boldsymbol{y}_{2L}^{(l)}$ are picked out. Then, the regenerated signal vector $\boldsymbol{y}_L^{(l)}$ of PW is represented by

$$\boldsymbol{y}_L^{(l)} = \left[\boldsymbol{y}(lL), \cdots, \boldsymbol{y}(lL + L - 1)\right]. \quad (19)$$

III. Computer Simulation

In this section, we describe computer simulation. The simulation conditions are as follows: cutoff frequency of analog filter is 200Hz, sampling frequency is 480Hz, cutoff frequency of digital filter is 10Hz, block length $2L$ is 5760, step gain η is 0.01, threshold th is 400. Moreover, an initial separation matrix is as follows:

$$\boldsymbol{W}_{2,2}^{(l)}(k) = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}. \quad (20)$$

Fig.1 shows simulation result. It is observed and separated signals in the driving state. Conventional BSS in the Fig.1 is used the method of Ref. [3]. Two sensors (#0, #1) are the observed signals obtained from the left thigh. A sensor (#Finger) is a data attached to the left index finger for confirmation. Here, we used a piezoelectric PW sensor (AYA-P, TAIYO YUDEN CO., LTD.) for observation.

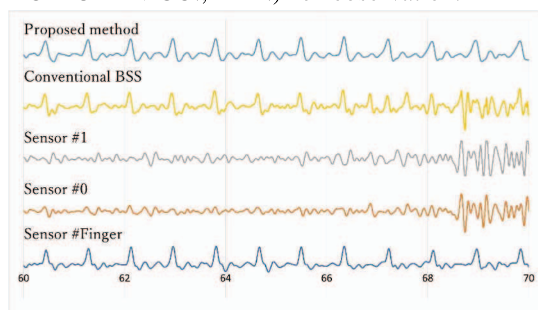


Fig. 1. Raw waveforms and separation results during driving

In the Fig.1, the NSnoise is generated at 67-70sec. The conventional method can't remove NSnoise. On the other hand, the proposed method can remove it. From the result, the proposed method can obtain PW with high separation precision compared to the conventional BSS.

IV. Conclusion

The conventional method is assumed that observed signal is instantaneous mixture model. Therefore, the separation precision is lower. Then, we assume that observed signal is convolutive mixture model, and we propose separation algorithm for convolutive mixture model. When the proposed method is used, the separation precision is improved compared with the conventional method. The future work is to consider the decision method of threshold for NSnoise reduction.

References

[1] S Esaki, G Obinata, S Tokuda, N Mori, M Makiguchi "Estimation of Driver's Fatigue Using Physiological Measures and Principal Component Analysis" *The Japan Joint Automatic Control Conference* Vol. 54, pp. 61-61, 2011

[2] S. Amari "Natural gradient works efficiently in learning" *Neural Computation* Vol. 10, pp. 271-276, 1998

[3] Hotaka Takada, Takashi Ishiguro, Tomomi Ogawa, Hiroki Matsumoto "A Proposal of Extraction Method of Pulse Wave during Driving of a Vehicle" *RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP) 2018 international Symposium on* March, 2018

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

A Supervised Learning Method for the Design of Linear Phase FIR Digital Filter Using Keras

Chien-Cheng Tseng
Depart. of Computer and Communication Engineering
National Kaohsiung University of Sci. and Tech.
Kaohsiung, Taiwan
tcc@nkust.edu.tw

Su-Ling Lee
Depart. of Computer Science and Information Engineering
Chang-Jung Christian University
Tainan, Taiwan
lilee@mail.cjcu.edu.tw

***Abstract*—In this paper, a supervised learning method for the design of linear phase FIR digital filter using Keras is presented. First, the design problem of the linear phase finite impulse response (FIR) digital filter is transformed to a supervised learning problem. Then, the optimizers in Keras framework are used to determine the filter coefficients by minimizing the mean squared error (MSE) loss function. The widely-used optimizers include adaptive moments (Adam) algorithm and stochastic gradient descent (SGD) with momentum algorithm. Finally, the numerical design examples of low-pass and high-pass FIR digital filters are demonstrated to show the usefulness of the supervised learning method with Keras framework.**

Keywords—digital filter; FIR filter; linear phase; supervised learning; deep learning

I. INTRODUCTION

In recent years, the deep learning has received great attentions in the research areas of machine learning and artificial intelligence [1]-[4]. The well-known deep learning methods are supervised leaning, unsupervised learning and semi-supervised learning etc. So far, the deep convolutional neural networks (CNN) with supervised learning algorithms have been successfully applied to object detection, object location, image classification, image segmentation, and face recognition etc. The popular frameworks to implement the deep CNN are Caffe, Theano, Tensorflow, Keras, CNTK and Torch. Because the Keras framework provides a convenient way to define and train deep networks, it is interesting to use Keras to develop a supervised learning method for designing linear phase FIR digital filter in the digital signal processing area. Digital filters are useful tools in digital signal processing applications [5]. Some typical examples are listed below. The frequency-selective filters (low-pass, band-pass, band-stop, high-pass) can be used to extract the wanted frequency components from the received signal. The notch filter can be applied to remove the narrowband power-line interference superimposed on the wideband biomedical signal. The digital differentiator can be employed to compute the velocity signal from the measured position signal. Thus, it is an interesting topic that uses the supervised learning algorithms in Keras framework to solve digital filter design problem. This topic will be studied in this paper.

II. FIR FILTER DESIGN USING SUPERVISED LEARNING

The transfer function of an FIR digital filter is given by

$$H(z) = \sum_{n=0}^{N} h(n) z^{-n} \tag{1}$$

If the symmetric condition $h(n) = h(N-n)$ is satisfied and the degree N is even, the FIR filter is a type I linear phase digital filter. Let $z = e^{j\omega}$, the frequency response can be written as

$$H(e^{j\omega}) = \left(h(\tfrac{N}{2}) + 2\sum_{k=1}^{\frac{N}{2}} h(\tfrac{N}{2} - k)\cos(k\omega) \right) e^{-j\frac{N}{2}\omega} \tag{2}$$

Clearly, the phase response is linear and amplitude response is given by

$$A(\omega) = \sum_{k=0}^{M} a_k \cos(k\omega) \tag{3}$$

where $M = \frac{N}{2}$, $a_0 = h(M)$, and $a_k = 2h(M-k)$ for $k = 1,2,\cdots,M$. Now, the design problem is how to determine the coefficients a_k such that the amplitude response $A(\omega)$ approximates the ideal response $D(\omega)$ as well as possible. That is, the following equation is wanted to be satisfied:

$$D(\omega) = A(\omega) = \sum_{k=0}^{M} a_k \cos(k\omega) \qquad \omega \in \Omega \tag{4}$$

where Ω is the interested frequency band depending on the specification of the digital filter. Let ω_ℓ $(\ell = 1,2,\cdots,L)$ be the uniform dense grid points in the interested band Ω, then equation (4) can be approximated by

$$D(\omega_\ell) = \sum_{k=0}^{M} a_k \cos(k\omega_\ell) \qquad \ell = 1,2,\cdots,L \tag{5}$$

According to the equation (5), Fig.1 depicts a supervised learning neural architecture to find the filter coefficients a_k. The activation function is chosen as a linear function $f(x) = x$. And, a_0 is the bias term in the architecture. The desired target response is $D(\omega_\ell)$ and the input training data vector is $[1, \cos(\omega_\ell), \cdots, \cos(M\omega_\ell)]^T$ for $\ell = 1,2,\cdots,L$. In this paper, the following mean squared error (MSE) loss function is minimized to train the architecture in Fig.1:

$$J = \frac{1}{m}\sum_{\ell \in S}\left|D(\omega_\ell) - \sum_{k=0}^{M} a_k \cos(k\omega_\ell)\right|^2 = \frac{1}{m}\sum_{\ell \in S}|e(\omega_\ell)|^2 \quad (6)$$

where m is batch size, S is the sampled set of data batch, and the amplitude response error is $e(\omega_\ell) = D(\omega_\ell) - \sum_{k=0}^{M} a_k \cos(k\omega_\ell)$. The optimizer often uses the following steepest descent algorithm to update the filter coefficients a_k by minimizing the MSE loss function:

$$a_k(n+1) = a_k(n) - \mu \frac{\partial J}{\partial a_k} \qquad k = 0,1,2,\cdots,M \quad (7)$$

where n is the iteration number and parameter μ is the learning rate. So far, two widely-used optimizers to accomplish the coefficient updates in (7) are the adaptive moments (Adam) algorithm and stochastic gradient descent (SGD) with momentum algorithm. Now, the sequential model on Keras framework is used to implement the architecture in Fig.1. The related Keras code is listed below:

```
model=sequential()
model.add(Dense(1, input_dim=M))
model.add(Activation('linear'))
model.compile(loss='mse', optimizer='adam')
model.fit(Φ , d , batch_size=m, epochs=T)
```

In the above, the matrix Φ and vector d are given by

$$\Phi = \begin{bmatrix} 1 & \cos(\omega_1) & \cdots & \cos(M\omega_1) \\ 1 & \cos(\omega_2) & \cdots & \cos(M\omega_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos(\omega_L) & \cdots & \cos(M\omega_L) \end{bmatrix} \quad d = \begin{bmatrix} D(\omega_1) \\ D(\omega_2) \\ \vdots \\ D(\omega_L) \end{bmatrix} \quad (8)$$

Until now, the design of linear phase FIR digital filter using supervised learning method with Keras has been described. In next section, the examples of low-pass and high-pass FIR filter deigns are demonstrated to show its effectiveness.

III. DESIGN EXAMPLES

In this section, the design examples and comparisons of the proposed filter design method are presented.

Example 1: In this example, the low-pass filter is designed. The ideal amplitude response is given by

$$D(\omega) = \begin{cases} 1 & 0 \le \omega \le \omega_c \\ 0 & \omega_c < \omega \le \pi \end{cases} \quad (9)$$

where ω_c is the cutoff frequency. Thus, the interested frequency band is selected as $\Omega = [0, \omega_c - \Delta] \cup [\omega_c + \Delta, \pi]$. The parameters are chosen as $M = 20$, $\omega_c = 0.5\pi$, $\Delta = 0.05\pi$, $L = 2837$, $\mu = 0.01$, $m = 4$, and $T = 10$. Fig.2(a) shows the amplitude response of the designed low-pass filter. Next, the proposed method is compared with conventional window method in [5] whose filter coefficients are given by

$$h(n) = \frac{\sin(\omega_c(n-M))}{\pi(n-M)} \quad (10)$$

Fig.2(b) depicts the amplitude response of the low-pass filter designed by the window method. Obviously, the proposed supervised learning method has smaller approximation error than conventional window method.

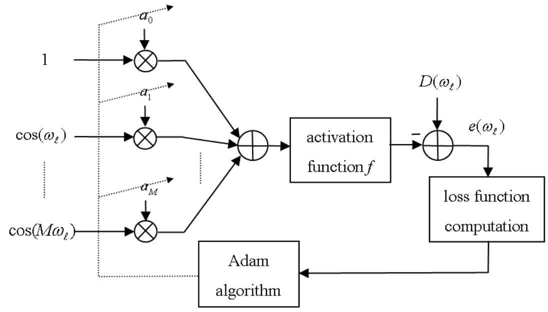
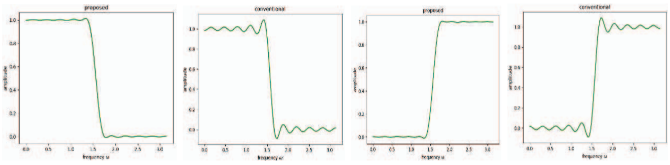


Fig.1 A supervised learning architecture for the design of linear phase FIR digital filters using Keras. Note that $f(x)$ is a linear activation function, a_0 is the bias term and $a_1, a_2, \cdots, a_M$ are the weights.



(a) (b) (c) (d)

Fig.2 Design results of FIR filters. (a) The amplitude response of the low-pass filter designed by proposed method. (b) The amplitude response of the low-pass filter designed by window method. (c) The amplitude response of the high-pass filter designed by proposed method. (d) The amplitude response of the high-pass filter designed by window method.

Example 2: In this example, the high-pass filter is designed. The ideal amplitude response is given by

$$D(\omega) = \begin{cases} 0 & 0 \le \omega < \omega_c \\ 1 & \omega_c \le \omega \le \pi \end{cases} \quad (11)$$

where ω_c is the cutoff frequency. If the parameters are chosen as $M = 20$, $\omega_c = 0.5\pi$, $\Delta = 0.05\pi$, $L = 2837$, $\mu = 0.01$, $m = 4$, and $T = 10$, Fig.2(c)(d) show the amplitude responses of high-pass filters designed by the proposed method and conventional window method. Clearly, the proposed method has smaller error than conventional window method.

IV. CONCLUSIONS

In this paper, the linear phase FIR digital filter has been designed. However, only linear phase FIR filter design is studied here. Thus, it is interesting to use a supervised learning method to design nonlinear phase FIR filters in the future.

REFERENCES

[1] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, The MIT Press, 2016.

[2] C.C. Aggarwal, *Neural Networks and Deep Learning: A Textbook*, Springer, 2018.

[3] F. Chollet, *Deep Learning with Python*, Manning, 2018.

[4] *R. Atienza,* Advanced Deep Learning with Keras, Packt, 2018.

[5] S.K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, Fourth Edition, McGRAW-Hill 2011.

Improvement of Noise Suppression Performance of SEGAN by Sparse Latent Vectors

Minami Sakuma, Yosuke Sugiura, Tetsuya Shimamura
Saitama University, Saitama, Japan
E-mail: {sakuma, sugiura, shima}@sie.ics.saitama-u.ac.jp

***Abstract*—For the purpose of speech enhancement, SEGAN, which is one of deep generative models, has attracted attention due to its high performance. In this paper, we propose a method to sparse latent vectors to further enhance the noise suppression effect of SEGAN.**

***Index Terms*—speech enhancement, generative model, SEGAN, sparse modeling**

I. Introduction

In recent years, with the spread of mobile communication terminals such as smartphones and tablet-type terminals, the opportunity to make a call in a noisy environment has increased, and the improvement of the noise suppression performance is desired. Speech enhancement has been studied for a long time, and various methods have been proposed. A classical approach is based on estimating and removing noise components contained in the input signal in the frequency domain.

Typical approaches for speech enhancement include the Wiener filter, the MMSE-STSA method [1], and so on. These techniques are known to be effective against stationary noise. On the other hand, their performance in non-stationary noise is degraded, but improved by combining a weighted noise estimation method [2], etc. Inherently, it is difficult to accurately estimate the spectrum of non-stationary noise, and problems remain due to the estimation error.

Recently, in such a situation, a deep neural network (DNN) approach has attracted attention. In the approach, the noise-containing signal is given as the input and the clean signal is used as the target output for learning. As a result, the difference between noise and voice characteristics can be automatically acquired in the network, and the non-stationary noise can be removed accurately.

Among the methods involving deep learning, Speech Enhancement GAN (SEGAN) [4] based on a generative adversarial network (GAN) [3] is particularly effective. There are two networks in SEGAN; Generator (G) and Discriminator (D). G is a network that generates enhanced signals, and D is a network that determines whether the input signal is a signal generated by G or the clean signal. While G tries to minimize the objective function, D tries to maximize the objective function. By this minimax principle, learning progresses while G and D compete with each other, and as a result, it is possible to generate a highly emphasized signal.

In this paper, we propose a method to sparsify latent vectors in order to further enhance the noise suppression effect of SEGAN.

II. SEGAN and Sparsification

A. Overview of SEGAN

Here we give an overview of SEGAN. Structure of SEGAN's G is different from GAN's, which is based on Auto-Encoder. Speech enhancement is achieved by inputting the noise-containing speech waveform to G and learning so that the clean speech waveform is output. The objective function of SEGAN is based on Least Squares GAN (LSGAN) [5] in which the square error is applied to the objective function of GAN.

However, the objective function for G adds the L_1 norm term of the error between the G-made signal and clean signal to reduce their difference. The objective functions in D and G are shown below, respectively,

$$\begin{aligned}\min_D V_{\text{SEGAN}}(D) = \frac{1}{2}\mathbb{E}_{\mathbf{x},\tilde{\mathbf{x}}\sim p_{\text{data}}(\mathbf{x},\tilde{\mathbf{x}})}[(D(\mathbf{x},\tilde{\mathbf{x}})-1)^2] \\ + \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z}),\tilde{\mathbf{x}}\sim p_{\text{data}}(\tilde{\mathbf{x}})}[D(G(\mathbf{z},\tilde{\mathbf{x}}),\tilde{\mathbf{x}})^2]\end{aligned} \tag{1}$$

$$\begin{aligned}&\min_G V_{\text{SEGAN}}(G) = \\ &\qquad \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z}),\tilde{\mathbf{x}}\sim p_{\text{data}}(\tilde{\mathbf{x}})}[(D(G(\mathbf{z},\tilde{\mathbf{x}}),\tilde{\mathbf{x}})-1)^2] \\ &\qquad + \lambda\|G(\mathbf{z},\tilde{\mathbf{x}})-\mathbf{x}\|_1\end{aligned} \tag{2}$$

where $\mathbf{x}$ is the clean signal, $\tilde{\mathbf{x}}$ is the noisy signal, and λ is a hyper-parameter that determines the magnitude of the L_1 norm.

B. Structure of Generator

Fig. 1 shows the specific structure of SEGAN's Generator. SEGAN's G is composed of a fully convolutional network using only dilated convolution, and there are no fully connected layers or pooling layers.

After the noisy signal is input to G, we have the following procedures.

1) The noisy signal $\tilde{\mathbf{x}}$ is input to the G's encoder (Genc), and the dimension reduction is compressed by the convolutional layer.

2) After the dimension reduction, the vector $\mathbf{c}$ is output by the encoder, the size of which is the same as the

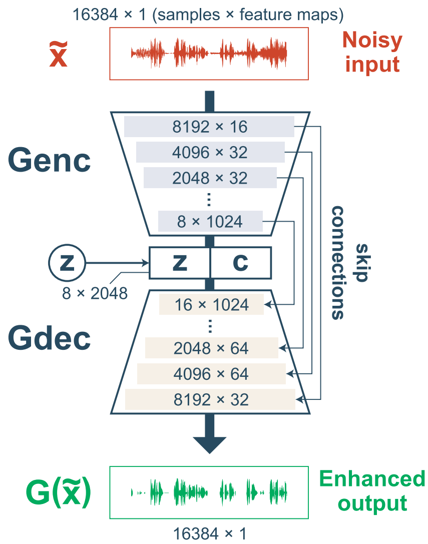


Fig. 1. Structure of SEGAN's Generator

random vector $\mathbf{z}$. $\mathbf{c}$ and $\mathbf{z}$ are concatenated to obtain a latent vector.

3) The obtained latent vector is input to G's decoder (Gdec), and dimensional restoration is performed to obtain the enhanced signal $\mathrm{G}(\tilde{\mathbf{x}})$. At this time, the encoder sends the fine structure of the voice to the decoder by skip connections.

The generated signal is convoluted again by D, and it is judged whether it is the clean signal or a signal created by G. Based on the error, Back Propagation makes the enhanced signal created by G close to the original signal. In this way, SEGAN learning is performed. SEGAN can exhibit the high rejection performance against non-stationary noise, but weak noise can not be removed, which may remain in the enhanced signal.

C. Sparsification of Latent Vectors

To sparsify the latent vector, we add the L_1 regularization term of the latent vector to the objective function of G. That is, the objective function is newly given by the following equation:

$$\min_G V_{\mathrm{SEGAN}}(G) = \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z}),\tilde{\mathbf{x}}\sim p_{\mathrm{data}}(\tilde{\mathbf{x}})}[(D(G(\mathbf{z},\tilde{\mathbf{x}}),\tilde{\mathbf{x}})-1)^2] + \lambda_1\|G(\mathbf{z},\tilde{\mathbf{x}})-\mathbf{x}\|_1 + \lambda_2\|\mathbf{c}\|_1 \tag{3}$$

where λ_2 represents a hyper-parameter that determines the strength of regularization for the latent vector.

The regularization term added here is minimized together with the original objective function by Back Propagation. As a result, the component of the latent vector with small values becomes 0, and only the component with large values functions to contribute to the reconstruction in the decoder. However, if the value of λ_2 is set to be large, the important component of the latent vector will be close to 0, thus the performance may be deteriorated. Therefore, proper adjustment of λ_2 is important.

TABLE I
OBJECTIVE EVALUATION (PESQ, SNR)

	PESQ	SNR [dB]
Input	2.67	5.50
Conv.	2.76	16.48
Prop. ($\lambda_2 = 1$)	2.74	**17.19**
Prop. ($\lambda_2 = 1.05$)	**2.87**	17.15

III. Experiment

The data set used in the experiment is the same as Pascual et al. [4]. 14 male speakers and 14 female speakers are used as training data, one male and one female are used as test data.

Both the conventional SEGAN and the proposed method are trained for 100 epochs. The value of λ_1 during training is set as 100 according to [4]. In addition, the value of λ_2 for the proposed method is prepared as $\lambda_2 = 1, 1.05$.

Table I shows PESQ and SNR in each method. Each value is an average of evaluations for all test data. Focusing on PESQ, the proposed method (Prop.) ($\lambda_2 = 1.05$) provides 0.11 points better than the conventional method (Conv.). Prop. ($\lambda_2 = 1$) has a slightly worse value than Conv. On the other hand, in SNR, Prop. ($\lambda_2 = 1$) improves by $0.71\,\mathrm{dB}$ over Conv., which is the best result. In SNR, Prop. ($\lambda_2 = 1.05$) also shows better results than Conv. From the above, it can be seen that the noise suppression performance of the proposed method is superior to that of the conventional method in terms of PESQ and SNR.

IV. Conclusions

In this paper, we have proposed a method of sparsifying latent vectors in order to improve the noise suppression performance of SEGAN. By adding the L_1 norm regularization term to the objective function and sparsifying the latent vectors, the components of the latent vectors unnecessary for restoration become 0, leading to a performance improvement. We confirmed the effectiveness of the proposed method by experiments.

Future work will be to conduct experiments while changing the value of λ_2, and to investigate the numbers that give better results. It is also necessary to evaluate subjectively and improve the performance furthermore.

References

[1] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator", IEEE Trans. on Acoust., Speech and Signal Process., ASSP-33, no. 2, pp. 443-445, 1985.
[2] M. Kato, A. Sugiyama, and M. Serizawa, "Noise suppression with high speech quality based on weighted noise estimation and MMSE STSA", IEICE Trans. Fund. vol. E85-A, no. 7, pp. 1710-1718, 2002.
[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", Proc. of NIPS, 2014.
[4] S. Pascual, A. Bonafonte, and J. Serra, "SEGAN: Speech enhancement generative adversarial network", Proc. of INTERSPEECH 2017, pp. 3642-3646, 2017.
[5] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks", Proc. of 2017 IEEE International Computer Vision, pp. 2813-2821, 2017.

Novel Deblocking Method for Cropped Video

Kenta Hashimoto
Department of Informatics
Kogakuin University
Tokyo, Japan
em19016@ns.kogakuin.ac.jp

Seiichi Gohshi
Department of Informatics
Kogakuin University
Tokyo, Japan
gohshi@cc.kogakuin.ac.jp

***Abstract*— Due to a large amount of data, videos are encoded (data volume reduction) before transmission and storage. However, coded videos have degradations. One of them is called 'block noise' that appears block shape degradation on the image. Block noise can be usually reduced by a deblocking filter that reduces the block boundaries. However, the conventional deblocking filter has an issue that it cannot be applied to cropped video with the block noise. Although the deblocking filter reduces the block noise, it also causes blur on the edges in the video. In this paper, we propose a novel deblocking method that can reduce block noise for cropped videos by keeping the edges.**

Keywords—signal processing, block noise, smoothing

I. Introduction

Videos contain substantial data. Therefore, they are encoded to reduce data volume before transmission and storage. H.264/MPEG-4 AVC is a commonly used encoding standard [1]. However, encoded videos cause degradations disturbing our watching. One such degradation called "block noise" causes block-shaped degradation in the video.

Block noise can usually be reduced using a deblocking filter[2]. A deblocking filter reduces block boundaries via the smoothing process. However, conventional deblocking filters blur the edges on the cropped video. No method exists for reducing block noise on cropped videos at present.

Herein, we propose a novel deblocking method that can reduce block noise in cropped videos while retaining the edges.

II. Problems of Conventional Methods

This chapter explains conventional deblocking filters and their problems. In the case of uncropped videos, block boundaries appear every 8 pixels or 4 pixels both horizontally and vertically at equal intervals. Conventional deblocking filters smooth these [2]. This paper discusses the case of 8 pixel intervals. The smoothing strength of the filter is determined using the quantization scale (QS), which is a parameter generated when the video is encoded and saved or transmitted with the encoded data. QS is related to the compression level, which is nearly equivalent to the degradation level. QS is used to decode the video; however, it is lost after decoding. Conventional deblocking filters use QS during video decoding.

However, conventional methods cannot be applied to cropped videos. It blurs edges of the video. Although block boundary positions in cropped videos possess a cyclic pattern, they also possess a phase shift at the end of the screen. Additionally, QS, which is necessary to determine the smoothing strength, is unavailable in cropped videos because the video is already decoded before cropping. Fig. 2 shows a zoomed image of a part of the image processed using the conventional deblocking method that does not consider the block boundary phase shift. Here, the smoothing strengths are set to constant values at all block boundaries. Edges such as

Fig.1. Block noise

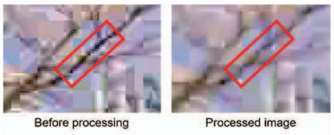


Fig.2. The result of the conventional deblocking filter (cropped video)

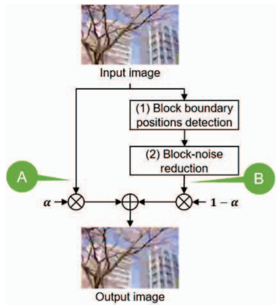


Fig.3. Block diagram of the proposed method

the tree branch (shown in the red frame) are blurred as compared to the image before processing.

III. Proposed Method

This chapter proposes a novel deblocking method that can be applied to cropped videos. The proposed method first detects block boundaries, and then applies a smoothing filter to the detected block boundaries while excluding the edges from the smoothing process.

Fig. 3 shows the block diagram (from the input image to the output image) of the proposed method. First, block boundaries in the input image are detected (1) using the existing block boundary detection algorithm [3]. Then, a deblocking filter is applied to the detected block boundaries (2). Since the input image is already decoded and cropped, the QS is unavailable. Therefore, each smoothing strength is determined using the differences between two-pixel values facing the block boundary. The input image (A) has unblurred edges that the original image possesses, but which suffers from noticeable block noise. Conversely, the deblocked image (B) has blurry edges but reduced block noise. Using these features of images (A) and (B), we can exclude the edges from the deblocking process.

By compositing both images, (A) is assigned to areas having edges and (B) is assigned to areas having no edges. The composition ratios are determined on a per-pixel basis using the coefficient α related to the edge detection results. Subsequently, (A) is multiplied by α per pixel and (B) is multiplied by $1 - \alpha$ per pixel. Then, both images are added. The α coefficients are generated by nonlinearly converting the outputs of the edge detection and clipping them between 0 and 1. In areas having edges, α values are close to 1; therefore, final output pixel values are similar to the input image pixel values. In contrast, in areas having no edges, α values are close to 0; thus, final output pixel values are similar to the deblocked image pixel values. The abovementioned process is repeated for each frame of the video sequence. In the next section, we explain the details of edge detection during the composition.

A. *Edge Detection*

We can detect the edges by applying the Sobel filter [4]. Results of the Sobel filter contain real edges that the original image possesses together with block boundaries generated via video encoding. Fig. 4 shows the results of edge detection for

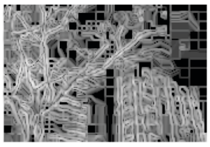


Fig.4. The result of Sobel Filter (combine horizontal with vertical)

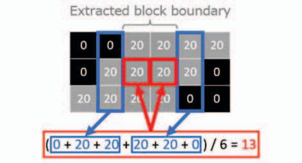
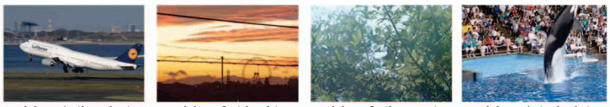


Fig.5. Block boundary removal



video 1 (jumbo) video 2 (dusk) video 3 (leaves) video 4 (whale)

Fig.6. Video sequences used in the experiment

TABLE I PSNR IMPROVEMENT

Video sequence number	Maximum PSNR improvement [dB]		Average PSNR improvement [dB]	
	Conventional method (64 patterns)	Proposed method	Conventional method (64 patterns)	Proposed method
1	+0.04 – +0.45	+0.45	+0.01 – +0.33	+0.36
2	0.00 – +0.49	+0.73	-0.29 – +0.34	+0.46
3	+0.04 – +0.49	+0.58	-0.01 – +0.37	+0.49
4	+0.01 – +0.31	+0.24	+0.00 – +0.17	+0.15

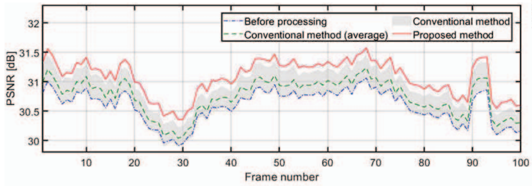


Fig.8. PSNR fluctuations in video 3

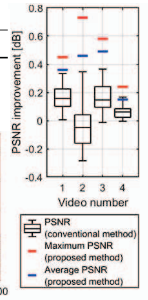


Fig.7. PSNR improvement

the image shown in Fig. 1; here, the brighter is the pixel, the higher is the output value. The image shown in Fig. 4 have real edges such as tree branches and block boundaries that generally appear as a lattice; e.g., the lattice that can be observed in the right upper side of the image. The block boundaries in the edge detection results cause unnecessary deblocking exclusion at the block boundaries in the composition process. Therefore, they must be eliminated. During block boundary elimination, real edges that cross or overlap with the block boundaries are considered.

Fig. 5 shows an example of block boundaries elimination from the edge-detected image. It presents a zoomed-in view of the area near the horizontal block boundary. Each cell represents a pixel and provides the corresponding Sobel filter output values. Two columns in the center show the extracted block boundary together with a real edge diagonally crossing the boundary. To eliminate the block boundaries, we overwrite block boundary pixel values in the red frame with an average of six adjacent pixel values in the blue frame (reference pixel values). In Fig. 5, the values in the red frame are overwritten by the values calculated as (0 + 20 + 20 + 20 + 20 + 0) / 6 = 13. When the reference pixel values are large (indicating the presence of some edges near the block boundary), overwritten values become large; consequently, real edges can be retained. Conversely, when the reference pixel values are small, overwritten values become small; thus, block boundaries are eliminated. The same process is applied to all horizontal and vertical block boundaries. Using the abovementioned process, we can eliminate block boundaries from the edge-detected image while retaining the real edges.

IV. EXPERIMENT AND RESULT

Herein, an experiment was conducted to evaluate the proposed method's effectiveness on deblocking cropped videos. Block noise was generated in the input video sequences by encoding them with H.264/MPEG-4 AVC, followed by cropping. We applied both conventional method by all block positions (8x8=64 patterns) and proposed method to the cropped videos. We used PSNR (Peak Signal to Noise Ratio) for image quality evaluation. Four video sequences comprising 100 frames and 4SIF (704 × 480 pixel) resolutions were used (Fig. 6); panning and camera shaking was included.

Table I shows the maximum and the average PSNR improvement. Column 1 presents the video sequence number. Columns 2-5 list PSNRs or their range measured after being processed by either the conventional method and proposed method. For example, the maximum improvement of PSNR in video 2 stays between +0.00dB and +0.49dB by the conventional method (64 patterns). However, the proposed method improves PSNR to +0.73dB at most. Although the average improvement of PSNR stays between +0.29dB and +0.34dB by the proposed method, the proposed method improves average PSNR to +0.46dB. Also in other video sequences, the proposed method improved the PSNR better than the conventional method. Fig.7 shows PSNR improvement of both methods. Vertical and horizontal axis indicates PSNR improvement and video number, respectively. The average PSNR improvement of the conventional method (64 patterns) is presented by box and whisker plot. Red and blue lines indicate the maximum and average PSNR improvement by the proposed method, respectively. PSNRs of the proposed method are higher than those of the conventional method. Fig.8 shows PSNR fluctuations in video 3. The horizontal and vertical axes indicate the frame number and PSNR. The blue dash-dotted line shows the PSNR before processing, and the red line shows the PSNR of the proposed method. The gray area shows range of the PSNR by 64 patterns of the conventional method, and the green dashed line represents the average of that. In every frame, the proposed method improved the PSNR.

V. DISCUSSION

Experimental results demonstrate that the proposed method improved the PSNR, thereby proving its effectiveness. Where the conventional method's smoothing filter blurred the edges because of the block boundary phase shift, the proposed method retained the edges via block boundary detection and smoothing process exclusion. In particular, the PSNR of video 2 (dusk), which contained many horizontal edges, showed better effects than others. Similar results are expected for videos containing numerous vertical edges. However, the PSNR of video 4 (whale), which image is complicated, stays around +0.2dB improvement. Also, though processing time is about 24 minutes by the conventional method that processed all 64 block position patterns, the proposed method costs only about 50 seconds.

VI. CONCLUSION

In this study, we proposed a novel deblocking method for cropped videos while retaining the edges. From experimental results, we concluded the proposed method is effective. Although, further algorithm improvement is necessary for videos which has a complicated image pattern.

[1] ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services", 2017

[2] A. Norkin, et al., "HEVC Deblocking Filter", IEEE Trans. On Circuits Systems for Video Technology, vol.22, no.12, 2012

[3] Fujitsu, "Block-Noise Detection Hardware, Block-Noise Detection Method, and Block-Noise Detection Program", JP2009270349, 2009

[4] Pooja Sharma, et al., "Different Techniques Of Edge Detection In Digital Image Processing", International Journal of Engineering Research and Applications Vol.3, Issue 3, 2013, pp.458-461

Design of Graph Filter Using Spectral Transformation and Window Method

Chien-Cheng Tseng
Depart. of Computer and Communication Engineering
National Kaohsiung University of Sci. and Tech.
Kaohsiung, Taiwan
tcc@nkust.edu.tw

Su-Ling Lee
Depart. of Computer Science and Information Engineering
Chang-Jung Christian University
Tainan, Taiwan
lilee@mail.cjcu.edu.tw

Abstract—**In this paper, the design of graph filter using spectral transformation and window method is presented. First, the specification of graph filter is converted to the one of digital filter using spectral transformation. Then, the window method is employed to design the digital filter with converted specification. Next, the binomial series expansion is used to obtain the coefficients of the graph filter from those of digital filter. Finally, the numerical design examples are demonstrated to show the effectiveness of the proposed graph filter design method.**

Keywords—graph filter; spectral transformation; digital filter; window method; binomial series expansion

I. INTRODUCTION

In recent years, conventional regular time-domain digital signal processing (DSP) has been extended to irregular vertex-domain graph signal processing (GSP) such that various complex network problems can be solved by using signal processing techniques [1]-[3]. Applications include sensor, transportation, brain, molecular, brain and biological networks. Two basic tools used in GSP are graph Fourier transform (GFT) and graph filter. Thus, it is interesting to study the efficient computation of GFT and the design of graph filter. In this paper, a simple design method of graph filter using spectral transformation will be investigated. Although the least squares (LS) method can be applied to design graph filter, it suffers from the numerical stability problem when filter order is greater than 12. This is because the LS method needs to solve the matrix inversion which is ill-conditioned when filter order is large. The proposed method only uses the window method and spectral transformation to determine filter coefficients, so it is a numerical stable closed-form design and does not need to solve the matrix inversion. Thus, the proposed method can be employed to design a high-order graph filter with small spectral approximation errors.

II. GRAPH FILTER DESIGN

In this section, graph filter design problem is described. A graph $G=(V,\boldsymbol{A})$ is defined by the vertex set $V=\{v_1,v_2,\cdots,v_N\}$ and the $N\times N$ adjacency matrix $\boldsymbol{A}$. If diagonal matrix $\boldsymbol{D}=diag[d_1,d_2,\cdots,d_N]$ is the degree matrix, then the normalized Laplacian matrix is given by $\boldsymbol{L}=\boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$ where $\boldsymbol{I}$ is identity matrix. The eigen-decomposition of the $\boldsymbol{L}$ is $\boldsymbol{L}=\boldsymbol{U\Lambda U}^T$ where unitary matrix $\boldsymbol{U}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_N]$ is composed of the eigenvectors and the diagonal matrix $\boldsymbol{\Lambda}=diag[\lambda_1,\lambda_2,\cdots,\lambda_N]$ is composed of the eigenvalues. The properties $\boldsymbol{UU}^T=\boldsymbol{I}$ and $0\le\lambda_n\le 2$ are satisfied. Let transfer matrix of graph filter be denoted by

$$\boldsymbol{H}=\sum_{k=0}^{M} b_k\boldsymbol{L}^k \tag{1}$$

where M is the order. Using eigen-decomposition, it yields

$$\boldsymbol{H}=\sum_{k=0}^{M} b_k\boldsymbol{U\Lambda}^k\boldsymbol{U}=\sum_{k=0}^{M} h(\lambda_k)\boldsymbol{u}_k\boldsymbol{u}_k^T \tag{2}$$

The above graph filter spectral response $h(\lambda)$ is defined by

$$h(\lambda)=\sum_{k=0}^{M} b_k\lambda^k=\boldsymbol{b}^T\boldsymbol{c}(\lambda) \tag{3}$$

where $\boldsymbol{b}=[b_0,b_1,\cdots,b_M]^T$ and $\boldsymbol{c}(\lambda)=[1,\lambda,\cdots,\lambda^M]^T$. Given the ideal spectral response $h_d(\lambda)$, the graph filter design problem is how to determine the filter coefficients $b_0,b_1,\cdots,b_M$ such that actual response $h(\lambda)$ approximates the ideal response $h_d(\lambda)$ as well as possible for $0\le\lambda\le 2$. If the least squares (LS) method is applied to solve this problem, the filter coefficients can be determined by minimizing the following cost function

$$J=\int_0^2\left(h_d(\lambda)-\boldsymbol{b}^T\boldsymbol{c}(\lambda)\right)^2 d\lambda=\boldsymbol{b}^T\boldsymbol{Q}\boldsymbol{b}-2\boldsymbol{b}^T\boldsymbol{p}+\rho \tag{4}$$

where $\boldsymbol{Q}=\int_0^2\boldsymbol{c}(\lambda)\boldsymbol{c}(\lambda)^T d\lambda$, $\boldsymbol{p}=\int_0^2 h_d(\lambda)\boldsymbol{c}(\lambda)d\lambda$ and $\rho=\int_0^2 h_d(\lambda)^2 d\lambda$. Thus, the optimal solution to minimize the function J is given by $\boldsymbol{b}_{opt}=\boldsymbol{Q}^{-1}\boldsymbol{p}$. Clearly, the LS method needs to solve the matrix inversion. The element of matrix $\boldsymbol{Q}$ at *i*-th row and *j*-th column can be calculated as

$$\boldsymbol{Q}(i,j)=\int_0^2\lambda^{i+j}d\lambda=\frac{2^{i+j+1}}{i+j+1}\qquad 0\le i,j\le M \tag{5}$$

Fig.1(a) shows the logarithm of condition number of matrix $\boldsymbol{Q}$ with respect to filter order M. The condition number is

3.613×10^{18} for $M = 10$. It can be observed that the matrix $\boldsymbol{Q}$ is ill-conditioned when M is large. Fig.1(b) and (c) show the designed results of low-pass filter with ideal response

$$h_d(\lambda) = \begin{cases} 1 & 0 \le \lambda \le \lambda_c \\ 0 & \lambda_c < \lambda \le 2 \end{cases} \tag{6}$$

Because the result for $M = 20$ case is poor, the LS method can be only used in low-order graph filter design. In the following, the spectral transformation method is presented to design high-order graph filter.

III. PROPOSED DESIGN METHOD

Here, the spectral transformation method is used to design low-pass graph filter. Three steps involved are listed below:

Step 1: The frequency λ of graph filter and frequency ω of conventional digital filter can be linked by the equation:

$$\lambda = 1 - \cos(\omega) \tag{7}$$

Fig.2 shows the spectral transformation relation from a graph low-pass filter to a digital low-pass filter using (7). Thus, the low-pass graph filter design problem in (6) can be solved by digital low-pass filter design with ideal response

$$H_d(\omega) = \begin{cases} 1 & 0 \le \omega \le \omega_c \\ 0 & \omega_c < \omega \le \pi \end{cases} \tag{8}$$

Step 2: Using the inverse discrete-time Fourier transform, the ideal impulse response of low-pass filter in (8) is given by

$$h_d(k) = \frac{\sin(\omega_c k)}{\pi k} \tag{9}$$

Using the window method, the transfer function of digital filter is given by

$$H(z) = \sum_{k=-M}^{M} h(k) z^{-k} \tag{10}$$

where filter coefficients are $h(k) = h_d(k) w(k)$. The $w(k)$ is rectangular or Hamming window in [4]. Because $h(k)$ is even symmetric, the frequency response of digital filter is

$$H(e^{j\omega}) = \sum_{k=-M}^{M} h(k) e^{-j\omega k} = \sum_{k=0}^{M} a_k \cos(k\omega) \tag{11}$$

where $a_0 = h(0)$, and $a_k = 2h(k)$ for $k = 1, 2, \cdots, M$. Using the Chebyshev polynomial $T_k(x)$, (11) can be rewritten as

$$H(e^{j\omega}) = \sum_{k=0}^{M} a_k \cos(k\omega) = \sum_{k=0}^{M} a_k T_k(\cos(\omega)) \tag{12}$$

Step 3: Using (7), the graph filter can be obtained from digital filter by the following equation

$$h(\lambda) = H(e^{j\omega})\Big|_{\cos(\omega)=1-\lambda} = \sum_{k=0}^{M} a_k T_k(1-\lambda) \tag{13}$$

Let the Chebyshev polynomial be denoted by

$$T_k(x) = \sum_{n=0}^{k} t_{k,n} x^n \tag{14}$$

then (13) reduces to

$$h(\lambda) = \sum_{k=0}^{M} a_k \sum_{n=0}^{k} t_{k,n} (1-\lambda)^n = \sum_{k=0}^{M} b_k \lambda^k \tag{15}$$

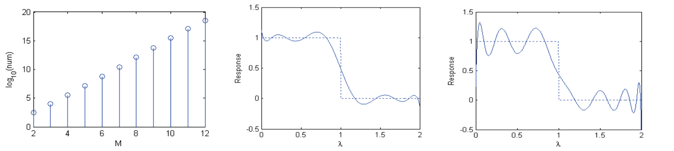


Fig.1 The least squares design method. (a) The logarithm of condition number $\log_{10}(num)$ of matrix $\boldsymbol{Q}$. (b) The spectral response (solid line) of designed filter with $M = 10$ and $\lambda_c = 1$. The dashed line is ideal response. (c) The spectral response (solid line) of designed filter with $M = 20$ and $\lambda_c = 1$. Because matrix $\boldsymbol{Q}$ is ill-conditioned, the obtained spectral response is not good in the case.

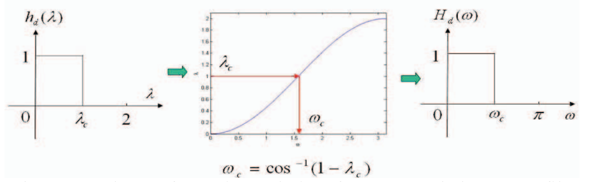


Fig.2 The spectral transformation relation from a graph low-pass filter to a digital low-pass filter using the equation $\lambda = 1 - \cos(\omega)$ in (7).

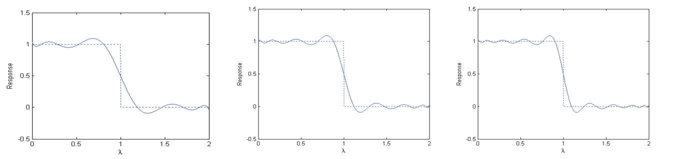


Fig.3 The designed results of the proposed method with $\lambda_c = 1$. (a) The spectral response (solid line) of designed filter with $M = 10$. The dashed line is ideal response. (b) The spectral response of designed filter with $M = 15$. (c) The spectral response of designed filter with $M = 20$.

In (15), using the binomial expansion for the term $(1-\lambda)^n$, the coefficients b_k can be obtained from a_k and $t_{k,n}$. Now, one design example is illustrated to show the effectiveness of the proposed method. If $\lambda_c = 1$ is chosen, then $\omega_c = \frac{\pi}{2}$. Fig.3(a)-(c) show the spectral responses of the designed low-pass graph filter for M=10, 15 and 20. Rectangular window is used here. Clearly, the proposed method can be applied to design graph filter without suffering from the numerical stability problem in the LS design for M=20 case.

IV. CONCLUSIONS

In this paper, the graph filter has been designed by spectral transformation and window method. However, only FIR graph filter design is studied here. Thus, it is interesting to use spectral transformation to design IIR graph filter in the future.

REFERENCES

[1] P.M. Djuric and C. Richard, *Cooperative and Graph Signal Processing: Principles and Applications,* Academic Press, 2018.

[2] B.S. Manoj, A. Chakraborty and R. Singh, *Complex Networks: A Networking and Signal Processing Perspective*, Prentice-Hall, 2018.

[3] A. Ortega, P. Frossard, J. Kovacevic, M.F. Moura and P. Vandergheynst, Graph signal processing: overview, challenges, and applications *Proceedings of the IEEE*, vol.106, no.5, pp.808-826, May 2018.

[4] S.K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, Fourth Edition, McGRAW-Hill 2011.

Pre-Inverse Active Noise Control System with Virtual Sensing Technique for Non-Stationary Path

Kohei Matsuhisa
Graduate School of Engineering
Tottori University
4-101 Koyama-minami, Tottori 680-8552 Tottori, Japan
M19J4046H@edu.tottori-u.ac.jp

Keisuke Okano
Graduate School of Engineering
Tottori University
4-101 Koyama-minami, Tottori 680-8552 Tottori, Japan
d18t2001b@edu.tottori-u.ac.jp

Naoto Sasaoka
Graduate School of Engineering
Tottori University
4-101 Koyama-minami, Tottori 680-8552 Tottori, Japan
sasaoka@tottori-u.ac.jp

Yoshio Itoh
Graduate School of Engineering
Tottori University
4-101 Koyama-minami, Tottori 680-8552 Tottori, Japan
itoh-y@tottori-u.ac.jp

***Abstract*— ANC (Active Noise Control) reduces noise in an acoustic field. In general, the ANC system needs to always place a microphone at the location where you want to reduce noise. Therefore, the ANC cannot achieve sufficiently noise reduction performance far from a microphone. A virtual sensing technique has been proposed to reduce the noise in a zone of quit (ZoQ) where there is no error microphone. Unfortunately, the conventional virtual microphone cannot track the non-stationary path. In this paper, we propose an ANC with virtual sensing to track the non-stationary path between ZoQ and microphone.**

Keywords—active noise control, virtual sensing technique, tertiary path

I. Introduction

ANC has been proposed to reduce noise by giving it the same amplitude and opposite phase simulated noise [1]. In a general ANC system, ZoQ is generated around an error microphone because the ANC system minimizes the power of the signal obtained from the error microphone. Therefore, it is difficult to reduce the noise far from the error microphone. Thus, a virtual sensing technique has been proposed to separate ZoQ from an error microphone[2]-[4]. However, the conventional method is required to estimate the path, which is from an error microphone to ZoQ, by placing a microphone in the ZoQ in advance, and then it is difficult to track the non-stationary path.

Therefore, this paper proposes an ANC with virtual sensing to track a non-stationary path. The proposed method places a virtual microphone between a loudspeaker and an error microphone. As a result, we can configure the structure of an ANC system, which can track the non-stationary path.

II. Proposed Method

Figs. 1 and 2 respectively show the block diagrams in tuning and control stages of a proposed method. The proposed method uses a pre-inverse ANC system that can control the primary and secondary paths separately [5]. $H_p(z)$, $H_S(z)$ and $H_t(z)$ show the transfer functions of the primary, secondary, and tertiary paths, respectively. $\hat{H}'_p(z)$ is the transfer function of the control filter. $\hat{H}_S(z)$ is the transfer function of the secondary pass modeling filter. $\hat{H}'_S(z)$ and $\hat{H}'_t(z)$ is the inverse transfer function of secondary path modeling filter $\hat{H}_S(z)$ and tertiary path $H_t(z)$. $\hat{H}'_S(z)$ is copied as a pre-inverse filter after the control filter $\hat{H}'_p(z)$.

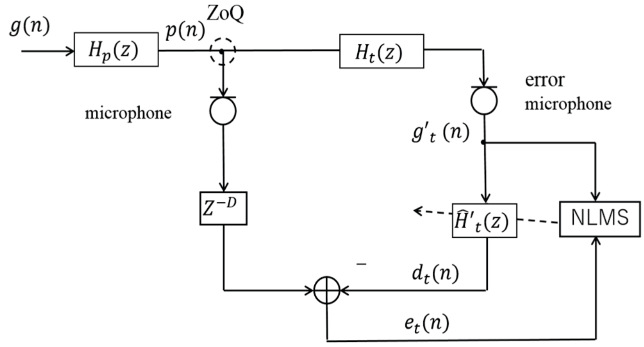


Fig. 1. Structure of the proposed method in the tuning stage.

$g(n)$, and $p(n)$ are respectively the noise source, and the noise passing through the primary path. $d(n)$ is the output signal of the secondary path. $y(n)$ is a cancellation error signal at ZoQ. $y'(n)$ and $y''(n)$ are respectively the input and output of the filter whose transfer function is inverse of the tertiary path. $e(n)$ and $e'(n)$ are the error signals for the secondary path estimation filter $\hat{H}_S(z)$ and the pre-inverse filter $\hat{H}'_S(z)$. $w(n)$ is the auxiliary noise, whose characteristic is white. In the tuning stage of Fig. 2, $g'_t(n)$ and $d_t(z)$ and $e_t(n)$ are the noise detected error microphone, the output signal of $\hat{H}'_t(z)$ and error signal of $\hat{H}'_t(z)$.

In the tuning stage, an error microphone is placed at the ZoQ. This stage estimates the inverse transfer function of the tertiary path $H_t(z)$. An adaptive filter $\hat{H}'_t(z)$ uses a normalized least mean square (NLMS) algorithm. When $\hat{H}'_t(z)$ sufficiently converges, $\hat{H}'_t(z)$ is expressed as

$$\hat{H}'_t(z) = Z^{-D}/H_t\ (z), \tag{2}$$

where Z^{-D} denotes delay. Next, in the control stage, the microphone at ZoQ is removed. Then, $\hat{H}'_t(z)$, which is estimated in the tuning stage, is copied to the filter $\hat{H}'_t(z)$ behind the error microphone in the control stage. If the inverse transfer function of the tertiary path is estimated sufficiently in the tuning stage, the influence of the tertiary path can be canceled from (2). Thus, the signal $y''(n)$ is given by

$$y''(n) = y(n-D). \tag{3}$$

Therefore, the proposed method adds delay Z^{-D} to the input signal of secondary path estimation filter and adaptive

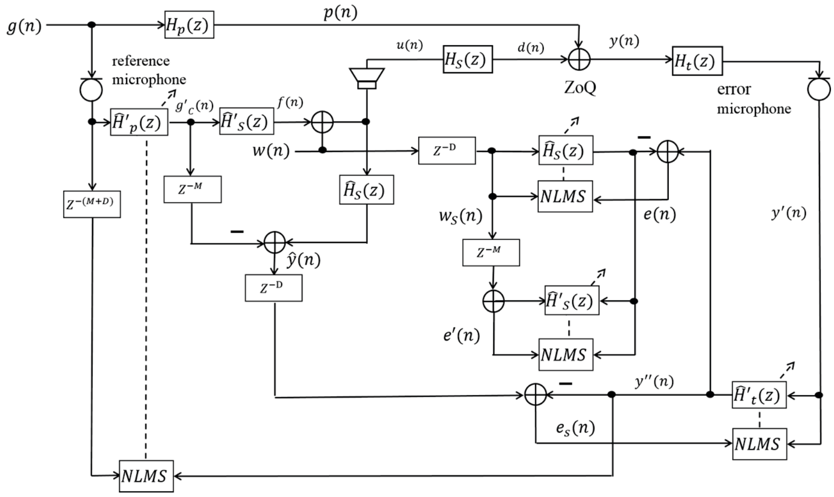


Fig. 2. Structure of the proposed method in the control stage.

algorithm of the control filter. When the ANC system converges sufficiently, $P_{\hat{y}}(n) = P_y$. At this time, ANC system starts updating $\hat{H}'_t(z)$, and stops updating $\hat{H}'_p(z), \hat{H}_S(z), \hat{H}'_S(z)$.

III. Simulation results

The performance of the proposed method is evaluated by computer simulations. The evaluation value (EV) as the following equation is used to assess the noise reduction performance.

$$EV(l) = 10\log_{10}\frac{E\left[\sum_{n=lL}^{(l+1)L-1} y^2(n)\right]}{E\left[\sum_{n=lL}^{(l+1)L-1} p^2(n)\right]}[dB], \quad (4)$$

where L is the block size and set to 256. l is the number of samples and set to 10000. Fig. 3 shows the duct for measuring the impulse responses of the primary, secondary and tertiary paths. In the simulation, the impulse responses of the tertiary path are multiplied by -1 at iteration number l = 5,000 to evaluate the tracking performance of the ANC systems. Fig. 4 shows the simulation result. We show the results by the pre-inverse ANC without virtual sensing, whose error microphone is placed at ZoQ. It can be seen that the proposed method has the potential to reduce the noise, and then the proposed method is comparable to the ANC without virtual sensing. Besides, the proposed method tracks the non-stationary path.

IV. Conclusion

In this paper, we proposes the ANC system with a virtual microphone can track the non-stationary path. The tuning stage estimates the inverse transfer function of the tertiary path, and the control stage tracks the non-stationary path and reduces noise. The simulation results show the effectiveness of the proposed method in a non-stationary environment. In the future, we will evaluate the DSP implementation of the proposed method.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 17K06461.

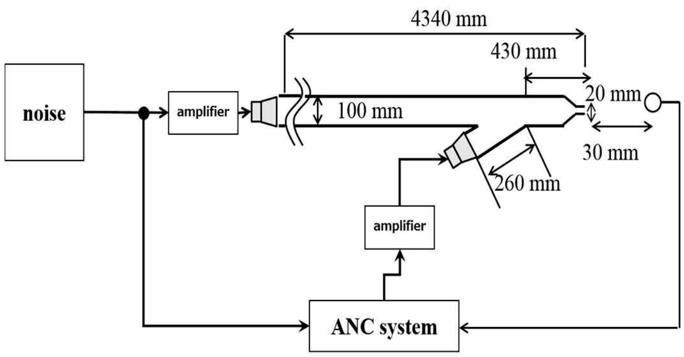


Fig. 3. Duct for detecting acoustic paths

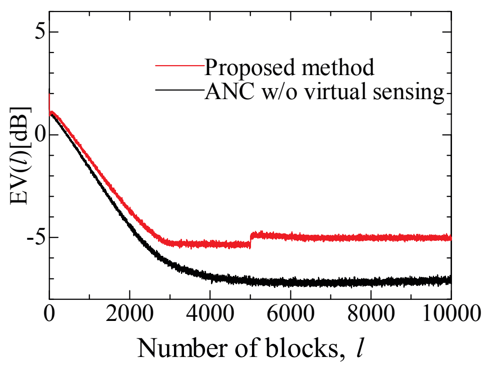


Fig. 4. Simulation result of noise reduction performance.

References

[1] S. M. Kuo, D. R. Morgan, Active Noise Control Systems: Algorithms and DSP Implementations, John Wiley & Sons, New York, 1996.

[2] N. Miyazaki, and Y. Kajikawa, "Head-mounted active noise control system with virtual sensing technique," J. Sound Vib., vol. 339, pp. 65-83, Mar. 2015.

[3] S. Edamoto, C. Shi, and Y. Kajikawa, "Virtual Sensing Technique for Feedforward Active Noise Control," Proc. The 5th Joint Meeting ASA. and ASJ., Honolulu, USD, Nov. 2016

[4] R. Pal, M. K. Sharma, S. Thangjam, "Ambulance Siren Noise Reduction Using Virtual Sensor Based Feedforward ANC System," Second International Conference on Advances in Computing and Communication Engineering, 2015

[5] Y. Tanaka, N. Sasaoka, Y. Itoh and M. Kobayashi, "Active Noise Control with Bias Free Pre-inverse Adaptive System," Proc. 2012 IEEE ISCAS, pp.3222-3225, May 2012.

A Nail Image Analysis Method to Evaluate Accumulated Stress Using Fuzzy Reasoning

Kazuki Shimamoto
Graduate school of Advanced Technology and Science
Tokushima University
Tokushima, Japan
c501837062@tokushima-u.ac.jp

Shin-ichi Ito
Graduate school of Advanced Technology and Science
Tokushima University
Tokushima, Japan
s.ito@tokushima-u.ac.jp

Momoyo Ito
Graduate school of Advanced Technology and Science
Tokushima University
Tokushima, Japan
momoito@is.tokushima-u.ac.jp

Minoru Fukumi
Graduate school of Advanced Technology and Science
Tokushima University
Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

***Abstract*—In this paper, we propose a nail image analysis method to evaluate accumulated stress using fuzzy reasoning. The proposed method consists of three stages: measurement, feature extraction, and stress evaluation. In the measurement, we take a nail image. In the feature detection, we extract the lunula of the nail to calculate its height. In the stress evaluation, we evaluate accumulated stress using fuzzy reasoning. The experimental results suggest that the proposed method can determine the presence or absence of accumulated stress.**

***Index Terms*—fuzzy reasoning, stress evaluation, nail image**

I. Introduction

In recent years, a social problem such as the increase in the number of mental illnesses and suicides have occurred. There have been many studies that measured stress. Most of them were measured stress using blood, saliva, and urine. These samples of stress measurement were difficult to measure accumulated stress because these samples were subject to daily fluctuations and only reflected cortisol levels relative to the minutes to hours prior to collection [1]. Therefore, hair samples and nail samples for measurement stress are considered. In particular, a nail sample is considered to be useful samples for measurement stress because it was easy to collect and had little deterioration [2]. However, a method using the nail sample is difficult to measure stress on a daily basis. Then, we focus on the change of state on the nail. The change of state on the nail was used to confirm the health status and predict disease [3]. Therefore, it has a chance of evaluating the accumulated stress by analyzing the changes of state on the nail. Although the relationship between nail changes and stress is unknown, in this paper, we propose a method to evaluate accumulated stress using the change of nail for college students.

II. Experiments

The proposed method consists of three stages: measurement, feature detection, and stress evaluation.

A. Proposed method

In the measurement, we take nail images of thumb using the iPhone 6s and fix the environment. In the feature detection, we extract the lunula of the nail to calculate its height [4]. In the stress evaluation, we employ a fuzzy reasoning technique. The proposed method creates membership functions for fuzzy reasoning based on the changes in height during experiments. The mean value of the extraction height is calculated weekly. The membership functions are composed of three factors; a difference between max-min height, a sum of the derivative of the change in the mean on a weekly basis, and a degree of the accumulated stress. The difference values between max-min height(Diff) are calculated by computing the difference between the maximum and minimum values of the lunula height during experiments. The proposed method has two membership functions for the Diff. The membership function whether the Diff is large or not(DiffLarge) and the Diff is small or not(DiffSmall) are as follows:

$$DiffLarge = \begin{cases} 0 & Diff \le TH1 \\ slop \times Diff + inter & Diff \le TH2 \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

$$DiffSmall = \begin{cases} 1 & Diff \le TH1 \\ slop \times Diff + inter & Diff \le TH2 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $Diff$, $slop$, and $inter$ are the Diff, the slope of the membership function, and the intercept, respectively. The sum of the derivative of the change in the mean on a weekly basis(SDC) is calculated by computing on mean values of height values during experiments. The proposed method has two membership functions for the SDC. The membership function whether the SDC is large or not(SDCLarge) and the

SDC is small or not(SDCSmall) are as follows:

$$SDCLarge = \begin{cases} 0 & SDC \leq TH1 \\ slop \times SDC + inter & SDC \leq TH2 \\ 1 & \text{otherwise} \end{cases} \quad (3)$$

$$SDCSmall = \begin{cases} 1 & SDC \leq TH1 \\ slop \times SDC + inter & SDC \leq TH2 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where SDC is the sum of the derivative of the change in the mean on a weekly basis. The $slop$ and $inter$ as the same as the variables in equation 1, respectively. Table I shows the values assigned to parameters $slop$, $inter$, $TH1$, and $TH2$, respectively. In the proposed method, four rules are used. Listing 1 shows the source code of the fuzzy rules.

```
NoneGrade = AND(DiffLarge, SDCLarge);
LowGrade = AND(DiffSmall, SDCLarge);
ExiGrade = OR(AND(DiffLarge, SDCSmall), AND
    (DiffSmall, SDCSMall));
```

Listing 1. Fuzzy Rules

where the $DiffLarge$, $DiffSmall$, $SDCLarge$, and $SDCSmall$ are the degree of conformity whether the Diff is large or not, the Diff is small or not, the SDC is large or not, and the SDC is small or not, respectively. The $NoneGrade$, $LowGrade$, and $ExiGrade$ are the degree of conformity: the accumulated stress is little or none, low, and existent, respectively. In the degree of the accumulated stress, the degrees are calculated by computing the center of a trapezoid. Listing 2 shows the source code of center calculation trapezoid. The center of a trapezoid is calculated as follows:

```
for (double i : setSyn)
    num = num + i * size;
    den = den + i;
    size += 0.05;
if (den != 0.0)
    cent = num / den;
else
    cent = 0.0;
```

Listing 2. Center calculation of trapezoid

where $setSyn$ and $cent$ are the synthetic trapezoid and the center of the trapezoid, respectively.

B. Experimental conditions

The subjects were 7 volunteers (mean age: 22.4 years) from Tokushima University in Japan. All subjects consented to take pictures of the nails of the thumb of Dominant hand. The experiments adopted the writing graduation thesis as the stress task because almost all students felt strong stress during making the graduation thesis. Then the volunteers divided into subjects who gave the stress task and subjects who did not give the stress task.

III. Experimental Results and Discussions

Table II shows the Diff, SDC, and accumulated stress degree of each subject. Deg. Stress in Table II indicates the accumulated stress degree. The mean of the reasonable results in the subjects who gave the stress task was 0.66. Then, the mean of the reasonable results in the subjects who did not give stress tasks was 0.44. The Diff of the subject who gave the stress task was larger than the Diff of the subject who did not give the stress task. This result suggested that the change of lunula may occur by giving a stress task. Then, the mean of the accumulated stress degree of the subjects A-D was larger than the mean of the accumulated stress degree of the subjects E-F. These results suggested that the proposed method can classify the subjects who gave the stress task and the subjects who did not give the stress task. Thus, it is considered the proposed method expressed the stress received every day as 0.42-0.46. On the other hand, the stress level of the subjects who gave the stress task was higher than 0.46, because the subjects A-D had the stress received every day in addition to the stress task. Therefore, these results suggest that the proposed method can evaluate the accumulated stress degree.

TABLE I
VALUES ARE SET TO SLOP, INTER, TH1, AND TH2.

Listing	ang	b	TH1	TH2
1	1/30	-2/3	20	50
2	-1/30	5/3	20	50
3	1/16	1/2	-8	8
4	-1/16	1/2	-8	8

TABLE II
DIFF, SDC, AND ACCUMULATED STRESS DEGREE OF EACH SUBJECTS.

Task	Subject	Diff	SDC	Deg. Stress
stress task	A	48.7	-9.3571	0.79
	B	48.3	-8.5547	0.79
	C	37.5	-3.5	0.55
	D	42	-4.57	0.52
nonstress task	E	24.3	3.3215	0.46
	F	24.4	5.5357	0.46
	G	36.5	2.5356	0.42

IV. Conclusion

In this paper, we proposed a method to evaluate accumulated stress for university students using a nail image. The proposed method consisted of three stages: measurement, feature extraction, and stress evaluation. In the measurement, we took a nail image. In the feature detection, we extracted the height of lunula. In the stress evaluation, we evaluated accumulated stress using fuzzy reasoning. The experimental results suggest that the proposed method can determine the presence or absence of accumulated stress.

References

[1] Frug é , Andrew D., et al.: "Fingernail and toenail clippings as a non-invasive measure of chronic cortisol levels in adult cancer survivors." Cancer Causes & Control 29.1, pp.185-191, 2018.

[2] Shuhei IZAWA, Reina YOSHIDA, Masako OHIRA, Ayumi YAMAGUCHI, Shusaku NOMURA: "Quantitative Measurements of Fingernail Cortisol: Effects of Ground-fingernail Grain Size and Extraction Time", Physiological psychology and psychophysiology, 1615tn, 2018.

[3] Basu, Pallavi, and Philip R. Cohen. "Macrolunula: Case Reports of Patients with Trauma-associated Enlarged Lunula and a Concise Review of this Nail Finding" Cureus 10.7 2018.

[4] Kazuki shimamoto, Sin-ichi Ito, Momoyo Ito, Minoru Fukumi: "A Method to Extract change of Lunula of the Nail", Proceedings of SAMCON2019, TT9-2-1 - TT9-2-4, 2019.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

Distributed compressed video sensing based on convolutional sparse coding

Tomohito Mizokami and Yoshimitsu Kuroki
Natinal Institute of Technology (KOSEN), Kurume College, Fukuoka, Japan
a3525tm@kurume.kosen-ac.jp, kuroki@kurume-nct.ac.jp

***Abstract*—This paper discusses a Distributed Compressed Video Sensing (DCVS) framework using Convolutional Sparse Coding (CSC). CSC is a technique to represent a signal as convolutions of filters and corresponding coefficients. Conventional block based DCVS methods divide a given video sequence into key and non-key frames. The key frames are decoded independently like still images, and the non-key frames use Side Information (SI) generated with previously decoded key frames. The sparse dictionaries of the non-key frames are designed with the SIs. However, in CSC based methods, a non-key frame can use the dictionary of the nearest key frame in the temporal domain since the dictionary filters, namely features, are robuster against motions than those of block based methods.**

***Index Terms*—distributed compressed video sensing, dictionary learning, convolutional sparse coding**

I. Introduction

The cardinal features of Distributed Compressed Video Sensing (DCVS) are: (a) Computational burden is transferred from the encoder side to the decoder side; (b) DCVS divides a video sequence into Groups Of Pictures (GOPs) which consist of a key frame and non-key frames and decreases the dimension of images with random projection to send the decoder side; and (c) In the decoder side, the low dimensional signal is recovered with compressed sensing. Please note that the dimension reduction is a kind of compression, but this paper denotes this compression as random projection or dimensional reduction to avoid confusion with compressed sensing. Decoeders of conventional block based DCVS reconstruct the randomly projected non-overlapped blocks of key frames independently. On the other hand, the non-key frames are reconstructed with the dimension reduced signal and Side Information (SI) which is generated with already-decoded key frames to design dictionaries of the non-key frames. This paper focuses on Convolutional Sparse Coding (CSC) [2] for sparse representation of compressed sensing. In CSC, the dictionaries indicate the features of a signal, and the coefficients express the distribution of the features. In adjacent frames, the features do not change drastically, and motions of objects can be expressed with changes of feature distribution. Then, CSC based DCVS can directly apply key frame dictionaries to non-key frames, which means this method does not need SI whereas block based DCVS needs SI to compensate motions of objects. Furthermore, CSC makes no block artifact and increases subjective quality of reconstructed images. This paper narrows the target only non-key frame reconstruction.

II. Preliminaries

Block based DCVS decoders divide each frame into non-overlapped blocks. The i-th block of a non-key frame is denoted as $\boldsymbol{f}_{NK,i}$ and its dimension is decreased with the random projection as $\boldsymbol{y}_{NK,i} = \Phi_{NK}\boldsymbol{f}_{NK,i}$. The encoders only send $\boldsymbol{y}_{NK,i}$ to decoders. The decoders try to reconstruct $\boldsymbol{f}_{NK,i}$ under the assumption that $\boldsymbol{f}_{NK,i}$ is expressed with the sparse dictionary D as $\boldsymbol{f}_{NK,i} = D\boldsymbol{x}_i$; then, the reconstruction is formulated as follows:

$$\begin{aligned}\arg\min_{\boldsymbol{f}_{NK,i},D,\boldsymbol{x}_i} &\frac{1}{2}\left|\left|\boldsymbol{y}_{NK,i} - \Phi_{NK}\boldsymbol{f}_{NK,i}\right|\right|_2^2 + \\ &\frac{\lambda_1}{2}\left|\left|\boldsymbol{f}_{NK,i} - D\boldsymbol{x}_i\right|\right|_2^2 + \lambda_2\left|\left|\boldsymbol{x}_i\right|\right|_1 + \lambda_3 h(\boldsymbol{f}_{NK,i}),\end{aligned} \tag{1}$$

where $\boldsymbol{x}_i$ is a coefficient of D on $\boldsymbol{f}_{NK,i}$ and λ_1, λ_2, and λ_3 are weight parameters. The dictionary D is learned with SI. The last term of the above equation $h(\boldsymbol{f}_{NK,i}) = ||\Psi(\boldsymbol{f}_{NK,i} - \boldsymbol{f}_{SI,i})||_1$ is a regularization term inspired from the frequency correlation noise model [1], where Ψ is the Discrete Cosine Transform (DCT) matrix and $\boldsymbol{f}_{SI,i}$ is the same position block in $\boldsymbol{f}_{SI}$ as $\boldsymbol{f}_{NK,i}$.

In CSC, convolution of dictionary filters $\boldsymbol{d}_m$ and sparse coefficients $\boldsymbol{x}_m$ to express a signal $\boldsymbol{s}$ is shown as

$$\boldsymbol{s} = \sum_m^M \boldsymbol{d}_m * \boldsymbol{x}_m, \tag{2}$$

where M is the number of dictionary filters and coefficients. For a fixed dictionary, the sparse coefficients $\boldsymbol{x}_m$ is estimated from

$$\arg\min_{\boldsymbol{x}_m} \frac{1}{2}\left|\left|\sum_m^M \boldsymbol{d}_m * \boldsymbol{x}_m - \boldsymbol{s}\right|\right|_2^2 + \mu\sum_m^M \left|\left|\boldsymbol{x}_m\right|\right|_1, \tag{3}$$

where μ is a weight parameter. To solve $\boldsymbol{x}_m$ in the Fourier domain, let us define matrixes D_m such that $D_m\boldsymbol{x}_m = \boldsymbol{d}_m * \boldsymbol{x}_m$, and denote D_m, $\boldsymbol{x}_m$, and $\boldsymbol{s}$ in the Fourier domain as $\widehat{D}_m$, $\widehat{\boldsymbol{x}}_m$, and $\widehat{\boldsymbol{s}}$ respectively. By defining

$$\widehat{D} = \begin{pmatrix}\widehat{D}_0 & \widehat{D}_1 \dots\end{pmatrix} \;\; \widehat{\boldsymbol{x}} = \begin{pmatrix}\widehat{\boldsymbol{x}}_0^T & \widehat{\boldsymbol{x}}_1^T \dots\end{pmatrix}^T \;\; \boldsymbol{x} = \begin{pmatrix}\boldsymbol{x}_0^T & \boldsymbol{x}_1^T \dots\end{pmatrix}^T, \tag{4}$$

Eq. (3) is rewritten as

$$\arg\min_x \frac{1}{2}\left|\left|\widehat{D}F\boldsymbol{x} - \widehat{\boldsymbol{s}}\right|\right|_2^2 + \mu\left|\left|\boldsymbol{x}\right|\right|_1, \tag{5}$$

where F is the Fourier transform matrix. The solution of Eq. (5) indicates the distribution of the feature $\boldsymbol{d}_m$ on $\boldsymbol{s}$.

III. Proposed Method

The proposed method replaces linear combination of the dictionary filters and the coefficients in DCVS by CSC. As objects in a non-key frame have the same features as the previous key frame, the dictionary filters of the previous frame are effective to reconstruct the current frame. Thus, Eq. (1) is changed as

$$\underset{\boldsymbol{x}_m}{\arg\min} \frac{1}{2}\left|\left|\Phi \sum_{m}^{M} \boldsymbol{d}_m * \boldsymbol{x}_m - \boldsymbol{y}_{NK}\right|\right|_2^2 + \mu \sum_{m}^{M} ||\boldsymbol{x}_m||_1 , \quad (6)$$

where $\boldsymbol{y}_{NK}$ is the randomly projected whole non-key frame which is not a block-divided, Φ projects an original signal to $\boldsymbol{y}_{NK}$, and $\boldsymbol{d}_m$s are the dictionary filters learned in the reconstructed key frame. The dictionary filters for the non-key frame in Eq. (6) are designed for the previous key frame as follows:

$$\underset{\boldsymbol{d}_m, \boldsymbol{x}_m}{\arg\min} \frac{1}{2}\left|\left|\sum_{m}^{M} \boldsymbol{d}_m * \boldsymbol{x}_m - \boldsymbol{f}_K\right|\right|_2^2 + \mu \sum_{m}^{M} ||\boldsymbol{x}_m||_1 \quad (7)$$
$$\text{s.t.}\ ||\boldsymbol{d}_m||_2 = 1,$$

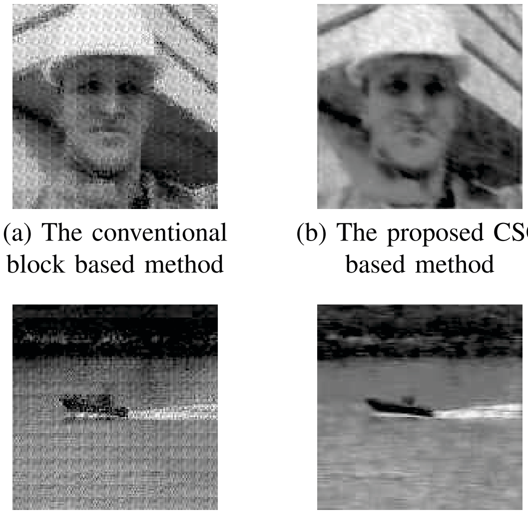
where $\boldsymbol{f}_K$ is the reconstructed key frame. The solution $\boldsymbol{d}_m$ of the above equation should have the same features as the current non-key frame. By using Eq. (5), Eq. (6) is rewritten as

$$\underset{\boldsymbol{x}}{\arg\min} \frac{1}{2}\left|\left|\boldsymbol{y}_{NK} - \Phi F^* \widehat{D} F \boldsymbol{x}\right|\right|_2^2 + \mu ||\boldsymbol{x}||_1 , \quad (8)$$

where F^* is the inverse Fourier transform matrix. Formulation of Eq. (8) is called Least Absolute Shrinkage and Selection Operator (LASSO) which consists of a data-fidelity term and an l_1 norm regularization term. Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [3] is one of iterative methods which can solve LASSO. Then, we solve Eq. (8) with FISTA.

IV. Experiments

To compare the reconstructed non-key frames of the proposed method with those of a conventional method, we conduct experiments with two video sequences "foreman" and "coastguard". Each GOP consists of a key frame and a non-key frame. The key frames are compressed and reconstructed by JPEG to 20% quality, and the compression rates of both Φ and Φ_{NK} are the same at 0.5. We solve Eq. (8) to obtain the solution of coefficients $\boldsymbol{x}$ with FISTA and reconstruct the non-key frame $\boldsymbol{f}_{NK}$ using the solution $\boldsymbol{x}$ and $\boldsymbol{d}_m$ designed for the previous key frame. In experiments, we refer convolutional sparse coding and dictionary learning library "SPORCO" [4] to implement our method and compare the proposed method with the conventional block based method which uses the key frame as SI. In general, block based DCVS makes SI with motion compensation vectors between the previous and the following key frames of the current non-key frame; however, this study refers only the previous key frame as SI to evaluate the robustness of convolutional filters against motions. Figure 1 shows reconstructed images of both methods. The left images include block noise and the boundaries between the blocks appear. On the other hand, the right images have no block noise and preserve smoother texture. Table 1 summarizes average PSNRs of 10 reconstructed non-key frames in each method. The results show that the proposed method reconstructs better euqality non-key frames than the conventional block based method. In particular, CSC represents simple features, e.g. surface of a sea and surface of a wall, well. The above results show that the convolutional dictionaries based DCVS reconstructs the compressed images smoother without block noises than the conventional block based DCVS without the SI.



(a) The conventional block based method (b) The proposed CSC based method

(c) The conventional block based method (d) The proposed CSC based method

Fig. 1: Reconstructed non-key frames

TABLE I: Average PSNR of 10 GOPs [dB]

Dataset	The block based method	The proposed method
foreman	21.41	21.70
coastguard	22.26	24.47

V. Conclusion

This paper proposes a DCVS framework using CSC. The proposed method expresses the motion of the objects with CSC without SI, and the reconstructed frames have no block noise and preserve smoother texture. We confirm that our method is effective to represent simple features in a signal particularly. One of our future works is key frame reconstruction with CSC.

References

[1] F. Tian, J. Guo, B. Song, H.Liu, and H. Qin, "Distributed compressed video sensing with joint optimization of dictionary learning and l_1-analysis based reconstruction," IEICE Trans. inf. & Syst., volE99-D, no.4, pp.1202-1211, 2016. DOI:10.1587/transinf.2015EDP7373.

[2] B.Wohlberg, "Efficient algorithms for convolutional sparse representation," IEEE Trans. on Image Processing, vol.25, no.1, pp.301-315, Jan. 2016.

[3] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imag. Sci., Vol. 2, No.1, pp.183-202, 2009. DOI.10.1137/080716542

[4] B. Wohlberg, "SParse Optimization Research COde (SPORCO)," Softwarelibrary available from http://purl.org/brendt/software/sporco, 2016.

Personal Authentication by Walking Motion using Kinect

Chunyu Guo
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
c501837004@tokushima-u.ac.jp

Shin-ichi Ito
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
s.ito@tokushima-u.ac.jp

Momoyo Ito
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
momoyo@is.tokushima-u.ac.jp

Minoru Fukumi
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

***Abstract*— In recent years, with the rapid development of the information society, the importance of personal authentication has become higher and higher. This paper focuses on the use of a Kinect sensor to obtain walking characteristics for personal authentication. In terms of the proposal method, Kinect is used to obtain body's physical feature quantity, such as the angle of joint bending when a person walks, the displacement of coordinates. In terms of learning recognition, the support vector machine and the obtained feature amount are used for personal authentication. We measured 3 subjects data 5 times a day for 4 days, and obtained an average recognition accuracy of 77.4% using cross-validation.**

Keywords—Personal authentication, Walking motion, Kinect, Support vector machine

I. INTRODUCTION

With the development of the information society, the importance of personal authentication is increasing. Personal authentication is divided into an authentication method using physical characteristics and an authentication method using action characteristics. The most basic behavioral characteristics is the characteristics of walking. Therefore, this paper propose a personal authentication method based on walking movements. The walking based personal authentication directly uses the physical information of avoiding direct contact with the body.

Moreover, in recent years, the performance of cameras has been significantly improved, and information can be captured at a long distance; the price is low, and they can be widely used. This paper uses Kinect V2 to measure and acquire walking motion data. Kinect V2 can obtain the 3 dimensional coordinates of 25 joints [1]. It only uses the information of the three-dimensional position of the human body, and utilizes the angle change of joints and the feature quantity of coordinate transformation which ensures the privacy of information and individuals.

II. PROPOSED METHOD

The proposal method consists of three parts: an input , a pre-processing , and an authentication sections.

A. The input department

The input section uses the Kinect V2 to obtain skeletal information of human body. The acquired information consists of x, y, and z coordinates. This paper used the skeleton information of the 19 red joint points in Fig.1.

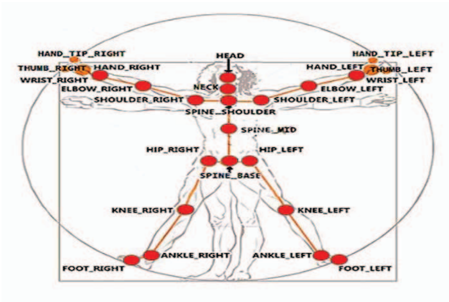


Fig.1 Joint point used [1]

B. The pre-processing section

1) Linear interpolation

In this paper, the linear interpolation is used to unify the data length, and the longest data is used as the standard to unify the length of all used data. Suppose we know the coordinates (x0, y0) and (x1, y1) to obtain the value of a position x in the [x0, x1] interval on the line. Since x is known, the value of y can be obtained from the formula (1).

$$y = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) \tag{1}$$

2) Noise processing

In noise processing we use a moving average filter to smooth the data by replacing each data point with the average of adjacent data points [2]. Smoothing is the removal of unimportant data, leaving important data. The formula is shown in the following.

$$y_{s(i)} = \frac{1}{2N+1}(y(i+N) + y(i+N-1) + \cdots + y(i-N)) \tag{2}$$

i is the number of 3D data, N is the filter size.

3) Average 0

The coordinate values of the data are obtained from the Kinect. The units of the x, y and the z coordinates are different, the unit of the coordinates of x and y are pixel, and the unit of the z coordinate is mm. We think that the same unit is more convenient for observing the distribution and fluctuation of data. Therefore, the unity of units is necessary. This method only changes the size of the data, without changing the law of data distribution and fluctuations.

4) Joint angle feature

The joint angle feature quantity is obtained by taking each joint point as a center point(x1,y1,z1) and then taking two points(x2,y2,z2)，(x3,y3,z3) adjacent thereto to calculate the joint angle according to the space vector formula. The formulas are as follows. We calculate the angle using (5) from (3) and (4)[3]. We used the 19 red joint points marked in Fig.1 to form 20 joint angle features.

$$\vec{a} = (x2-x1, y2-y1, z2-z1) \tag{3}$$

$$\vec{b} = (x3-x1, y3-y1, z3-z1) \tag{4}$$

$$\theta = \cos^{-1}\left(\frac{\vec{a}.\vec{b}}{|\vec{a}||\vec{b}|}\right) \tag{5}$$

5) Frequency domain feature extraction

This section converts the angle of the extracted joint into a frequency domain feature [4]. The amplitude spectrum is obtained using a discrete Fourier transform. The formula is shown in the following.

$$F(\mathrm{t}) = \sum_{x=0}^{N-1} f(x) e^{-i\frac{2\pi t x}{N}} \tag{6}$$

Where N is any natural number, e is the number of Napiers, and i is the imaginary unit.

C. Authentication section

The Authentication section learns and recognizes the joint angle feature amount obtained from the pre-processing section by the support vector machine(SVM).

1) Support Vector Machines

SVM is a pattern recognition model that uses supervised learning. SVM is one of the best discriminative machine learning methods for two class separation of pattern recognition in many methods currently known.

2) Kernel function

The choice of SVM kernel functions is crucial for their performance, especially for linearly inseparable data. In this experiment, we choose the Gaussian radial basis function, which is a locally strong kernel function and can map a sample into a higher dimensional space. This kernel function is the most widely used one. It has a better performance, even with relatively few samples.

III. Experiment

A. Experimental condition

As shown in the above, this paper uses Kinect V2 to measure the trajectory of the three-element coordinates of 19 joint points. The experimental subjects consisted of subjects A, B and C in the age of 20. The experiment was conducted for 4 days. 3 people were tested every day, and each person tested 5 times a day (data on the first day is data1, data on the second day is data2, data on the third day is data3,and data on the fourth day is data4). The Kinect V2 was placed approximately 70 cm from the ground and the subject stood facing it 4 m away from the Kinect V2. The specific experiment is that the experimental subject starts straight from 4 meters, and stops the experiment at 2 meters. The start and end time are manually performed. Fig.4 show the experimental environment.

Fig.4 Experimental environment

B. Cross-validation

Measurement data were divided into learning data and test data. First, we separated the data for each measurement day. The 1-day data were used as test data, the rest day data were as learning data for cross validation. We can find out adversely affecting data in identification accuracy by using this. Table I. shows the combination for cross-validation.

TABLE I. CROSS-VALIDATION

Data	Data1	Data2	Data3	Data4
Text Data	Day1	Day2	Day3	Day4
Learning Data	Day2,3,4	Day1,3,4	Day1,2,4	Day1,2,3

IV. Result

In terms of results, we used the method of cross-validation as shown in Table I. The accuracy of 4 days was obtained and the average accuracy was 77.4%. The result of identification rate obtained by this experiment are shown in Table II.

TABLE II. IDENTIFICATION RATE

Text Data	Day1	Day2	Day3	Day4
Accuracy	77.4%	72.1%	80.3%	79.6%
Average77.4%				

V. Conclusion

In this paper, we proposed a personal authentication method for walking exercise, using Kinect V2 to measure the three-dimensional data of joint points, performing noise processing, linear interpolation, etc. as preprocessing, and calculating the angle of each joint as the feature quantity. The feature values obtained by the preprocessing were learned and identified by the SVM. The accuracy of the experiment averaged 77.4%. The recognition rate of the experiment was not very high. It may be caused by fewer experimental objects and too little experimental data. In addition, when the discrete Fourier transform was carried out, the period of walking was not judged, it maybe cause the low recognition rate. Therefore, we continue to improve our research.

References

[1] “Thorough comparison of Kinect v1 and Kinect v2” https://www.buildinsider.net/small/kinectv2cpp/01

[2] Tera Ando, Mimora Fukumi, Momoyo Ito, Shin-ichi Ito, “ Personal identification for surveillance systems using Kinect” , Proceedings of the Institute of Electrical and Electronics Engineers of Japan, 1-4, September 2018, in Japanese.

[3] Satoshi Yoshida, Masao Izumi, Hiroshi Tsuji, “A research on the ability of Kinect to discriminate people” ,Technical Report of the Institute of Image Information and Television Engineers, Vol.36,No.8,2012,in Japanese

[4] Kenichi Takada, Teruaki kitasuka, Masayoshi Aritsugi, “ A Consideration of an Indiviual Identification Method by Gait Using a Markerless Motion Capture Device ” IPSJ SIG Technical Report,Vol.2012, No.9,pp.1-7,2012,in Japanes

Data Retrieval from Printed Image Using Image Features and Data Embedding

Takuhiro Nishikawa*, Mitsuji Muneyasu*, Yuuki Nishida* Soh Yoshida* and Kosin Chamnongthai†
*Faculty of Engineering Science, Kansai University
3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680 Japan
Email: {muneyasu, k997761, k783893, sohy}@kansai-u.ac.jp
†Faculty of Engineering, King Mongkut's University of Technology Thonburi
126 Pracha-uthit Rd., Bangmond, Tungkru, Bangkok 10140, Thailand
Email: kosin.cha@kmutt.ac.th

***Abstract*—Data retrieval methods from printed images have recently been proposed. However, these methods have problems in terms of data capacity and flexibility. We propose a method of solving these problems using image recognition based on image features and a data-embedding method in the frequency domain. This makes it possible to include different information in the same image and greatly increase the amount of information that can be acquired. In the experiment, we show the effectiveness of the proposed method.**

***Index Terms*—data-embedding, DCT, AKAZE, image recognition, printed image**

I. Introduction

The technology of obtaining information through printed images is attracting attention [1], [2]. Such methods include embedding information in the printed image itself [1] and extracting information by identifying the image using image recognition [2]. The former technique can embed different information even in the same image, but the amount of data obtained from the image is only 14 bytes, which is insufficient for several applications. The insufficient amount of data is because of the severe deterioration of the image resulting from printing of the image and capturing by mobile terminals, and the amount of embedding should be limited to keep the detection rate high. The latter technique can obtain a large amount of information because information is extracted from a database in which the image and the information are related. However, since only the information related to the image in the database can be supplied, different information cannot be obtained from the same image.

In this paper, we propose a method in which a data-embedding technique [1] and an image identification technique based on image recognition [2] are integrated. In this method, image features are extracted from the embedded images and the images are identified by matching the vector consisting of the numbers of keypoints of image features in each local region using a database that has been previously prepared. Then the embedded data is extracted from the image so that different information can be obtained from the same image. Furthermore, a method of improving the data detection rate of the data-embedded image using information of an original image obtained by image identification is also considered. Finally, experimental results show the effectiveness of the proposed method.

II. Proposed method

A. Embedding algorithm

Figure 1 shows an overview of the data-embedding algorithm in the proposed method. The data-embedding is the same as in the method described in [1]. First, the number of block divisions is set to N,

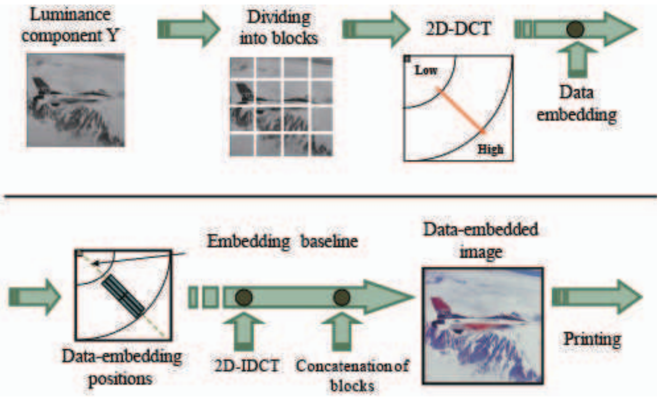


Fig. 1. Overview of the proposed embedding algorithm.

and the luminance component of the original image is divided into blocks of $N \times N$. After a two-dimensional discrete cosine transform (2D-DCT) is performed on each block, data of 7 bits are embedded into each block. From the data, the bit sequence to be embedded is generated by a method based on multiplexing of Walsh codes of 64 bits. Data-embedding is performed by adding the generated bit sequence multiplied by a gain to the DCT coefficients at a specific position of the block in the original image. This gain controls the image quality and detection accuracy. When the embedding is complete, an inverse 2D-DCT is performed, and each block is integrated to obtain an image in which data is embedded. Hereinafter, this image is referred to as an embedded image. Also, a frame border is added around the embedded image to be used for image correction during detection.

B. Detection algorithm

Figure 2 shows an overview of the detection algorithm in the proposed method. The detection algorithm first performs frame detection for the luminance component of a captured image. A corrected image is obtained by restoring the geometric shape of the image by the correction process. The correction process consists of lens distortion correction and projective transformation. After the image correction, the frame border is removed.

First, the corrected image is converted to a predetermined resolution (600×600) for feature detection and then divided into blocks of 8×8. The image feature, AKAZE (Accelerated KAZE) [3], is used for each block to calculate local features and to obtain a vector in which the numbers of obtained keypoints are arranged in raster scan order. The correlation coefficient between this vector and the feature

Part of this research was financially supported by a JSPS Grant-in-Aid for Scientific Research (C) (No. 17K06450).

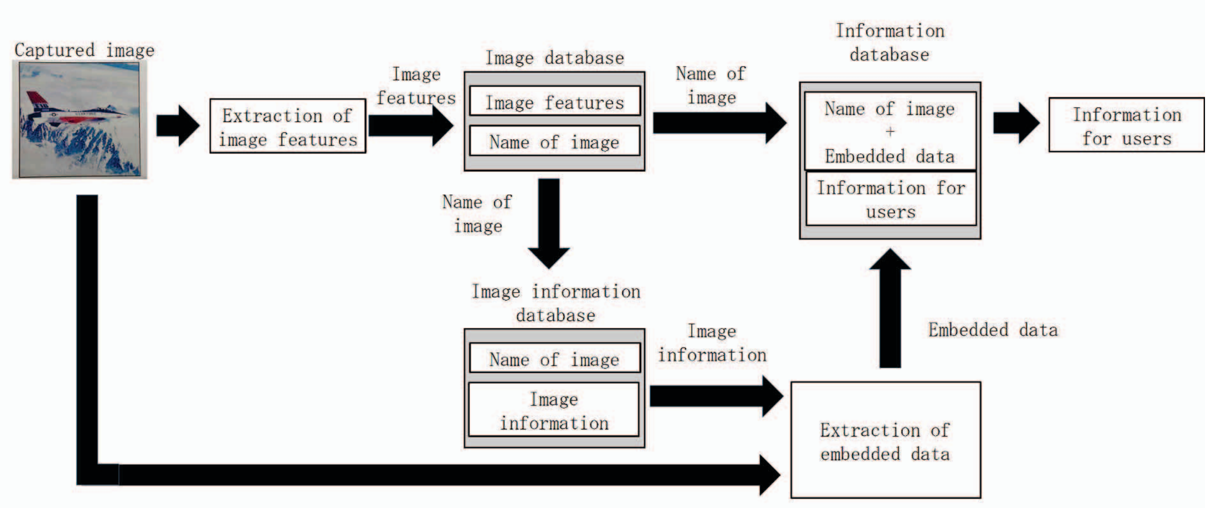


Fig. 2. Overview of the proposed detection algorithm.

vector stored in the database in advance is calculated, and the image name with the maximum correlation coefficient is extracted from the image database.

We also extract the embedded data from the captured image. Here, we use a modified version of the method described in [1]. The corrected image is converted to the same resolution as that when the data was embedded. This image is divided by the same number of blocks as that when the data was embedded, and a 2D-DCT is performed on each block. The DCT coefficients at the position where the bit sequence is embedded are extracted from each DCT-transformed block and are denoted by D_W.

In this method, since the image name is identified by feature detection, the DCT coefficients of the original image, which act as noise, can be used in the detection. This information is stored in the image information database and extracted as image information.

The detection of the 7-bit data is accomplished by calculating the correlation between the extracted DCT coefficients and the multiplexed Walsh codes in the same way as in [1]. Here, instead of D_W, we use $\hat{D}_W$, which reduces the influence of the original image. $\hat{D}_W$ is calculated by

$$\hat{D}_W = D_W - \alpha D, \quad (1)$$

where D is the DCT coefficient of the original image and α is the scaling parameter determined by the kind of image and the average intensity of the captured image.

III. Experiment

To verify the effectiveness of the proposed method, an experiment was performed on six color images (512 × 512 pixels, 256 gray levels for each color). The following equipment was used for the experiment: a Canon LBP9600C printer and a Samsung SM-T700 tablet for image capture. The data to be embedded was the same, 1110000, for each block. The experiment was performed ten times for each image. We evaluate the detection performance of the embedded data using the detection rate D_r:

$$D_r[\%] = \frac{B_d}{B_t} \times 100, \quad (2)$$

where B_d is the number of bits detected correctly and B_t is the number of embedded bits (in this case, a total of 112 bits are embedded).

TABLE I
Average detection rates and processing times.

	detection rate (%)		processing time (ms)	
Image	[1]	proposed	[1]	proposed
Airplane	95.88	98.04	792	1786
Lenna	94.64	99.64	625	1689
Mandrill	96.79	99.11	977	1913
Milkdrop	93.21	95.71	560	1813
Pepper	97.23	98.57	610	1650
Sailboat	78.30	96.70	573	1539

Table I shows the average detection rates of 10 trials by the conventional [1] and proposed methods and each processing time. In the experiment, the names of all images were correctly identified by the image recognition method in the proposed method, and image information was extracted successfully. From this table, we see that the average detection rates of the proposed method are superior to those of the conventional method with reasonable processing times.

IV. Conclusions

We proposed a method for acquiring data from images by integrating data-embedding technology and image identification technology using image features. In the proposed method, a system that can present different information to the user even with the same image can be constructed by integrating the two technologies. In addition, on the premise of a database search for image identification, we can reduce the effect of the DCT coefficients of the original image, which act as noise in embedded data detection. The experimental result shows the effectiveness of the proposed method.

References

[1] A. Hiyama and M. Muneyasu, “Improved Method of Detecting Data in Data-Embedded Printed Images Considering Mobile Devices,” IEICE Trans. Fundamentals. Vol. E99-A, No. 11, pp. 2000–2002, 2016.
[2] R. Mori and M. Muneyasu, “Discrimination of Printed Images Using Image Features and Its Application,” IEICE Technical Report, Vol. 116, No. 482, SIS2016–45, pp. 19–24, 2017 (in Japanese).
[3] P. Alcantarilla, J. Nuevo and A. Bartoli, “Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces,” Proceedings of the British Machine Vision Conference 2013, pp. 13.1–13.11, 2013.

Study on Discrimination of Finger Motions based on EMG signals

Xiao Shan
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
c501837048@tokushima-u.ac.jp

Shin-ichi Ito
Information Science and Intelligent Systems
Tokushima University
Tokusihima, Japan
s.ito@tokushima-u.ac.jp

Momoyo Ito
Information Science and Intelligent Systems
Tokusima University
Tookushima, Japan
momoito@is.tokushima-u.ac.jp

Minoru Fukumi
Information Science and Intelligent Systems
Tokushima University
Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

***Abstract*—In recent years, biological signals have attracted attention as tools for human interfaces. Researches on biological signals have been actively conducted. In this paper, we propose a method which distinguishes ten motions, such as "One" "Two" "Three" "Four" "Five" "Six" "Seven" "Eight" "Nine" and "Ten" by measured the electromyogram of the wrist. We measure data by installing 8 dry type sensors on the right wrist. We carry out frequency analysis using FFT and try to take 3 kinds of methods to remove noise. Finally, we use Support Vector Machine (SVM) for identification and classification. We conducted experiments with four subjects. In the experimental result, the accuracy of finger motions recognition was 65%. In the future, we will also add more methods to remove noise, and try to find other methods to improve the accuracy in the research.**

Keywords—* *Biological signals, Electromyogram, Support Vector Machine

I. INTRODUCTION

Recent years, bio-signals are receiving attention as a tool of human interfaces. Biological signals such as brain waves, the pulse wave and electromyogram (EMG) have been actively researched. Above all, EMG has already applied to many researches. Most of the researches of EMG have measured the EMG on the shoulders or on the arms because there are a lot of muscles [1]. However, in daily life, it is inconvenient to carry out experiments because there are clothes on the shoulders or arms. In addition, wet type sensors were often used to measure EMG. If wet type sensors are dried, those need to be changed to new ones [2]. Therefore, wet type sensors take a high cost. From these background, in this paper we measure EMG by attaching dry type sensors to wrist, and then distinguish ten motions, such as "One" "Two" "Three "Four" "Five" "Six" "Seven" "Eight" "Nine" and "Ten" by using them (Fig.1). Our final purpose is the detailed discrimination of finger motions by using wrist EMG.

II. PROPOSED METHOD

The proposed method is divided into three parts: an input, a pre-processing and a discrimination section(Fig.2). In this paper, we implement the proposed method by Python programming language.

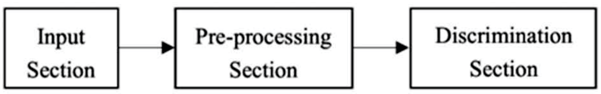


Fig.2. Flow of the Proposed Method

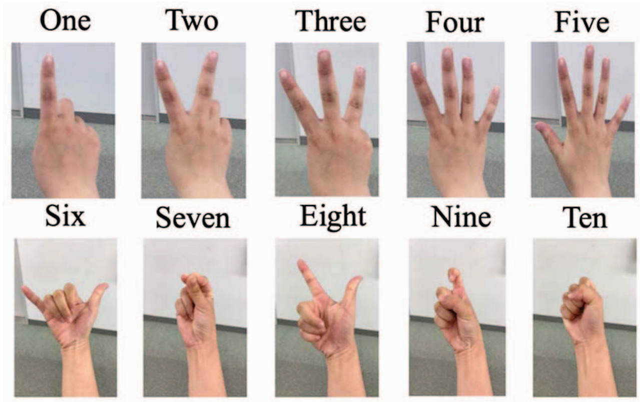


Fig.1. Discrimination of finger motions

A. Input section

The Input section of the proposed method measures EMG by 8 channels of dry type sensors. We attach sensors around a wrist (Fig.3). We use P-EMG plus (Fig.4) for measuring EMG. Sampling rate is 1 kHz in measuring EMG.

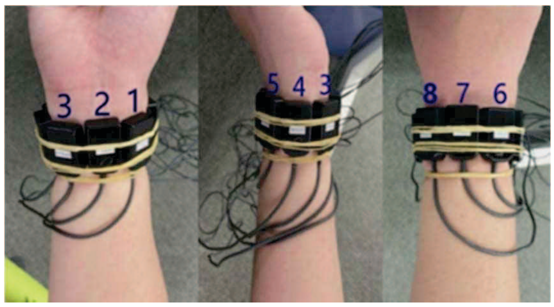


Fig.3. Sensor Position of the Proposed Method.

Fig.4. P-EMG plus.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

B. *Preprocessing Section*

In the preprocessing section, we mainly deal with noise in EMG data. First, in order to observe the noise more convenient, we carry out frequency analysis using FFT. Second, hum noise is mixed from AC power source, and it is 60 Hz in West Japan. Therefore, we change a value of about 60 Hz of FFT spectra to 0. Third, we can clearly observe the drift noise below 20 frequencies, so we apply a high pass filter to remove the noise below 20 frequencies. We judge outliers on the basis of the following formula.

$$x > a + 3b \qquad (1)$$

where "a" means an average value and "b" means standard deviation. The average value and the standard deviation are calculated using values in 1,024 points. The average value is substituted for outliers. Finally, we use inverse FFT [2].

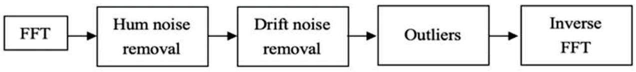


Fig.5.The flow of pre-processing section.

C. *Discrimination*

The discrimination section, we use SVM, which divides finger motions into ten categories: “One” “Two” “Three” “Four” “Five” “Six” “Seven” “Eight” “Nine” and “Ten”.

III. EXPERIMENT

The subjects of experiments are 5 people. The subjects repeated the finger motions 10 times about every three seconds. We conducted this measurement for 3 days. Therefore, we gathered 1500 data in total. EMG data used for experiments is a part of each time series data in Fig.6, which shows how to cut out identification data.

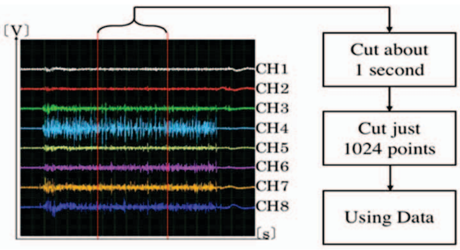
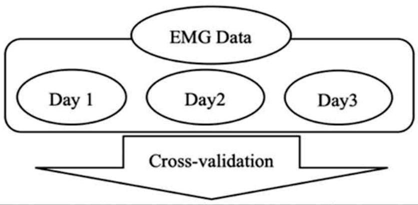


Fig.6. How to cut out identification data.

In addition, measurement data is divided into learning data and test data. First, we separate the data for each measurement day. The 1-day data are used as test data, the rest day data are as learning data for cross validation[4].



Dataset Name	Data 1	Data 2	Data 3
Test Data	Day 1	Day 2	Day 3
Learning Data	Day 2+3	Day 1+3	Day 2+3

Fig.7. Cross validation

IV. RESULT AND CONSIDERATION

A. *Result*

Learning time of SVM took about 2.0 seconds. Table 1 shows the results of finger motions recognition.

Table.1 The results

Subject	Dataset 1	Dataset 2	Dataset 3	AVE
A	50%	57%	65%	57%
B	65%	72%	68%	68%
C	67%	70%	62%	63%
D	68%	60%	70%	66%
E	70%	75%	68%	71%
AVE	64%	61%	67%	65%

B. *Consideration*

First, the accuracy of finger motions recognition was low, we think its reasons are the number of parameters is large and poor learning.

Second, we think there are still much noise in the EMG data. We guess that there are many reasons. For example, while measuring data, all of the circumstances, the subject’s sitting position, the sensitivity of the sensors and the state of the experiment, can cause noise included in the EMG data.

In future work, we will try to use multi-class classification in consideration of versatility and consider using a layer structure to improve the difficulty level of learning.

V. CONCLUTION

In this paper, we proposed a method which could discriminate finger motions by using the wrist EMG. We focused on ten motions. The proposed method was divided into three parts: an input, a pre-processing and a learning recognition sections. In the input section, we measure EMG data by installing 8 dry type sensors on the wrist and P-EMG plus. In the pre-processing section, we have tried to take 3 kinds of methods to remove noise.

In the learning and recognition section, we used support vector machine (SVM). We conducted experiments with five subjects. In the experimental result, the accuracy of finger motions recognition was 65%. In future work, we will add the number of subjects, create a layer structure of CNN and think out more methods to remove noise.

REFERENCES

[1] Daiki Hiraoka, Momoyo Ito, Shin-ichi Ito, Minoru Fukumi, “Japanese Janken Recognition Based on Wrist EMG Analysis by CNN and SVM”, 4th International Conference on Advanced Technology & Sciences (ICAT’Rome), Rome, Italy, pp.323–328, Novemver 23–25, 2016

[2] Ryohei Shioji, Shin-ichi Ito, Momoyo Ito and Minoru Fukumi, “Personal Authentication Based on Wrist EMG Abalysis by a Convolutional Neural Network”, Proceedings of the 5th IIAE International Conference on Intelligent Systems and Image Processing 2017, Hilton Waikiki, Hawaii, USA, pp.12–18, September 7–12, 2017

[3] Ryohei Shioji, Shin-ichi Ito, Momoyo Ito and Minoru Fukumi, “Personal Authentication and Hand Motion Recognition based on Wrist EMG Analysis by a Convolutional Neural Network”, SCIS&ISIS 2018, Toyama, Japan, pp.–, December 5–8, 2018

[4] Takahide Funabashi, Yohei Takeuchi, Momoyo Ito, Koji Kashihara and Minoru Fukumi: Recognition of Finger Motion by Wrist EMG, Proceeding of 2012 International Workshop on Nonlinear Circuits, Communication and Signal Processing NCSP’12, p.433-436, Honolulu, Mar, 2012.

Residual Concatenated Network for ODBTC Image Restoration

Alim Wicaksono
Department of Informatics
Universitas Sebelas Maret
Surakarta, Indonesia
wicayudha.wy@gmail.com

Heri Prasetyo
Department of Informatics
Universitas Sebelas Maret
Surakarta, Indonesia
heri.prasetyo@staff.uns.ac.id

Jing-Ming Guo
Department of Electrical Engineering
National Taiwan Univ. Sci. Technol.
Taipei, Taiwan
jmguo@seed.net.tw

***Abstract*—This paper proposes Residual Concatenated Network (RCN) for improving the quality of Ordered Dither Block Truncation Coding (ODBTC) decoded image. This method inherits the effectiveness of Convolutional Neural Networks (CNN) for suppressing the impulsive noise occurred in decoded image. It suppresses the noise by applying a series of convolutional operations. The network directly performs learning process via an end-to-end mapping approach. The experimental results reveal that the proposed approach yields a promising result in the ODBTC image restoration.**

Keywords—image restoration, odbtc, deep learning, convolutional neural network

I. INTRODUCTION

The ODBTC method compresses a grayscale image in efficient way [1,2]. To achieves good compression quality, this method transforms image into bitmap image and two extreme color quantizers. The bitmap image introduces the visual illusion which effectively removes the false contour and blocking artifacts by utilizing the dithered halftoning image. Two extreme color quantizers replace the bitmap image at decoding stage. Those two representations reduce the required bits for storing an images.

In image compression task, the ODBTC decoded image often produces a slight unpleasant result. Since of ODBTC nature, the quality of decoded image is significantly degraded if we increase the size of image block. In addition, the impulsive-noise also occurs in the ODBTC decoded image due to the dithering process. Thus, to suppress the noise occurrence on ODBTC decoded image, the ODBTC image restoration is proposed in this paper.

II. DEEP LEARNING BASED IMAGE RESTORATION

This section presents the ODBTC image restoration using the deep learning-based approach. The proposed method exploits CNN for extracting the information of ODBTC decoded image, as well as for suppressing the occurred impulsive noise. This noise induces an unsatisfactory image quality. This noise is a result of the halftoning process of an input image.

A. Residual Concatenated Network (RCN)

The RCN consists of multiple convolution layers. This network concatenates each output features from previous layers, as input features, to the current layer. Suppose that a RCN composes of L convolution layers. The input features are denoted as x. The output features of d-th convolution layer in RCN can be formulated as

$$x_{d+1} = \delta(W_d * [x_d, x_{d-1}, \dots] + b_d), \tag{1}$$

where $d = 0, 1, 2, \dots, L$. δ denotes the Leaky-ReLU activation function. W_d and b_d are the weights and biases of the d -th convolution layer, respectively. $[x_d, x_{d-1}, \dots]$ denotes the features concatenation operation. The first convolution layer produces k output features, while the d-th convolution layer generates $k \times d$ features, except the L-th convolution layer output features. The number of features in the L-th convolution layer is same as x_0.

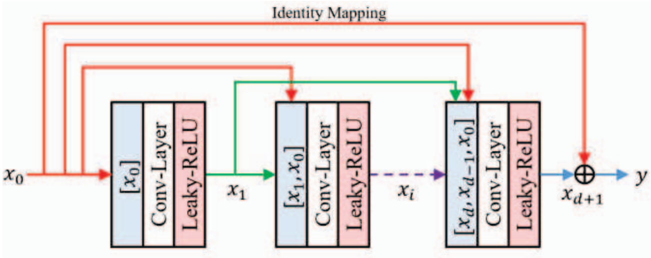


Fig. 1. Residual Concatenated Network (RCN) architecture

Subsequentlt, the output features of convolution d-th layer receive residual information from initial input features via shortcut connection which is an identity mapping as

$$y = x_L + x_0, \tag{2}$$

B. Residual Network of RCN (RRCN)

The RRCN requires an additional shortcuts connection between RCN modules. It has purpose to help the flow of information to the deeper layers of the network. In this architecture, we use identity mapping as shortcut connection.

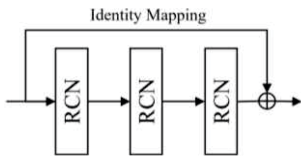


Fig. 2. Residual Network of RCN (RRCN) architecture

Thus, the output of L RRCN modules can be formulated as

$$y_L = y_l + \sum_{i=l}^{L-1} \mathcal{R}(y_i), \tag{3}$$

where $\mathcal{R}$ denotes the RRCN module. Also, the additional shortcut connection can help the flow of gradients during training process [3].

C. Network Architecture

This network consists of four parts, namely shallow feature extractor, deep feature extractor, deep feature reconstructor, and shallow feature reconstructor. The shallow feature extractor contains single convolution layer followed by Leaky-ReLU. The deep feature extractor has the RRCN modules with dilated kernels and downsampling operator, while the deep feature reconstructor owns the RRCN modules with regular kernels and upsampling operator. Whereas, the shallow feature reconstructor consists of two convolution layers followed by Leaky-ReLU at each layer

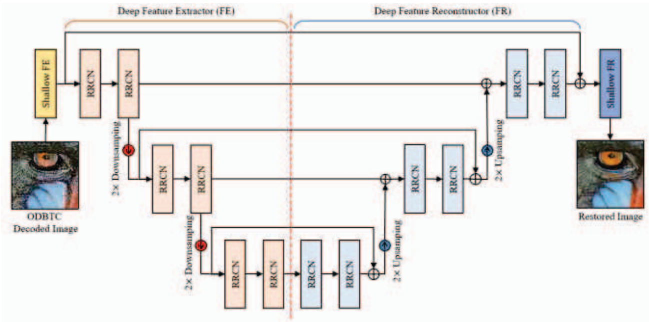


Fig. 3. Network for ODBTC Image Restoration

In this scenario, the downsampling operation is performed with single convolution layer with two stride units, while the upsampling operation is conducted with bilinear interpolation method followed by the convolution layer. The additional shortcut connections are also appended for connecting the features after and before entering the RRCN modules on each level of network.

D. Loss Function

In the end-to-end mapping learning, the network $\mathcal{F}$ optimizes its parameters by minimizing the pixel difference between the output image obtained from the network and the original training images. We utilize the Mean Squared Error (MSE) as the loss function, defined as follow:

$$L = \frac{1}{n}\sum_{i=1}^{n}\left\|\mathcal{F}\left(I_i^{odbtc}\right) - I_i^{original}\right\|^2, \tag{4}$$

where n is the number of training samples. Minimizing the loss function between two images (I_i^{odbtc} and $I_i^{original}$) induces the increased PSNR values.

III. EXPERIMENTAL RESULTS

This section reports the experiment results to examine and investigate the usability of the proposed image restoration method. The experiments were conducted on color images. The performance of the proposed method is objectively measured in terms of Peak Signal to Noise Ratio (PSNR) and Structural SIMilarity (SSIM) index. The PSNR is measured by averaging the PNSR scores over red, green, and blue color bands.

A. Experimental Setup

In the training process, the proposed network employs the DIV2K image dataset [4]. To generate training dataset, all images are firstly compressed with ODBTC, in which each compressed image is further divided as non-overlapping block of size 128×128 pixels. Herein, the block sizes of ODBTC are set as 8×8 and 16×16. For testing purpose, six teen standard images are investigated. The proposed network consists of three convolutional layers for RCN with kernel size 3×3 and the number of features is 32. The leak factor on Leaky-ReLU is set on 0.01. On Deep Feature extractor, we use 3, 2, 1 dilatation factors for each convolutional layer of RCN.

B. Performance of the Proposed Image Restoration

This subsection reports the performance of the proposed ODBTC image restoration. Fig. 4 shows the visual investigation of the proposed method on testing images.

Fig. 4. Visual evaluation on test image. From left to right are ODBTC decoded images, restored images using the proposed method, and original images. The ODBTC block size is 16×16.

Table I summarizes the performance comparisons with and without our proposed image restoration. As shown in the table, the proposed method yields higher average PSNR dan SSIM values. This finding indicates the quality improvement of ODBTC decoded images.

TABLE I. EVALUATION ON PROPOSED METHOD

Block Size	ODBTC		Wavelet [5]		Proposed	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
8×8	20.39	0.7439	25.11	0.8632	**29.59**	**0.9438**
16×16	14.64	0.5679	23.56	0.8118	**27.45**	**0.9170**

IV. CONCLUTIONS

A deep learning-based approach for restoring ODBTC decoded image has been presented in this paper. This method exploits the CNN to extract and to remove the impulsive noise. The proposed method offers a promising result on restoring the ODBTC decoded image.

REFERENCES

[1] J.M. Guo, and M.F. Wu, "Improved block truncation coding based on the void-and-cluster dithering approach," *IEEE Trans. Image Process*., vol. 18, no. 1, pp. 211-213, 2009.

[2] J.M. Guo, "High efficiency ordered dither block truncation with dither array LUT and its scalable coding application," *Digital Signal Process*., vol. 20, no. 1, pp. 97-110, 2010.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *arXiv:1512.03385 [cs]*, Dec. 2015

[4] R. Timofte *et al.*, "NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 1110–1121.

[5] H. Prasetyo and D. Riyono, "Wavelet-Based ODBTC Image Reconstruction," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 10, no. 2, pp. 95–99–99, Jul. 2018.

Trend Prediction of Influenza and the Associated Pneumonia in Taiwan Using Machine Learning

Sing-Ling Jhuo, Mi-Tren Hsieh, Ting-Chien Weng, Mei-Juan Chen and Chieh-Ming Yang
Department of Electrical Engineering
National Dong Hwa University
Hualien, Taiwan
E-mail: 410523006@gms.ndhu.edu.tw, 410423015@gms.ndhu.edu.tw
610723021@gms.ndhu.edu.tw, cmj@gms.ndhu.edu.tw
j82887@gmail.com

Chia-Hung Yeh
Department of Electrical Engineering
National Taiwan Normal University
Taipei, Taiwan
Department of Electrical Engineering
National Sun Yat-sen University
Kaohsiung, Taiwan
E-mail: chyeh@ntnu.edu.tw

***Abstract*—Trend prediction of influenza and the associated pneumonia can provide the information for taking preventive actions for public health. This paper uses meteorological and pollution parameters, and acute upper respiratory infection (AURI) outpatient number as input to multilayer perceptron (MLP) to predict the patient number of influenza and the associated pneumonia in the following week. The meteorological parameters in use are temperature and relative humidity, air pollution parameters are Particulate Matter 2.5 (PM 2.5) and Carbon Monoxide (CO), and the patient prediction includes both outpatients and inpatients. Patients are classified by tertiles into three categories: high, moderate, and low volumes. In the nationwide data analysis, the proposed method using MLP machine learning can reach the accuracy of 81.16% for the elderly population and 77.54% for overall population in Taiwan. The regional data analyses with various age groups are also provided in this paper.**

Keywords—influenza, flu, pneumonia, machine learning, multi-layer perceptron

I. Introduction

Influenza (flu) is a worldwide notorious viral respiratory infectious disease, commonly manifested by high spike fever, muscle soreness, and acute upper respiratory tract symptoms. There are four types of Influenza, identified as A, B, C and D, which are from RNA virus family. A mutation of the RNA may result in the influenza pandemic every 10-year[1]. Other viral infections and environmental factors such as air pollutants may contribute to the influenza viral mutation and spread [2]. Therefore, trend prediction of the epidemics of influenza and the associated pneumonia regionally and nationally can provide valuable information for taking preventive actions for public health [3].

Machine learning using big data training has been applied to predict clinical diseases [3][4]. This paper uses meteorological parameters, temperature and relative humidity, and air pollution parameters, PM 2.5 and CO, and the number of outpatients with AURI as the input data of multilayer perceptron (MLP) to relate to the tendency of the subsequent one-week outpatients and inpatients of influenza and the associated pneumonia in Taiwan. The regional data analyses with various age groups are also provided in this paper.

II. Materials And Methods

A. Data sets

The data of the temperature, relative humidity, PM2.5 and CO are from air quality monitoring data of the open website of Environmental Protection Administration in Taiwan [5]. The number of outpatients with AURI and the number of outpatients and inpatients of flu and the associated pneumonia are provided by the open website of Taiwan Centers for Diseases Control [6]. We use the temperature, relative humidity, PM2.5, CO and the number of AURI from Dec. 2009 to Dec. 2017 as input and the outpatients and inpatients volumes with flu and the associated pneumonia from Jan. 2010 to Jan. 2018 as the outcome. The distribution for the average weekly number of the patients is shown in Fig.1. In Taiwan, the flu is prevailing in winter and patient number increases gradually since Oct. and declines after March. According to the number of patients, we divide the number of outpatients and inpatients with flu and the associated pneumonia into three levels: low(0), moderate(1), high(2); the patient distribution of Taiwan and Eastern Taiwan (Yilan, Hualien and Taitung) are listed in Table I. We use the middle month of each season (January, April, July, and October) as the test data (33%) and the other months as the training data(66%).

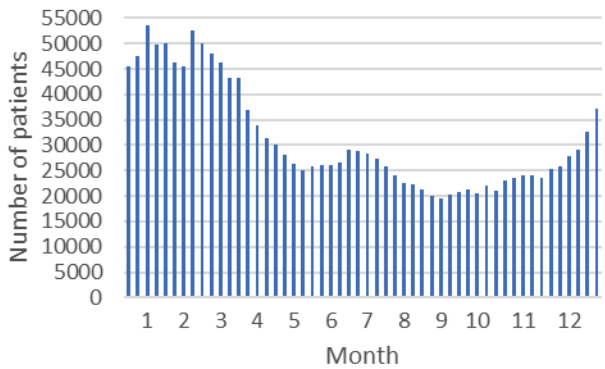


Fig. 1. Average weekly number of the outpatients and inpatients with flu and the associated pneumonia from Jan. 2010 to Jan. 2018 in Taiwan

B. Multilayer Perceptron(MLP)

Figure 2 illustrates the multi-layer perceptron (MLP) architecture with a forward and backward propagation learning algorithm used in this paper. The input parameters includes month (x_1) and week (x_2) of a year for the predicted week, the outpatients number of AURI in the previous week (x_3), daily average data of temperature, relative humidity, PM2.5, and CO for 30 consecutive days(x_4 to x_{123}) before the predicted week. The output is the trend prediction (low, moderate or high volume) of the number of outpatients and inpatients of flu and the associated pneumonia for the subsequent week. Taking a week of patients is because of the

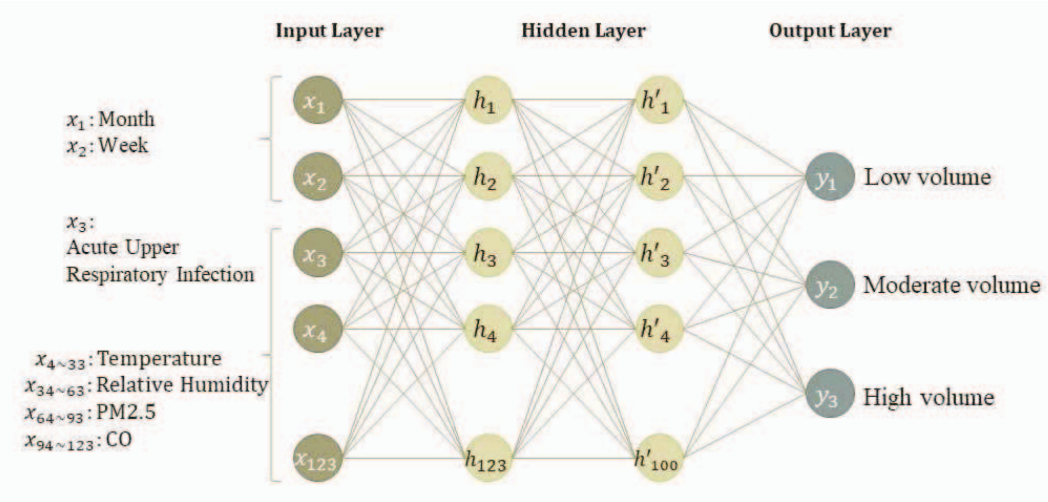


Fig. 2. Multilayer Perceptron (MLP) diagram for the proposed method

time lag effect of patients seeking the medical treatment. In the proposed MLP model, there are two hidden layers composed of 123 and 100 nodes, respectively, in deep learning. The MLP is implemented by Tensorflow software, and iteration number for training procedure is 2000 to determine optimal outcomes.

TABLE I. CLASSIFICATION OF THE NUMBER OF PATIENTS OF FLU AND THE ASSOCIATED PNEUMONIA (PERSONS)

Volume	Regions	
	Taiwan	Eastern Taiwan
Low	<22889	<648
Moderate	22889-30240	648-1014
High	>30240	>1014

III. EXPERIMENTAL RESULTS

Table II demonstrates the nationwide and regional prediction accuracy of proposed MLP machine learning. The patients of total population are divided into 5 age groups : over 65 years old, 25 to 64 years old, 15 to 24 years old, 5 to 14 years old, and 0 to 4 years old. From the results, we can see that the accuracy of all of Taiwan is higher than that of Eastern Taiwan. It can be that the larger the area is, the more data can be collected, and the accuracy will be higher. Eastern Taiwan has better air quality, thus the air pollution parameters have less impact during the MLP training processing.

In the nationwide data analysis, the proposed method by MLP machine learning achieves 77.54% prediction accuracy for the trend of influenza and the associated pneumonia volumes for overall population. The nationwide trend prediction has the best result for the elderly population (81.16%). This finding is consistent with our previous study [4], which could be explained by that the elderly have weaker immunity and many comorbidities, and thus are more likely to get the flu when climatic factors and multiple air pollutants are put into considerations. The prediction accuracy for the age groups of 25 to 64, 15 to 24, and 5 to 14 years old are similar. The prediction accuracy for the age group of 0 to 4 years old is lower not only because of fewer children samples but also the children of this age group are mostly indoor so the environmental factors are less influential.

TABLE II. PREDICTION ACCURACY(%) OF ALL OF TAIWAN AND EASTERN TAIWAN FOR THE POPULATIONS WITH VARIOUS AGE GROUPS

Age Group	Regions	
	Taiwan	Eastern Taiwan
All	77.54	66.67
>=65	81.16	64.49
25~64	74.64	60.87
15~24	76.09	58.70
5~14	70.29	59.42
0~4	56.52	55.07

IV. CONCLUSION

In this paper, the meteorological parameters (temperature and relative humidity), air pollution parameters (PM 2.5 and CO) of 30 consecutive days, and the number of AURI patient of previous week are utilized to forecast the trend of the patients of flu and viral pneumonia of subsequent week by MLP machine learning. This work can provide prevention actions for diffusion of influenza.

REFERENCES

[1] W. R. Dowdle, “Influenza pandemic periodicity, virus recycling, and the art of risk assessment,” Emerg. Infect. Dis., vol.12, no.1, pp.34-39, Januray 2006.

[2] H. Sooryanarain, and S. Elankumaran, “Environmental role in influenza virus outbreaks,” Annu. Rev. Anim. Biosci., vol.3, pp.347-373, February 2015.

[3] F. Saberian, A. Zamani, M. A. Shoorehdeli, M. M. Gooya, P. Hemmati, and M. Teshnehlab, “Prediction of seasonal influenza epidemics in Tehran using artificial neural networks,” The 22nd Iran. Conf. Ele. Eng., May 2014.

[4] M. J. Chen, P. H. Yang, M. T. Hsieh, C. H. Yeh, C. H. Huang, C. M. Yang and G. M. Lin. “Machine learning to relate PM2.5 and PM10 concentrations to outpatient visits for upper respiratory tract infections in Taiwan: a nationwide analysis,” World J. Clin. Cases, vol. 6, no.8, pp. 200-206, August 2018.

[5] Environmental Protection Administration. Available from: URL: https://taqm.epa.gov.tw/taqm/en/, accessed on May 2019.

[6] Taiwan Centers for Diseases Control. Available from: URL: https://www.cdc.gov.tw/En, accessed on May 2019.

A MODIFIED STRUCTURAL SIMILARITY INDEX WITH LOW COMPUTATIONAL COMPLEXITY

Woei-Tan Loh
Faculty of Engineering
Universiti Malaysia Sarawak
Kota Samrahan, Malaysia
lohwt91@yahoo.com

David Boon Liang Bong
Faculty of Engineering
Universiti Malaysia Sarawak
Kota Samrahan, Malaysia
bbldavid@unimas.my

***Abstract*—Structural Similarity Index (SSIM) has been a benchmark method for image quality assessment (IQA). This is due to its simplicity and good performance. In this paper, we propose a modified SSIM method that reduces the computational complexity with comparable performance. Instead of computing similarities on local windows, the proposed method computes global information similarities. The proposed method omits the luminance part similarities in SSIM due to its less crucial role in assessing image quality. From the presented results, the proposed method has a much lower computational time and comparable performance compared to SSIM.**

***Keywords*— image quality, computational complexity, SSIM**

I. INTRODUCTION

In the past decade, many image quality assessment (IQA) methods are proposed. Among them, Structural Similarity Index (SSIM) [1] is the most popular methods. Many researchers extend SSIM. Results of these works prove the capability and generality of SSIM. The proposed method in this paper is termed global-SSIM (GLOSS). In GLOSS, the quality index is calculated from global information. This is different from literature works [2-3]. Fast SSIM [2] optimizes computation speed using the integral image. Work by Bruni and Vitulano [3], on the other hand, utilizes a subset of image information that is computed from local windows. GLOSS will focus on all distortions of natural images that is dissimilar to other works [4-6].

II. PROPOSED METHOD

The main factor that increases the computational complexity of SSIM is the utilization of local windows. To speed up SSIM, GLOSS computes a quality index from the global aspect. This is inspired by the work by Larson and Chandler [7]. The effectiveness of global quality in accessing image quality is also proven in other works [8]. Different from these works, global quality is not computed from local quality indices, but directly computes global quality from every pixel in GLOSS.

Thus, mean, standard deviation, and covariance are computed from the whole image. For GLOSS, similarity equation from [1] is modified to:

$$GLOSS\_1(X,Y)=\frac{(2\mu_X\mu_Y+C_1)(2\sigma_{XY}+C_2)}{(\mu_X^2+\mu_Y^2+C_1)(\sigma_X^2+\sigma_Y^2+C_2)}, \tag{1}$$

where μ and σ are the mean and standard deviation of entire image X and Y. Parameter σ_{XY} is the covariance of entire image X and Y. Constants C_1 and C_2 are used to prevent instability.

In SSIM [1], weighting function w_i aims to remove blocking artifacts. The blocking effect is due to the local information computation. In GLOSS, no local information computation is involved and w_i could be omitted. According to [9], the luminance part of SSIM is less crucial in determining image quality. Hence, the luminance similarity term is omitted accordingly. GLOSS is now simplifies to:

$$GLOSS\_2(X,Y)=\frac{2\sigma_{XY}+C_2}{\sigma_X^2+\sigma_Y^2+C_2}. \tag{2}$$

In order for GLOSS to have values between zero and one, equation (2) is modified to use the absolute value of σ_{XY}. Finally, GLOSS is formulated as:

$$GLOSS(X,Y)=\frac{2|\sigma_{XY}|+C_2}{\sigma_X^2+\sigma_Y^2+C_2}. \tag{3}$$

III. EXPERIMENTAL RESULTS

Three publicly available benchmark image databases, LIVE [10], TID2013 [11], and CSIQ [12] were utilized for evaluating GLOSS. The performance of GLOSS is evaluated from two aspects. Spearman Rank Order Correlation Coefficient (SC) measures prediction monotonicity while prediction accuracy is evaluated by Pearson Linear Correlation Coefficient (PC). Before measuring PC, nonlinear regression of the predicted scores is computed by using a logistic regression function [13]. GLOSS is compared with SSIM and Fast SSIM. The method by Bruni and Vitulano [3] is not compared as its code is unavailable.

The results of performance comparison are shown in Table 1. The best results are highlighted. For LIVE image database, SSIM performs the best. Fast SSIM and GLOSS have very similar performances for all distortion types except AWN and FF. For TID2013 image database, GLOSS outperforms SSIM especially for SC. All methods have poor results for local block-wise distortions, contrast change, and change of color saturation. This may due to the less sensitivity of SSIM-based methods for these distortions. In CSIQ image database, GLOSS attains the best results for all evaluation criteria. In overall, GLOSS has good performances. Out of the 35 distortion types in all databases, GLOSS performs the best for 24 and 15 types of distortions for SC and PC. This shows GLOSS has better monotonicity and similar accuracy compared with SSIM.

Table I. Correlation Results on Three Image Databases.

	Distortion	SC			PC		
		SSIM	Fast SSIM	GLOSS	SSIM	Fast SSIM	GLOSS
LIVE	JP2K	**0.9614**	0.9232	0.9325	**0.9665**	0.9300	0.9330
	JPEG	**0.9764**	0.9660	0.9537	**0.9789**	0.9686	0.9590
	AWN	0.9694	0.8940	**0.9865**	**0.9701**	0.9130	0.9471
	GB	**0.9517**	0.8930	0.8990	**0.9451**	0.8758	0.8915
	FF	**0.9556**	0.9370	0.9209	0.9489	**0.9504**	0.9056
TID2013	AGN	0.8671	0.7123	**0.9135**	0.8180	0.7055	**0.8741**
	NC	0.7226	0.6347	**0.8606**	0.7729	0.6468	**0.8989**
	SCN	0.8515	0.7390	**0.8993**	0.8134	0.7237	**0.8473**
	MN	0.7767	0.7663	**0.8086**	0.7772	0.7872	**0.8053**
	HFN	0.8634	0.7566	**0.9080**	0.8661	0.7726	**0.9138**
	IN	0.7503	0.6568	**0.8825**	0.6818	0.6299	**0.8239**
	QN	0.8657	0.6974	**0.9000**	**0.8128**	0.7003	0.8090
	GB	**0.9668**	0.9466	0.9569	0.9281	**0.9414**	0.8858
	ID	0.9254	0.8545	**0.9551**	**0.9552**	0.8895	0.9383
	JPEG	0.9200	0.8828	**0.9467**	**0.9512**	0.9189	0.9496
	JP2K	0.9468	0.8905	**0.9577**	**0.9648**	0.9027	0.9609
	JGTE	**0.8493**	0.8304	0.8054	**0.9003**	0.8781	0.8401
	J2TE	0.8828	0.7637	**0.9291**	**0.8558**	0.7429	0.8356
	NEPN	**0.7821**	0.7756	0.7235	0.7594	**0.7942**	0.6590
	block	**0.5720**	0.5604	0.1870	0.5619	**0.5685**	0.1519
	MS	0.7752	**0.7758**	0.5396	**0.7488**	0.7402	0.4307
	CTC	0.3775	0.3657	**0.4615**	0.4963	0.5035	**0.6759**
	CCS	**0.4141**	0.1214	0.0435	**0.4385**	0.1082	0.1196
	MGN	0.7803	0.6261	**0.8850**	0.7531	0.6377	**0.8540**
	CN	0.8566	0.4511	**0.9125**	0.8800	0.3766	**0.8907**
	Lossy	0.9057	0.8543	**0.9377**	0.9164	0.8474	**0.9279**
	Dither	0.8542	0.6786	**0.9178**	0.8411	0.6729	**0.8522**
	CA	0.8775	0.8611	**0.8920**	**0.9738**	0.9580	0.9706
	SSR	0.9461	0.9156	**0.9554**	**0.9663**	0.9247	0.9635
CSIQ	AWN	0.8974	0.8413	**0.9524**	0.8972	0.8397	**0.9501**
	JPEG	**0.9546**	0.9476	0.9412	**0.9786**	0.9682	0.9588
	JP2K	0.9606	0.9135	**0.9649**	**0.9694**	0.9189	0.9639
	PGN	0.8922	0.8186	**0.9460**	0.8906	0.8241	**0.9445**
	GB	0.9609	0.9204	**0.9637**	0.9472	0.8978	**0.9474**
	CTD	0.7922	0.8135	**0.9376**	0.7889	0.8204	**0.9299**

The computational complexities in terms of computational time are also compared. They are tested on a laptop with a core I5 processor and 8GB of RAM. Each method is tested on LIVE database with 982 images. The computational time of Fast SSIM is not computed as it runs on C++ while the other methods run on MATLAB. Its improvement compared to SSIM is quoted [2]. The results of ratio computational time of GLOSS and Fast SSIM to SSIM are shown in Figure I. GLOSS and Fast SSIM is about ten times and three times faster than SSIM. Thus, GLOSS is about three time faster than Fast SSIM.

Figure I. Ratio Computation Time to SSIM.

IV. CONCLUSION

This paper proposed a modified SSIM for lower computational complexity, GLOSS. GLOSS can run about ten times faster than SSIM. Despite its low computational time, GLOSS has a slight performance drop in assessing qualities of images. Compare to Fast SSIM, a recent method in speeding SSIM, GLOSS runs faster and has comparable performance.

V. ACKNOWLEDGMENT

This work was supported by Ministry of Higher Education Malaysia through the provision of research grant F02/FRGS/1492/2016.

REFERENCES

[1] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE Transactions on Image Processing, 13(4), pp. 600-612, 2004.

[2] M.J. Chen, and A.C. Bovik, “Fast Structural Similarity Index Algorithm,” Journal of Real-Time Image Processing, 6(4), pp. 281-287, 2011.

[3] V. Bruni, and D. Vitulano, “An Entropy Based Approach for SSIM Speed Up,” Signal Processing, 135, pp. 198-209, 2017

[4] D.B.L. Bong and B.E. Khoo, “Objective blur assessment based on contraction errors of local contrast maps,” Multimedia Tools and Applications, 74(17), pp.7355-7378, 2015.

[5] D.B.L. Bong and B.E. Khoo, “Blind image blur assessment by using valid reblur range and histogram shape difference,” Signal Processing: Image Communications, 29(6), pp. 699-710, 2014.

[6] W.-T. Loh and D. B. L. Bong, "Quality Assessment for Natural and Screen Visual Contents," in International Conference on Image Processing (ICIP), IEEE, Taipei, Taiwan, pp. 3025-3026, September 2019.

[7] E.C. Larson, D.M. Chandler, “Most apparent distortion: full-reference image quality assessment and the role of strategy,” Journal of Electronic Imaging, 19(1), p. 011006, 2010.

[8] P.V. Vu, D.M. Chandler, “A fast wavelet-based algorithm for global and local image sharpness estimation,” IEEE Signal Processing Letters, 19(7), pp. 423-426, 2012.

[9] D.M. Rouse, and S.S. Hemami, “Understanding and simplifying the structural similarity metric,” in the 15th IEEE International Conference on Image Processing (ICIP), IEEE, California, United States, pp. 1188-1191, October 2008.

[10] H.R. Sheikh, H. R. “LIVE Image Quality Assessment Database Release 2. http://live. ece. utexas. edu/research/quality, 2005.

[11] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo “Image Database TID2013: Peculiarities, Results and Perspectives,” Signal Processing: Image Communication, 30, pp. 57-77, 2015.

[12] E.C. Larson and D.M. Chandler, "Most Apparent Distortion: Full-Reference Image Quality Assessment and the Role of Strategy," Journal of Electronic Imaging, 19 (1), p. 011006, 2010.

[13] Video Quality Expert Groups, “Final Report from the Video Quality Experts Group on the Validation of Objective Quality Metrics for Video Quality Assessment,” http://www.its.bldrdoc.gov/vqeg/projects/frtv-phase-i/frtv-phase-i.aspx, 2000.

Quality Enhancement of DDBTC Decoded Image

Joni Welman Simatupang
Study Program of Electrical Engineering
President University
Bekasi, Indonesia
joniwsmtp@president.ac.id

Heri Prasetyo
Department of Informatics
Universitas Sebelas Maret
Surakarta, Indonesia
heri.prasetyo@staff.uns.ac.id

***Abstract*— This paper presents two methods for improving the quality of Dot Diffused Block Truncation Coding (DDBTC) decoded image. The first method exploits the usability and effectiveness of decimated Discrete Wavelet Transform and stationary Wavelet Transform for reducing the half-toning artifact. While, the second method employs the Vector Quantization (VQ) approach for performing the image patch replacement and suppressing the unpleasant artifact of DDBTC decoded image. As documented in Experimental Section, these two methods performs well for improving the quality of DDBTC decoded image.**

Keywords*— *dot diffused, image enhancement, vector quantization, wavelet

I. Wavelet-Based Quality Improvement

Let I be an input image of size $M \times N$. Typically, the input image is in grayscale. However, the DDBTC method can be easily extended to the color image. DDBTC scheme firstly divides I into several non-overlaping blocks of size $m \times m$. Suppose that i is the processed image block. The DDBTC converts and encodes i into another form as follow [1]:

$$i \Rightarrow \{t_{min}, t_{max}, b\}, \tag{1}$$

where t_{min} and t_{max} denote the minimum and maximum quantizer, respectively. b is the bitmap image of size $m \times m$. In 8-bits pixel representation, DDBTC reduces the required bits for storing a single image block from $8m^2$ bits to be $2 * 8 + m^2$ bits. The DDBTC employs the class matrix and error diffused kernel for generating b [1-3].

In the decoding process, the DDBTC reconstructs t_{min}, t_{max}, and b into the original image block. This process is denoted as:

$$\hat{\imath} \Leftarrow \{t_{min}, t_{max}, b\}, \tag{2}$$

where $\hat{\imath}$ denotes the reconstructed image block of size $m \times m$. Each pixel in $\hat{\imath}$ is composed from t_{min} or t_{max} based on the value of b. This substitution process is given as follow:

$$\hat{\imath}(x,y) = \begin{cases} t_{min}, & \text{if } b(x,y) = 0 \\ t_{max}, & \text{if } b(x,y) = 1 \end{cases}, \tag{3}$$

for $x, y = 1,2,\dots,m$. One obtains the DDBTC decoded image $\hat{I}$ by collecting all image blocks $\hat{\imath}$.

The wavelet-based image enhancement aims to improve the quality of DDBTC decoded image by exploiting the decimated and stationary wavelet transform. Specifically, this process is denoted as follow:

$$\tilde{I} \Leftarrow \psi\{\hat{I}\}, \tag{4}$$

where $\tilde{I}$ and $\psi\{\cdot\}$ are the enhanced DDBTC decoded image and operator of wavelet-based image enhancement, respectively. The quality of enhanced image should be as similar as possible to that of the original image, i.e. $\tilde{I} \approx I$.

In the wavelet-based enhancement, the DDBTC decoded image $\hat{I}$ is firstly downsampled by factor 0.5 with bicubic interpolator. It indicates that the image size of $\hat{I}$ reduces from $M \times N$ into $\frac{M}{2} \times \frac{N}{2}$. Subsequently, we perform the wavelet transforms on $\hat{I}$ as:

$$\mathcal{W}\{\hat{I}\} \Rightarrow \{\hat{I}_{\phi} | \phi = (LL, LH, HL, HH)\}, \tag{5}$$

$$\mathcal{S}\{\hat{I}\} \Rightarrow \{\hat{I}_{\phi_s} | \phi_s = (LL, LH, HL, HH)\}, \tag{6}$$

where $\mathcal{W}\{\cdot\}$ and $\mathcal{S}\{\cdot\}$ denote the operator of decimated wavelet and stationary wavelet transform, respectively. In this paper, we utilize the Daubechies 1 (db1) for $\mathcal{W}\{\cdot\}$ and $\mathcal{S}\{\cdot\}$. The symbols $\hat{I}_{\phi}$ and $\hat{I}_{\phi_s}$ are the transformed DWT and SWT sub-bands, respectively. Both are of sizes $\frac{M}{4} \times \frac{N}{4}$ and $\frac{M}{2} \times \frac{N}{2}$, respectively. Each image sub-band $\hat{I}_{\phi}$ is further upsampling with factor 2 using bicubic interpolator to obtain the image sub-band of size $\frac{M}{2} \times \frac{N}{2}$.

Simple additive operation is then applied for each image sub-band $\hat{I}_{\phi}$ and $\hat{I}_{\phi_s}$ as bellow:

$$A_{\phi} \Leftarrow \lambda_{\phi} \hat{I}_{\phi} + (1 - \lambda_{\phi}) \hat{I}_{\phi_s}, \tag{7}$$

where A_{ϕ} is fusion between $\hat{I}_{\phi}$ and $\hat{I}_{\phi_s}$, for all $\phi = (LL, LH, HL, HH)$. The symbol λ_{ϕ} is the additive scaling factor controlling the strength of $\hat{I}_{\phi}$ and $\hat{I}_{\phi_s}$. The size of A_{ϕ} is $\frac{M}{2} \times \frac{N}{2}$. At the end, an inverse DWT operation is executed as follow:

$$\tilde{I} \Leftarrow \mathcal{W}^{-1}\{A_{\phi} | \phi = (LL, LH, HL, HH)\}, \tag{8}$$

where $\mathcal{W}^{-1}\{\cdot\}$ denotes the inverse DWT operator, and $\tilde{I}$ is the DDBTC reconstructed image of size $M \times N$. The wavelet-based approach effectively suppresses the occurred impulsive noise on DDBTC decoded image.

II. VQ-Based Quality Improvement

This section presents the quality enhancement of DDBTC decoded image using VQ approach. This method requires a set of training images denoted as $T = \{t_k | k = 1,2,\dots,N_T\}$, where N_T is the number of training images, i.e. original image without compression. Let $C = \{C_1, C_2, \dots, C_{N_c}\}$ be a codebook of size N_c obtained from VQ training over T. The symbol C_k denotes the k-th codeword of size $n \times n$, for $k = 1,2,\dots,N_c$. We firstly extract image patch p of $n \times n$ from the DDBTC decoded image $\hat{I}$. Herein, we use overlapping strategy on image patch extraction process. Subsequently, we perform similarity matching between p with C_k with Euclidean distance as follow:

$$k^* \Leftarrow \arg \min_{k=1,2,\dots,N_c} \|p - C_k\|_2^2, \tag{9}$$

where k^* is the index of the most similar codeword C_k in C. Then, the image patch p is simply replaced with codeword C_{k^*}, i.e. the most closest codeword, as follow:

$$\tilde{p} \Leftarrow C_{k^*}, \tag{10}$$

where $\tilde{p}$ is the replaced image patch.

The procedure of image patch replacement is performed over all overlapping image patches. By collecting all image patches $\tilde{p}$, we obtain an non-alignment DDBTC reconstructed image as follow:

$$\tilde{o}(x, y) \Leftarrow \cup_{\forall \tilde{p}} \tilde{p}. \tag{11}$$

An alignment process is then executed to obtain a correct DDBTC enhanced image. This process is defined as bellow:

$$\tilde{I}(x, y) \Leftarrow \frac{\sum \tilde{o}(x,y)}{\sum R^T(x,y)R(x,y)}, \tag{12}$$

where $R(x, y)$ is the image patch operator. At the end of VQ-based processing, we obtain the enhanced DDBTC decoded image $\tilde{I}$ of size $M \times N$.

III. Experimental Results

We investigate the performances over four color images as displayed in Fig. 1, each of size 512×512. Herein, the wavelet and VQ-based methods perform image enhancement over all color spaces, i.e. Red, Green, and Blue individually. The wavelet-based approach utilizes λ_ϕ as $\{0.8, 0.7, 0.7, 0.6\}$ for wavelet sub-bands $\{LL, LH, HL, HH\}$. Whereas, the VQ codebook is generated from standard images (Lena, Baboon, Lake, and Peppers) as training set.

A. *Visual Investigation*

The visual investigation between the wavelet and VQ-based methods is reported in this sub-section. We use the codebook of size $N_c = 256$ for VQ-based approach. The DDBTC image block size is set as $m \times m = 8 \times 8$. Fig. 2 demonstrates the enhancement results of DDBTC decoded image. The first row is the DDBTC decoded image. Whereas the second and third rows are the enhanced images from wavelet and VQ-based method, respectively. As shown in Fig. 2, these two aforementioned schemes improve the quality of DDBTC decoded image. The result from VQ-based scheme is visually better compared to that of the wavelet-based approach.

B. *Objective Evaluation*

This sub-section summarizes the performance of wavelet and VQ-based approaches in terms of objective image quality assessments. We consider two objective metrics, i.e. Peak-Signal-to-Noise-Ratio (PSNR) and Structural Similarity Index Metric (SSIM), to examine the performance. All images in Fig. 1 are turned as testing image. The average PSNR and SSIM scores are subsequently computed over all testing images. In this experiment, we employ several codebooks over different sizes, i.e. $N_c = 16, 32, \ldots, 256$. Table I tabulates the performance comparisons between the wavelet and VQ-based approaches on quality enhancement of DDBTC decoded images. These two aforementioned methods effectively improves the image quality as indicated with increasing average PSNR and SSIM values in comparison with the DDBTC decoded image. The VQ-based approach offers better performance compared to that of the wavelet-based method. Yet, the proposed quality enhancement methods yield acceptable results based on visual investigation as well as objective measurement.

References

[1] Y. F. Liu and J. M. Guo, "Dot-diffused halftoning with improved homogeneity," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4581-4591, Nov. 2015.

[2] J. M. Guo, H. Prasetyo and N. J. Wang "Effective image retrieval system using dot-diffused block truncation coding features," *IEEE Trans. Multimedia*, vol. 17, no. 9, pp. 1576-1590, Sep. 2015.

[3] H. Prasetyo and H. Kurniawan, "Reducing JPEG False Contour Using Visual Illumination," *Informations*, vol. 9, no. 2, pp. 41, 2018.

Fig. 1. A set of testing images in color space.

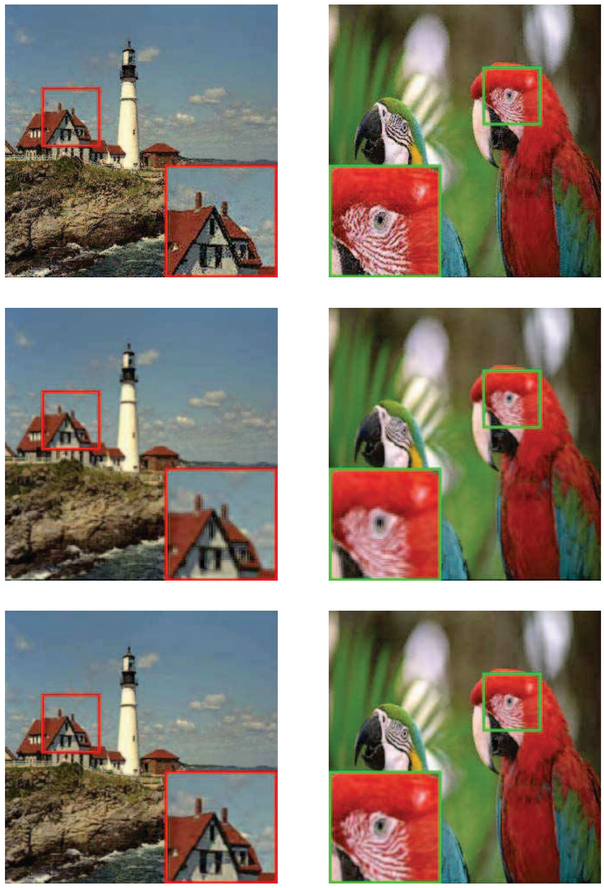


Fig. 2. The quality enhancement of DDBTC decoded images: (first row) DDBTC decoded image, (second row) enhanced image with wavelet-based approach, and (third row) enhanced image from VQ-based method.

TABLE I. Evaluations in Terms of Average PSNR and SSIM Scores

Method	Average PSNR	Average SSIM
DDBTC Decoded Image	20.52	0.74
Wavelet-Based Method	24.97	0.84
VQ-Based Method, $N_c = 16$	26.05	0.86
VQ-Based Method, $N_c = 32$	27.80	0.91
VQ-Based Methd, $N_c = 64$	28.44	0.92
VQ-Based Method, $N_c = 128$	28.89	0.93
VQ-Based Method, $N_c = 256$	29.25	0.94

A 10-bit 100-MHz Current-Steering DAC with Randomized Thermometer Code Calibration Scheme

Hsin-Liang Chen[1], Ming-Han Hsieh[2], and Jen-Shiun Chiang[2*]

[1]Dept. of Electrical Engineering, Chinese Culture Univ.
55, Hwa-Kang Rd., Yang-Ming-Shan, Taipei, Taiwan
E-mail: cxl7@ulive.pccu.edu.tw

[2]Dept. of Electrical and Computer Engr., Tamkang Univ.
Tamsui District, New Taipei City, Taiwan
E-mail: chiang@tku.edu.tw

Abstract

This paper presents a 10-bit 100 MHz digital-to-analog converter (DAC) by using a current-steering architecture. The proposed DAC suits for high speed and high resolution applications. Since the current-steering DAC suffers from the process mismatch which limits both the static and dynamic performances, this work employs a pseudo-random controller to improve the linearity. The random generator controls the selection of the element in the MSB part, and therefore, the harmonics caused by mismatch problem can be mitigated.

Key words: digital to analog converter, current-steering, dynamic element matching, pseudo-random number generator.

I. Introduction

Current-steering digital-to-analog converters (DAC) are widely used in wideband applications [1], where a high spurious-free dynamic rang (SFDR) is particularly preferred. Fig. 1 shows a conventional thermometer coded current-steering DAC. It consists of equally weighted current cells. Each current cell contains a current source, a MOSFET pair as a current switch, and a digital latch controlled by a clocked signal. The latch performs the data synchronization and adjusts the crossing point of the differential output of control signals. The complementary outputs of the latch control the current switch for directing the current source to OUT+ or OUT−.

One of the major obstacles of high-resolution current-steering DAC is element mismatch. Techniques of dynamic element matching (DEM) [2] have been successfully used to reduce the element mismatch effects. As a DEM method, data-weighted average (DWA) is widely used for oversampling rate applications [3]. In Nyquist rate applications, a random multiplying method is proposed for current-steering DAC. With a proper layout switching scheme, the effect of the serious mismatch problem caused by small transistors can be reduced, also a small area DAC can be implemented.

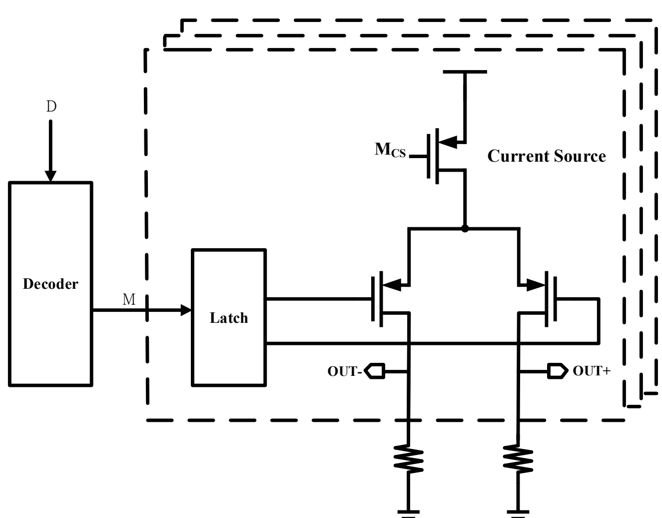


Fig. 1 Current-steering DAC.

The DAC static linearity problems, specified as differential nonlinearity (DNL) and integral nonlinearity (INL), are mainly determined by the mismatching of current source among different current cells. The cascode current source is employed to increase the output resistance. The size of the transistors in the current source must be large enough to make sure perfect current matching.

II. Proposed DAC Architecture

Fig. 2 shows the proposed 10-bit DAC architecture. The DAC is segmented into a 4-bit LSBs binary-weighted DAC and a 6-bit MSBs thermometer code DAC. The 6 MSBs are converted through two thermometer decoders. B4~B6 are converted into a 7-bit thermometer code and will be decoded by a randomizer. Then the input binary codes of the 6 MSBs are converted into 63-bit thermometer codes. Because the 6-to-63 decoder circuits are complicated, it divides it into Row-Column decoders which are two 3-to-7 thermometer decoders and the 4 LSBs are converted through a binary code buffer. In order to avoid causing a different decoder which is not synchronized on time, the buffer equalizes delay time paths between the 4 LSBs input signals and latch. Then the 63-bit thermometer code and 4 LSBs binary code will determine the control signals to control the switch of the current unit.

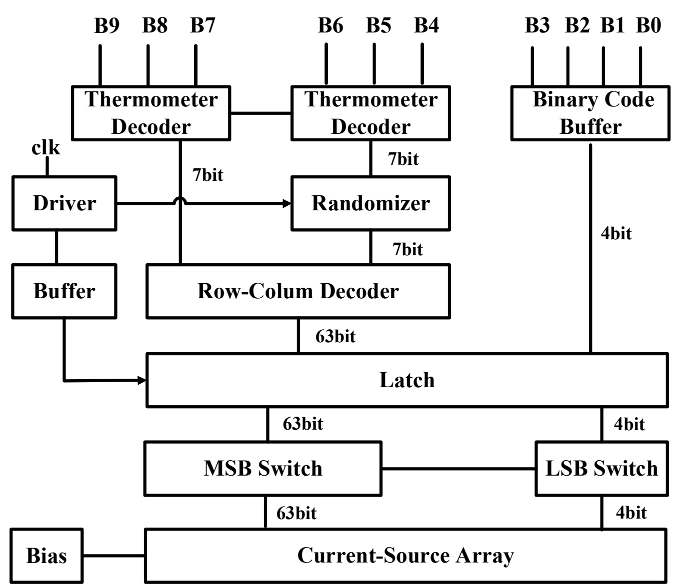


Fig. 2 Segmented 10-bit current-steering DAC.

III. Implementation of The Latch and Randomizer

A. Latch

The current-steering DAC will have dynamic performance degradation caused by three factors: (1) synchronization mismatch of control signals at the switch, (2) drain-voltage variation of the current-source transistors, (3) coupling of the control signals through C_{gd} capacitances of switch to the output. To synchronize the control signals, a high speed latch is employed and shown in Fig. 3.

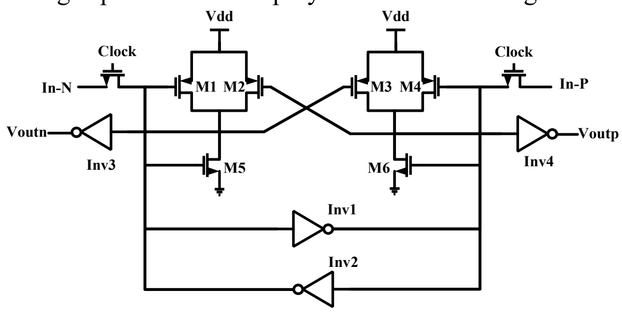


Fig. 3 The high-speed latch circuit.

The major function of this driver is to shift the crossing point of the control signals at the current source switch. Thus, the switch transistors will never be simultaneously in the off state. By sizing transistors M1~M6, the crossing point of the control signal at the switch transistors can be adjusted. A lower crossing point signal should be implemented because the PMOS current sources and PMOS switches are used. The feedback inverters, Inv1 and Inv2, can suppress the clock feedthrough from the pass transistors.

B. Randomizer

Fig. 4 shows the detailed randomizer circuit schematic [1], which is composed of MUX array and pseudo-random number generator (PRNG) to obtain a repeat-cycle of 2^n-1 with DFF. The PRNG is modified from a common linear feedback shift register (LFSR). According to the different application word lengths, this PRNG could be selected with the *sel* controlled bit.

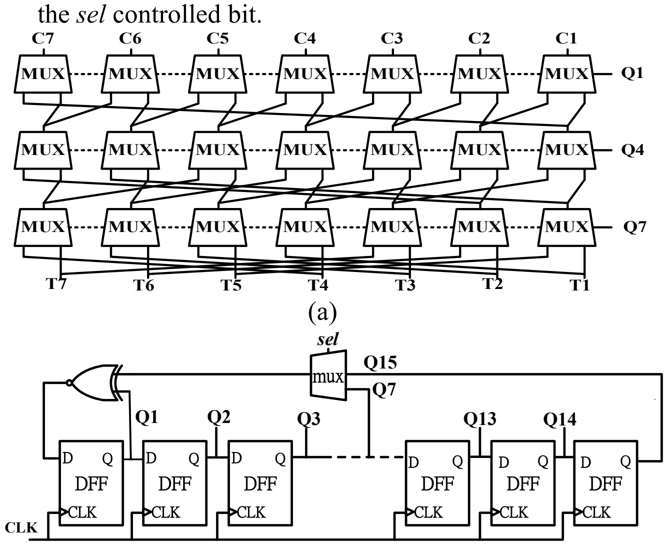


Fig. 4 (a) MUX array, (b) the PRNG circuit

IV. Simulation Results

The experimental circuit was designed and simulated by 0.18 um process. The simulation setup is based on a single-ended 50Ω terminated resistance with 1.8V of the supply voltage. The expected mismatched current source has 10% of maximum standard deviation. Fig. 5 shows that the simulated output spectrum of the DAC with 128 samples. It shows that the randomized calibration effectively suppresses the distortion tones by current mismatching. With 7-bit/15-bit randomized calibration, the SFDR is up to 70.6/71.6 dB, respectively. Fig. 6 shows the simulated output spectrum of the DAC with 256 samples. Since the 7-bit PRNG has 128 of repeat-cycle, in this case, 256 samples would be expected with worse linearity that caused by repeating random numbers and it would lead to reduce the SFDR. The result shows that the repeating pseudo-random number would increase only 0.5dB of distortion. Therefore, if the area cost is limited, the PRNG with lower bit could be traded off with the performance of SFDR.

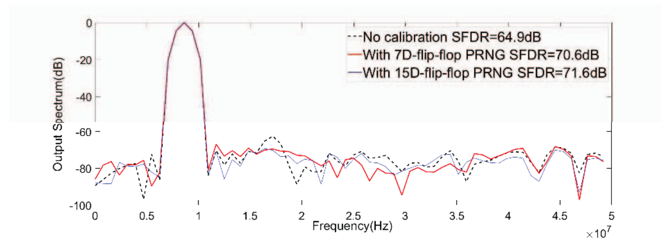


Fig. 5 DAC output spectrum corresponding to 128 samples

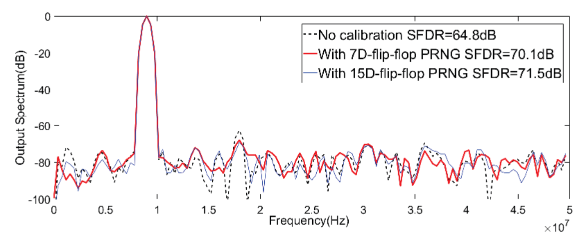


Fig. 6 DAC output spectrum corresponding to 256 samples

V. Conclusion

In this paper, a current steering DAC with randomized thermometer code is proposed. The 10-bit current-steering DAC is implemented in a 0.18 um CMOS process. The simulation results indicate that the randomized thermometer code calibration substantially reduces the effects of the large mismatch caused by small transistors. Also, the bit number of the PRNG can trade off with different data length of application. The DAC with the calibration scheme can be implemented with small-area while the good linearity is obtained.

Acknowledgment

This work is supported by the Ministry of Science and Technology, Taiwan, under Grand Number MOST 107-2218-E-034-001-MY2.

Reference

[1] W.-T. Lin, H.-Y. Huang, and T.-H. Kuo, "A 12-bit 40 nm DAC achieving SFDR > 70 dB at 1.6 GS/s and IMD <–61 dB at 2.8 GS/s with DEMDRZ technique," *IEEE J. Solid-State Circuits*, vol. 49, no. 3, pp. 708–717, Mar. 2014.

[2] K. L. Chan, J. Zhu, and I. Galton, "Dynamic element matching to prevent nonlinear distortion from pulse-shape mismatches in high-resolution DACs," *IEEE J. Solid-State Circuits*, vol. 43, no. 9, pp. 2067–2078, Sep. 2008.

[3] D. H. Lee and T. H. Kuo, "Advancing data weighted averaging technique for multi-bit sigma–delta modulators," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, no. 10, pp. 838–842, Nov. 2007.

[4] D. A. Johns and K. Martin, Analog Integrated Circuit Design, Wiley, 1996.

2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)

Reconfirm gestalt principles from scan-path analysis on viewing photos

Hsien Chih Chuang
Department of information communications
Chinese Culture University
Taipei, Taiwan
zxz7@ulive.pccu.edu.tw

Da Lun Tang
Department of mass communications
Tamkang University
Taipei, Taiwan
136216@mail.tku.edu.tw

***Abstract*— The advances in communication technology not only ushered in an era of visual image explosion, but also led to the emergence of the "visual generation" who are heavily image-reliant in their quotidian existence. Photography is a form of visual arts that sets great store by human visual perception. It brings aesthetic delight by connecting human vision with psychological experiences. Every captivating or heart-touching photograph has its sui generis context or story, and gestalt theory may be the key to this kind of mystery. However, people have simply a limited knowledge of their viewing mechanism, and compelling evidence for the links between viewers' inner feelings and visual cognition remains wanting. To remedy this deficiency, this study takes the angle of positive aesthetics and employs the eye-tracking method to investigate the influence of photographic compositions' gestalt on the subjects' visual cognition, so as to grasp how they look at, interpret and appreciate photographic works. By manipulating the gestalt elements, calculating the dispersion degree of the subjects' scan paths, and collecting the subjects' opinions on the aesthetic quality of each photographic work, we found that the works embodying the gestalt principles tend to confine the subjects' scan paths to a unified context, while those violating the principles do not. Enhancing the foundation of photography and developing a new eye-tracking method with positive aesthetics, this study will stimulate photographers' artistic creativity, promote arts education, improve students' aesthetic literacy, and produce useful teaching material on the pedagogy and theories of photography.**

Keywords*— *gestalt principles, eye-tracking, photography

I. INTRODUCTION

As a form of visual art that attaches great importance to human visual perception, photography provides viewers with aesthetic delight by connecting visual perceptions with psychological experiences. Photography has been closely tied to artists and artistic creation since the early development of gestalt theory. Examples of applying gestalt theory to photography courses can be found at home and abroad (e.g. Zakia, 2002). The theory-based teaching focuses not so much on blindly applying the principles of gestalt perception as on invoking them creatively and flexibly. Gestalt psychology and semiology thus significantly improve photographers' ability in visual expression. As a matter of fact, every captivating or heart-touching photograph has its sui generis context or story. Gestalt psychology and semiology may be the key to this kind of mystery apart from being helpful for beginners to overcome the problem of unrigorous composition. As a crucial psychological approach, the eye-tracking analysis has been employed extensively in applied psychology. Conducting a psychological eye-tracking experiment, this study seeks to investigate the variation in the frequency, position and distribution of the subjects' gaze on gestalt photographs, insofar as to grasp people's psychological processes and patterns of admiring photographic works.

II. RESEARCH METHOD

Adopting the eye-tracking method, this study not only analyzes the gestalt properties and the psychological process of vision, but also explores the connections between photographic techniques and viewing modes as well as those between gestalt composition and people's visual perception. By virtue of the eye-tracking experiment, this study manipulates the photographs' gestalt elements, so as to identify the key factors behind viewers' visual perception and subjective aesthetic judgement about photographic works.

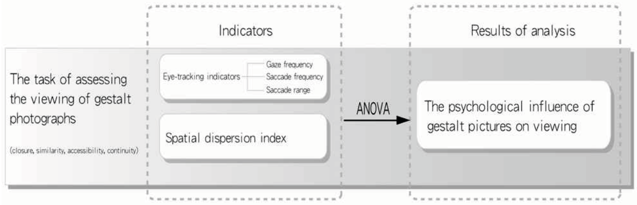


Fig. 1. The Research Design

This study recruits 30 university students for the experiment which features a random display of stimuli and a "within subject design," that is, each subject needs to partake in every step of the experiment. The pictures to be displayed are derived from renowned photographers' oeuvres and carefully selected by experts and scholars of photography, so as to make sure that they conform the gestalt principles. These pictures are divided into four groups according to four gestalt principles (i.e. closure, similarity, accessibility, and continuity), and each group consists of 10 photographs, hence a total of 40 pictures. Each of these pictures has a pixel resolution of 1024*768, and will be displayed for 10 seconds one by one. The scan paths of each subject will be recorded when the subject is viewing these pictures. This experiment takes 10 minutes.

The experiment manipulation is described as follows:

Independent variables: gestalt principles (closure, similarity, accessibility, and continuity)

Dependent variables: eye-tracking indicators (total gaze frequency, total saccade range, average saccade range, and saccade frequency)

Fig. 2. An example of the experiment stimuli (closure)

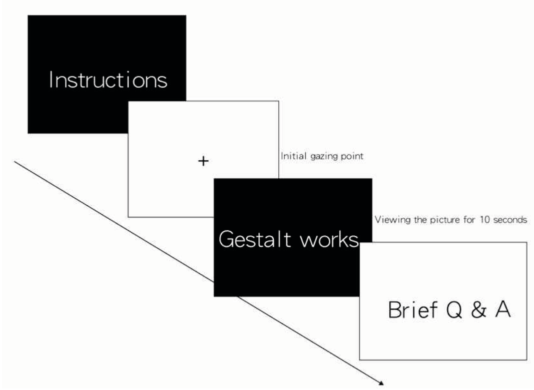


Fig. 3. A flowchart of the experimental steps.

III. CONCLUSION

This study provides empirical evidence from the perspective of experimental aesthetics. To draw a tentative conclusion, we argue that understanding human vision and gestalt principles, mastering photographic composition, grasping common viewing patterns and processes, and knowing how to attract viewers' attention are necessary commodities for photographers and photography learners if they are going to give viewers a common focus, enhance their visual perception and delight, and ergo create brilliant works in a more objective, accurate fashion that bring viewers better admiring experience and greater aesthetic pleasure.

REFERENCES

[1] Antes, J. R. (1974). The time course of picture viewing. Journal of Experimental Psychology, 103(1), 62-70.

[2] Arnheim, R. (1969). Visual thinking. Berkeley: University of California Press.

[3] Arnheim, R. (1974). Art and visual perception: The new version. Berkeley: University of California Press.

[4] Babcock, J. S., Pelz1, J. B., & Fairchild, M. D. (2003). Eye tracking observers during rank order, paired comparison, and graphical rating tasks. Proceedings of the 2003 PICS Digital Photography Conference, Rochester, NY.

[5] Berlyne, D. (1971). Aesthetics and psychobiology. New York: Appleton Century Crofts.

[6] Checkosky, S. F., Whitlock, D. (1973). The effects of pattern goodness on recognition time in a memory search task. Journal of Experimental Psychology, 100(2), 341-348.

[7] Duchowski, A. T. (2003). Eye tracking methodology: theory and practice. Verlag London Limited, 186-187.

[8] Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. Behavioral and Brain Sciences , 22, 661-674.

[9] Gandhi, N.J., & Keller, E.L. (1997). Spatial distribution and discharge characteristics of superior colliculus neurons antidromically activated from the omnipause region in the monkey. Journal of Neurophysiology, 78, 2221-2225.

[10] Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. International Journal of Industrial Ergonomics, 24(6), 631-645.

[11] Gombrich, E. H. (1982). The image and the eye. London: Phaidon Press.

[12] Gombrich, E. H. (1984). A sense of order. London: Phaidon.

[13] Gombrich, E. H. (1995). The story of art. London: Phaidon.

[14] Henderson, J. M. (2007). Regarding scenes. Current Directions in Psychological Science, 16, 219-222.

[15] Henderson, J. M., & Hollingworth, A. (1998). Eye movement during scene viewing: An overview. In Eye Guidance in Reading and Scene Perception, G. Underwood, Ed. Elsevier Science Ltd.

[16] Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. Journal of Experimental Psychology: Human Perception & Performance, 25(1), 210-228.

[17] Hong, J., & Lee, A. Y. (2010). Feeling mixed but not torn: the moderating role of construal level in mixed emotions appeals. Journal of Consumer Research, 37(3), 456-472.

[18] Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. Cognitive Psychology, 8(4), 441-480.

[19] Kingstone, A., Smilek, D., &Eastwood, J.D. (2008). Cognitive ethology: A new approach for studying human cognition. British Journal of Psychology, 99, 317-345.

[20] Koffka, K. (1935). Principles of Gestalt psychology. New York: Harcourt, Brace, & World.

[21] Liversedge, S. P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. Trends in Cognitive Sciences, 4(1), 1-14.

[22] Loftus, G. R. & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. Journal of Experimental Psychology: Human Perception and Performance, 4(4), 565-572.

[23] MackWorth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. Perception & Psychophysics, 2(11), 547-552.

[24] Maritain, J., M. Rader (Rd.) (1966). A modern book of Esthetics. New York: Holt, Rinehart & Winston.

[25] McManus, I., Edmondson, D., Rodger, J. (1985). Balance in pictures. British Journal of Psychology, 76(3), 311-324.

[26] McManus, I., Kitson, C. (1995). Compositional geometry in pictures. Empirical Studies of the Arts, 13(1), 73-94.

[27] Megaw, E.D. & Richardson, J. (1979). Target uncertainty and visual scanning strategies. Human Factors, 21(4), 302-315.

[28] Min-Yuan Ma, Hsien-Chih Chuang. (2015). A Legibility Study of Chinese Character Complicacy and Eye Movement Information. Perceptual and Motor Skills, 120, 232-246.

[29] Min-Yuan Ma, Hsien-Chih Chuang. (2015). How form and structure of Chinese characters affect eye movement control. Journal of Eye Movement Research, 8(3): 3, 1-12.

[30] Min-Yuan Ma, Hsien-Chih Chuang. (2017). An Exploratory Study of the effect of enclosed structure on type design with fixation dispersion: Evidence from eye movements. International Journal of Technology and Design Education, 27(1), 149-164.

[31] Newbury, D. (1995). Towards a new practice in photographic education (Unpublished doctoral dissertation). University of Central England, Birmingham.

[32] Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372-422.

[33] Reber, R., Winkielman, P., Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. Psychological Science, 9(1), 45-48.

[34] Rogers, A., & Allen, D. (1996). What's happening to photography. Journal of Art and Design Education, 15(1), 23-30.

[35] Salvucci, D. D., & Anderson, J. R. (1998). Tracing eye movement protocols with cognitive process models. In Proceedings of the Twentieth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

[36] Solso, R. L. (1997). Cognition and the visual arts. Cambridge, MA: MIT Press.

[37] Wilson, A., Chatterjee, A. (2005). The assessment of preference for balance: Introducing a new test. Empirical Studies of the Arts, 23(2), 165-180.

[38] Yarbus (1967). A.L. Eye movement and vision. P1enum Press.

[39] Zakia, R. D. (1993). Photography and visual perception. Journal of Aesthetic Education, 27(4), 67-81.

[40] Zakia, R. D. (2002). Perception and imaging (2nd ed.). Woburn, MA: Focal.

Controlling myopia progression in children by the rotary prism eye exercise device

Der-Chin Chen
Department of Electrical Engineering
Feng Chia University
Taichung, Taiwan
kanatasan.tw@yahoo.com.tw

Jih-Yi Liao
Chief Executive Officer
Loocadi ,INC. No.29199990
Taipei, Taiwan
id4chair1616@hotmail.com

Wei-hsin Chen
Department of Optometry
Hsin Sheng College of Medical Care and Management
Taoyuan, Taiwan
Wayne08560856@gmail.com

Po-Ting Liu
Professional Master's Program of Optometry Science and Technolgy
Feng Chia University
Taichung, Taiwan
boking8291@gmail.com

Shih-Tsung Chang
Optometry Science and Technolgy
Feng Chia University
Taichung, Taiwan
cheung2379@yahoo.com

***Abstract*—In this paper, the use of the rotary prism eye exercise device combined with dynamic modified eye charts to establish the children myopia control method by rotating prism speed with changing. In this program, clients typically range in age from 12 to 18 years of age. The eye exercise was performed on Saturday and Sunday of each week for three months. The visual acuity was measured by the use of an auto refraction device before and after two weeks of eye exercises. After three months, the improvement in visual acuity was average of 0.20D of both right and left eyes decreasing an indicating that eye exercises as vision training-based program could improve visual acuity for children with myopia.**

Keywords—myopia control, eye exercises, rotating prism

I. INTRODUCTION

Myopia commonly onsets in childhood and is due to a mismatch between the eyeball length and its optical power, resulting in light focusing in front of the retina and thus causing blurred distance vision. Currently, the few strategies used for myopia control have proven to be effective. These effective methods are the use of multifocal spectacles, orthokeratology contact lenses, soft bifocal contact lenses, and topical pharmaceutical agents such as atropine. [1,2] Atropine eye drops have been used for myopia control for many years, with effective short-term results. But use of these eye drops also has some drawbacks. The best way to take advantage of methods to control myopia is to detect nearsightedness early. The eye exercises can control myopia progression and cannot correct your eyesight. In this study, the use of the vision training device combined with dynamic modified eye charts to establish the children myopia control method by rotating prism speed with changing.

Prisms can be used to change the orientation of the light beam, as well as combining or splitting optical beams with partial reflecting surfaces. Ray angle deviation through a prism can be determined by tracing a light ray through the element and using Snell's law at each interface. For the prism shown at Fig.1, the indicated angles are given by equation (1).

$$\theta_2 = \theta_1' - \alpha \qquad (1)$$

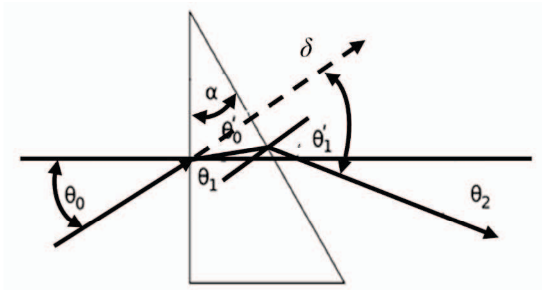


Fig. 1. ray deviation by prism

The angle of incidence θ_0 and prism apex angle α are both small for thin prism. The θ_1' is the angle of refraction.[3,4]

The deviation angle of thin lens depends on wavelength through n; the deviation angle varies with wavelength according to eq. (2):

$$\delta(\lambda) = [\, n(\lambda) - 1] * \alpha \qquad (2)$$

Where indices index is $n(\lambda)$, as the function of light wavelength, the n_d is the refractive index at the wavelength 587.56 nm.

II. SYSTEM STRUCTURE

Eye exercises can increase eye function and help you focus better, reduce eyestrain and sensitivity to light and help with other aspects of vision such as hand-eye coordination, depth perception, peripheral vision, etc. The eye exercises system in Fig.2 is compose of as followings: (a) Patient with nearsightedness, (b) dynamic modified eye charts (ppt), (c) Auto refractor, and (e) the rotary prism eye exercise device. The patient wears the eye exercise device, and looks at the dynamic modified eye chart placed at 6 meters position, then at near for individuals needing reading glasses.

A eye exercise device for exercising the muscles of the eye wherein the exercise consists of circulating movement of eye and prisms rotation speeds change, the set of device being attached to a patient's head and positioned in front of the patient 's eyes.

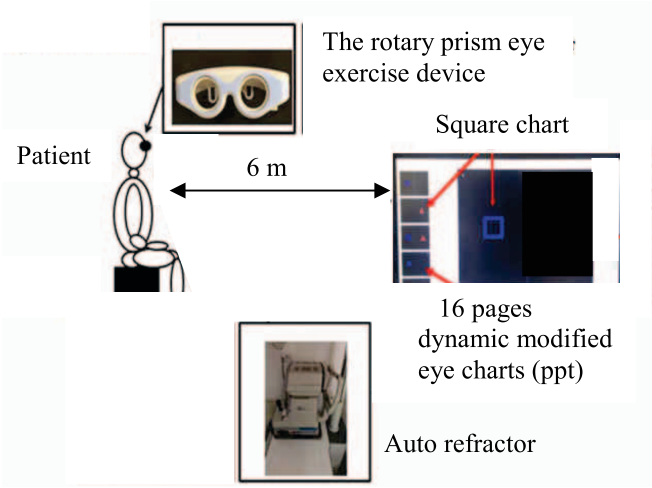


Fig. 2. the eye exercise system

It is comprised of : (a) a front frame part; (b) a back frame part; (c) thin prism of the left and right. ; (d) two DC motors ; (e) two gear sets providing the rotation of thin prism. The figure 3 shows the rotary prism's eye exercise device. The rotation speed of the thin prism should then be adjusted

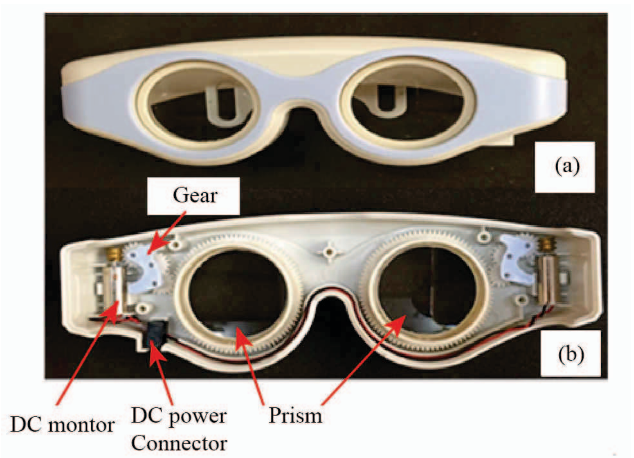


Fig. 3. eye exercise device

III. Experiments And Results

Prior to entering into the eye exercise program, each student completed a visual history questionnaire that asked about previous vision exam results, eye and hand dominance, significant ocular history, and the use of prescription glasses or contact lenses. The specification of rotating prism used for experiment are listed in Table1. These specifications consist of three parts, eye exercise device, Auto refractor and Auto Refract meter. Due to sophisticated experiments, experimental procedures are divided into two parts as following:

A. *Pre-Experiment: Eye exercises to improve eyesight*

1) Hold a pencil at an arm's length and focus on it.

2) Slowly bring it closer to your nose.

3) Move it farther from your vision until you can no longer keep it in focus.

4) Repeat about 10 times a day.

Table1 the rotary prism eye exercise device

Clear Aperture	90% of central circular dimension
Surface flatness and quality	λ/10 @633nm over clear aperture and 20/10
Material and size	Bk7 and 300mm
Prism power	5-8 power diopters
Lens power	1.00D-2.00D

B. *experiment with the rotary prism eye exercise device*

The eye exercises were grouped into saccadic eye movements, accommodation, convergence and divergence. Five students participated in the eye exercise by the rotary prism eye exercise device. The student is myopia and the duration of eye exercises is 7 minutes. The exercises process is described below:(a) Wear corrective lenses and an eye exercise device then look at the dynamic modified eye charts in the 6 meters distance.(b)Clockwise rotation for 0.5 minutes, then counterclockwise rotation for 0.5minutes and 7 consecutive periods stop after 7 minutes. (c)Due to dizziness and fatigue during the exercise, the patient needs rest. (d)The eye can become relaxed after resting commonly, feel bright and eye will be flexible. This eye exercise will increase the flexibility of eye muscles and improve eye vision. The data of visual acuity was recorded before and after the eye exercise over 12 weeks is shown as Table 2. After three months, the improvement in visual acuity was average of 0.10D of both right and left eyes decreasing.

Table2 Visual acuity of before and after the eye exercise

Week		1	2	3	4	5~6	7	8	9~10	11~12
VA before	OD	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6
	OS	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.4	0.5
VA after	OD	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5
	OS	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5

IV. Conclution

Studies have indicated that look at a 3C Industry product or near object for extended long time; can lose 1% to 2% of their vision per year. The minimal movement of the eye ball causes the muscles to atrophy resulting in diminishing vision. Children are especially vulnerable and can lose their vision at a rapidly rate. Using the proper eye exercises, the vision lost due to weakened eye muscles about the eye can be regained. After three months, the improvement in visual acuity was average of 0.20D of both right and left eyes decreasing an indicating that eye exercises as eye exercise-based program could improve visual acuity for children with myopia.

References

[1] Vitale S, Ellwein L, Cotch MF, Ferris FL 3rd, Sperduto R. Prevalence of refractive error in the United States,1999-2004, Arch Ophthalmol. 2008;126(8):1111-1119

[2] Richman JE, Cron MT., "Guide to Vision Therapy, " South Bend. IN: Bernell Corp, 1988.

[3] Streff J. Optical effects of plano prism with curved surfaces. J Am Optom Assoc 1972; 44: 717-21.

[4] Smith ,Warren J., Modern Optics Engineering, McGraw Hill, N.Y., 2000.

Merge Mode-based Data Embedding in SHVC Compressed Video

LieLin Pang
Faculty of Comp. Sci. & Info. Tech.
University of Malaya
Kuala Lumpur, Malaysia
adpangll@siswa.um.edu.my

KokSheik Wong
School of Information Technology
Monash University Malaysia
Bandar Sunway, Malaysia
wong.koksheik@monash.edu

Ryoma Ito
Graduate School of Sci. & Tech.
Shinshu University
Nagano, Japan
17w2801d@shinshu-u.ac.jp

***Abstract*—A merge mode-based data embedding technique is proposed for SHVC compressed video. Inter-prediction block merging candidate selection decision is manipulated to embed data based on the pre-defined mapping rules. Experimental results show that encouraging payload is achieved at the expense of slight bit rate increment and negligible degradation in perceptual video quality. In the best case scenario, the sequence *PartyScene* can embed 84.4 kbps with an average 1.1% bit rate overhead for the LDB configuration.**

***Index Terms*—Block merging, data embedding, SHVC, advance motion vector prediction**

I. Introduction

H.264/SVC [1] and Scalable High Efficiency Video Coding (SHVC) [2] are the current video coding standards that offer scalable features to support video communication over varying network conditions, unknown bandwidth, video of different quality, as well as devices with different hardware capabilities. As the number of video content increases, these scalable coded videos need to be managed effectively. Data embedding technique offers several ways to manage videos, including copyright protection, authentication, fingerprinting, hyper-linking related videos and tamper detection. Here, different data can be embedded in different ways to serve different purposes.

The conventional methods for HEVC and previous video standards manipulate the syntax elements in intra prediction, motion compensated prediction, as well as quantized coefficient to embed data [3]. However, these techniques cannot be directly applied to SHVC because SHVC exploits inter-layer redundancy by utilizing the reconstructed picture and motion information of the base layer (BL) picture as references to predict those in the enhancement layers (ELs). Our previous work [4] puts forward an error compensation embedding technique during intra prediction by cancelling errors introduced in the reference layer for SHVC video. Recently, we also manipulate the parity bit of the motion vector prediction indices in SHVC and utilize threshold-guided technique to improve payload [5].

II. Proposed Data Embedding Technique

In this work, the inter-prediction block merging candidate selection decision is manipulated to embed data based on

This work was partly supported by Fundament Research Grant Scheme (FRGS/1/2018/ICT02/MUSM/02/2) by Ministry of Education Malaysia.

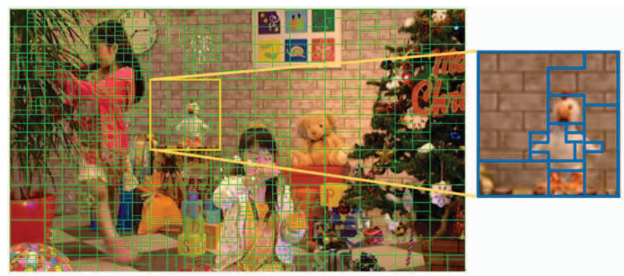


Fig. 1. Quad-tree Partitioning and Merging Block Structure.

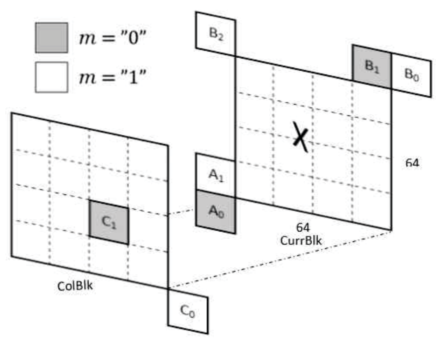
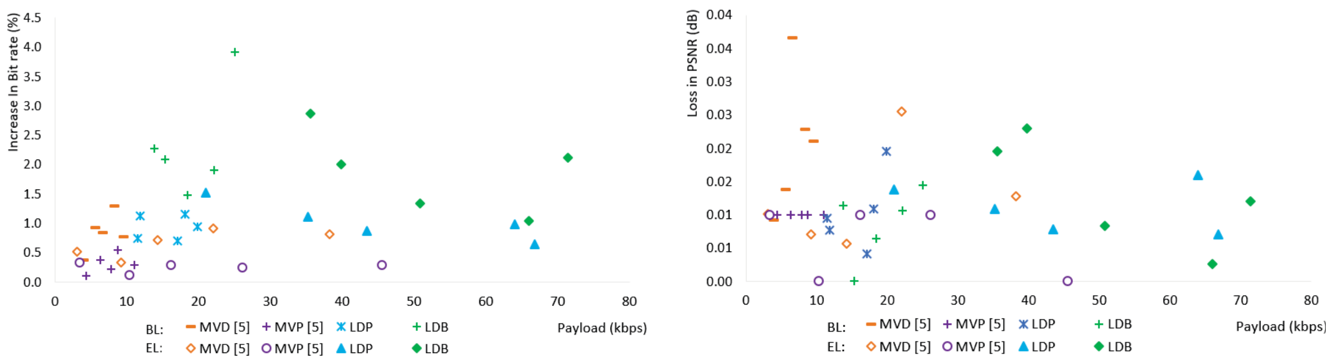


Fig. 2. Example of Mapping Rule for Merge candidates.

some pre-defined mapping rules. Note that adjacent blocks containing object with similar motion are likely to be predicted by using the same motion vector. Hence, merge mode in SHVC reuses motion information from the adjacent blocks. An illustration of possible merge block is depicted in Fig. 1, where the merge blocks are marked with blue borders. The advantage of manipulating merge block for data embedding is that the bit rate overhead can be kept to the minimum while payload can be increased because the number of blocks associated with skip/merge CUs is $> 50\%$ as compared to other prediction mode [6].

To embed data, merge candidates are divided into two groups, where one is associated with bit '0' and the other associated with bit '1'. An example of the mapping rule is shown in Fig. 2. To determine the merge candidate block for the current block (*CurrBlk*), the neighboring blocks which are

978-1-7281-3038-5/19/$31.00 ©2019 IEEE



(a) bit rate.

(b) video quality.

Fig. 3. Changes in bit rate and video quality after data embedding.

mapped to the message bit m (dependent on mapping rule) are considered, where the one with the best rate distortion (RD) cost is coded. The proposed method follows the steps below to embed data into a SHVC video:

1) Construct a list of merging candidates from spatial, temporal and inter-layer candidates as shown in Fig. 2.
2) For each merging candidate, check if the candidate block is mapped to m. If YES then go to step 3, otherwise proceed to the next merging candidate.
3) Perform motion compensation and residual prediction, then calculate RD cost.
4) If the RD cost is less than the best RD cost, then set the current block as the best merge candidate.
5) Repeat step 2$\sim$ 4 until the best merge candidate is selected for the PU.

III. Experiments

The proposed data embedding technique is implemented by modifying the SHVC reference software SHM-12.0 [7]. Five standard test video sequences (i.e., *RushHour*@30Hz (960×540, 1280×720), *FourPeople*@60Hz and *BlueSky*@24Hz (640×360 , 1280×720), and *BasketballDrill*@50Hz and *PartyScene*@50Hz (416×240, 832×480)) are utilized for evaluation. Experiments are conducted using two layers spatial scalability with GOP of 4. The zero motion setting for inter-layer reference picture is enabled. The remaining parameters are set to the SHM's default scalable configuration for LDP (low delay P) and LDB (low delay B).

The impact of data embedding to the bit rate for the proposed (labelled as LDP and LDB) and our previous method [5] are shown in Fig. 3(a). Note that [5] is based on *MVP*, and it manipulates the indices of motion vector predictor while the *MVD*-based technique manipulates the difference between motion vectors to embed data. Results suggest the proposed method achieves higher payload when compared to the *MVP*- and *MVD*-based techniques at the expense of some bit rate overhead. This is because there are more prediction blocks associated with skip/merge mode, which can provide more syntax elements for manipulation. Lower bit rate overhead is observed for [5] due to its low embedding capacity. The effect of the embedded payload on the video quality in term of PSNR is shown in Fig. 3(b). Both techniques show comparable video quality, and the degradation is negligible. For all cases, the drop is PSNR value is $\leq 0.04\%$ and the drop in SSIM value is insignificantly. In the best case scenario, the perceptual quality for the sequence *PartyScene* drops at most by 0.01 dB while embedding up to 84.4 kbps with an average 1.1% bit rate overhead for the LDB configuration.

IV. Conclusion

A merge mode-based data embedding technique is proposed for SHVC video. Merging candidate is chosen based on the predefined mapping rules to embed data without significantly compromising the perceptual video quality. Experimental results show that payload can be increased at the expense of higher bit rate as compared to [5].

As future work, we want to jointly utilize the other syntax elements in SHVC to further improve of payload and refine our work for actual application.

References

[1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h.264/avc standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sept 2007.
[2] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20–34, Jan 2016.
[3] Y. Tew and K. Wong, "An overview of information hiding in h.264/avc compressed video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 2, pp. 305–319, Feb 2014.
[4] L. Pang, K. Wong, and S. T. Liong, "Data embedding in scalable coded video," in *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Dec 2017, pp. 1190–1194.
[5] L. Pang and K. Wong, "A data embedding technique for spatial scalable coded video using motion vector predictor," in *2019 IEEE International Conference on Image Processing (ICIP)*, Sep. 2019, pp. 4050–4054.
[6] P. Helle, S. Oudin, B. Bross, D. Marpe, M. O. Bici, K. Ugur, J. Jung, G. Clare, and T. Wiegand, "Block merging for quadtree-based partitioning in hevc," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1720–1731, Dec 2012.
[7] SHM-12.0. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_SHVCSoftware/tags/SHM-12.0

User Trajectory Analysis within Intelligent Social Internet-of-things (SIoT)

Guang Xing Lye, Wai Khuen Cheng, Teik Boon Tan, Chen Wei Hung
Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman
Perak, Malaysia
simple@1utar.my, chengwk@utar.edu.my, tantb@utar.edu.my, hungcw@utar.edu.my

Yen-Lin Chen
Dept. of Computer Science and Information Engineering
University of Technology Taipei, Taiwan
ylchen@csie.ntut.edu.tw

***Abstract*— Despite the advancement of Internet-of-Things (IoT) and social network, one of the main challenges in SIoT domain is intelligent service discovery and composition. This paper presents an SIoT architecture with personalized recommendation in order to deliver better service discovery. Our proposed approach has outperformed other methods during the experiments.**

Keywords—Social Internet-of-Things (SIoT), user trajectory analysis, SIoT architecture, smart campus, personalized recommendation

I. INTRODUCTION

Social IoT (SIoT) is an emerging paradigm of IoT in which heterogeneous IoT devices are able to communicate, collaborate on behalf of their owners, establish relationships based on common interest, and perform service trading autonomously. In common, SIoT is expected to enhance the existing distributed system's features such as service discovery and composition, information management, and service trustworthiness management [1]. Although the adoption of SIoT is started to be used in some domains like smart vehicles, smart home and integrated transportation, it is still noted that the current SIoT systems encountered a number of challenges that affect their usability and reliability in several SIoT domains [2].

Personalization and recommendation are two important prerequisites that must be incorporated into the SIoT systems in order to deliver a promising service [3]. Both prerequisites are essential to produce a higher satisfaction level of SIoT solution that match the preferences of the user [4]. According to our previous work in paper [5], location-based smart information system could utilize the mobile trajectories of the users to recommend several point-of-interests according to user preferences and conditions. The trajectory data are obtained from various sensors embedded in the mobile phone, as well as the smart environment (e.g. devices, building and check-in points). Many existing solutions in the literature focus on improvising at the service layer, recommendation to the users is mainly based on static information, such as preload service details (e.g. location and type) in an area and current user position. Many times, limited choices are generated which does not characterize the actual human needs.

Our work is different, as we proposed a trajectory analysis framework which can apply trajectories of the users with similar behavior and movement patterns, and subsequently integrating user's beliefs for a personalized service recommendation. We have leverage the available trajectory-based and contextualized data for our personalized recommendation. The preliminary results showed the applicability of the proposed framework within a smart campus application, UniCAT [5]. However, the applicability of the proposed user trajectory framework crossing several SIoT domains are yet to examine.

The major contribution of this article is to present a generic architecture of SIoT for the smart campus which provides the reader with an understanding on the key features of this solution, with particular attention to an SIoT use case scenario.

II. SOCIAL INTERNET-OF-THINGS

A. SIoT Architecture

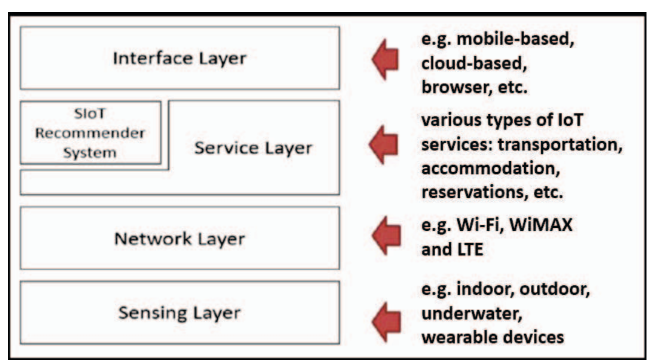


Fig. 1. The proposed SIoT architecture with recommender module.

In general, we adopt the SIoT architecture as illustrated in Fig. 1 for the system development [4]. Sensing devices at the bottom layer capture and transmit data to different IoT services through different network mechanism, such as LTE, WiMAX, WiFi, and satellite internet. Users have access to different IoT services by navigating through the respective Social Network

Service (SNS) interface layer. Practically, the selected SNS interface should forward the appropriate user's request to the SIoT recommender system in order to trigger the relevant IoT services. The recommender system is not only activating by system call but also matching the user profile with other related SNS users, and subsequently recommend any possible solutions to the service requestor. The following is a sample scenario describing the application of SIoT in the campus.

B. *Use Case Scenario*

Michael is a freshman in the university. He is not used to the environment in several issues. For example, access to accommodation, campus facilities – library, sports, etc. and transportation. At the beginning, everything is a mess to him simply because he has no idea whom to refer during the difficult situations. By exploiting various social relationships (course-mates and other campus communities) which develop through several IoT applications within the area, Michael finds himself was able to adapt to the environment more easily. A smart campus app which provides personalized recommendation regarding the best place to stay, to eat, visit or do sport, is highly needed. The smart campus app could provide recommendation to the users by considering several factors such as places near to his friends, facilities in the area, and the accessibility of transportation. By matching his needs or profile with the seniors with similar behavior and perhaps trajectory patterns, the app is able to remind and suggest the freshman with info such as the best bus route, the most preferred restaurants and so on. This is achieved by adopting a reliable intelligent SIoT platform within the campus.

III. User Trajectory Analysis

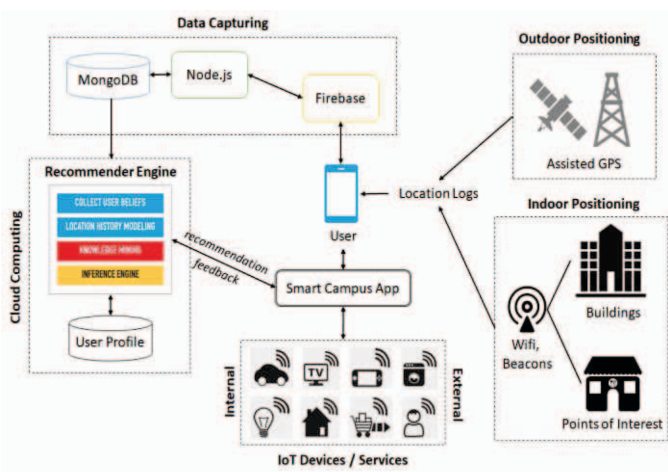


Fig. 2. The overall SIoT architecture with different sensing inputs, data capturing module, recommender engine and internel/external IoT services.

Fig. 2 illustrates how the overall implementation of personalized user trajectory analysis within a smart campus. User movement trajectories are mainly gathered through the indoor and outdoor positioning. Data from the location logs will be delivered to the Firebase (mobile & web platform) for further processing. The figure also shows that the internal/external IoT devices and services are able to communicate with each other through web service call (e.g. REST and SOAP), as well as the smart campus app. Recommendations to the user are generated by the recommender engine with possible integration of various inputs from the external SIoT services. For instance, recommend a dining place through the trajectory analysis meanwhile referring to the restaurant ranking provided by the Foursquare in the area. Also, user personal preferences could be taken part for further recommendation filtering.

IV. Experimental Results

In order to examine the user satisfaction of the proposed trajectory analysis in SIoT, we used 2 datasets, namely UniCAT and Weeplaces, for precision-recall analysis. Food recommendations are generated from the smart campus app and further verified with the user satisfaction. From the experimental results as shown in Fig. 3, our proposed SIoT architecture (*Hybrid*) has successfully delivered the personalized recommendation with higher precision and recall values, 0.43 and 0.62, respectively. Two other approaches, Link Analysis (*LA*) and Rank by Frequency (*RF*) are selected as the baseline methods.

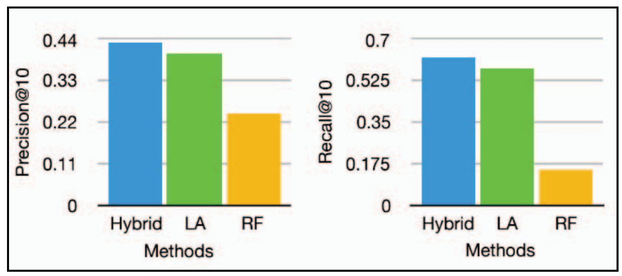


Fig. 3. Comparison of precision-recall values between different recommendation approaches – Hybrid, LA and RF.

V. Conclusion

This project presented the feasibility of our proposed SIoT architecture in a smart campus. Personalized recommendation with further inputs from relevant autonomous SNS certainly produce higher satisfaction to users. Due to the complexity of social and IoT structures, service discovery and composition in different domains can be the extension of current work.

References

[1] L. Atzori, A. Iera, and G. Morabito, "From "smart objects" to "social objects": The next evolutionary step of the internet of things," Communications Magazine, IEEE, vol. 52, no. 1, pp. 97–105, 2014.

[2] L. Atzori, A. Iera, and G. Morabito, "Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm," Ad Hoc Networks, vol. 56, pp. 122–140, 2017.

[3] Roopa M.S., Santosh Pattar, Rajkumar Buyya, Venugopal K.R., S.S. Iyengar, L.M. Patnaik, "Social Internet of Things (SIoT): Foundations, thrust areas, systematic review and future directions," Computer Communications, vol. 139, pp. 32-57, 2019.

[4] W.K. Cheng, A.A. Ileladewe, and T.B. Tan, "A Personalized Recommendation Framework for Social Internet of Things (SIoT)," International Conference on Green and Human Information Technology (ICGHIT), Malaysia, 2019.

[5] G.X. Lye, A.A. Ileladewe, W.K. Cheng and T.B. Tan, "A Personalized Recommendation Framework with User Trajectory Analysis Applied in Location-based Social Network (LBSN)," Proceedings of the 3rd IEEE International Conference on Engineering Technologies and Social Sciences (ICETSS), pp. 1-6, Thailand, 2017.

2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)

Crowd Behavior Classification based on Generic Descriptors

Pei Voon Wong
Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman
31900 Kampar, Perak, Malaysia
wongpv@gmail.com

Norwati Mustapha, Lilly Suriani Affendey , Fatimah Khalid
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.
norwati@upm.edu.my;
lilly@upm.edu.my;
fatimahk@upm.edu.my

Yen-Lin Chen
Department of Computer Science and Information Engineering
National Taipei University of Technology,
1, Sec. 3, Chung-hsiao E. Rd., Taipei, Taiwan 10608
ylchen@csie.ntut.edu.tw

Crowd behavior analysis plays an important role in high security interests in public areas such as railway stations, shopping centres, and airports, where large populations gather. The crowded scenes vary in various densities, structures and occlusion. It brings enormous challenges in identifying generic descriptors to describe motion dynamics caused by pedestrians walk in different directions with extremely diverse behaviors. Therefore, this research is proposal an approach for crowd behavior analysis to recognize the common properties across different crowded scenes. The recognized common properties are then used to identify generic descriptors from group-level for crowd behavior classification.

Crowd behavior classification, generic descriptors, crowded scenes, group-level, motion.

I. Introduction

The study of group or crowd analysis for crowded scenes are generally modelled in low, middle and high levels. At the low-level, moving objects are discovered and tracked to extract the crowd motion features from each of the video frames. Motion features such as particle flow, streak flow, spatio-temporal [1], and trajectory or tracklet are extracted. At the mid-level, motion pattern segmentation is used in crowd analysis by grouping the features into similar categories through some resemblance measures or probabilities. The high-level focuses on discovery of crowd motion descriptors based on low-level motion features in order to facilitate understanding crowd behavior.

Therefore, various crowd motion descriptors such as social force, potential field, chaotic invariants, and collectiveness have been suggested from different perspectives for scene-specific crowd behavior analysis. These crowd motion descriptors are restricted to holistic perspective. Recently, some researchers proposed generic descriptors, such as curl, divergence, collectiveness, uniformity, stability, and conflict, from the computer vision point of view to describe crowd behavior for different crowded scenes [2-3]. However, these generic descriptors cannot perform well for the motion dynamics caused by pedestrians who walk in different directions with extremely diverse behaviors; such as pedestrians in streets or shopping malls.

In addressing the above limitations, we propose a group motion pattern mining and prediction approach to identify generic descriptor from group-level for crowd behavior classification.

II. Generic Descriptors Based on Group Motion Pattern Mining and Prediction

Crowd behavior analysis framework can be divided into low-level, mid-level and high-level. At the low-level, motion feature extraction is performed to extract trajectories from each of the video frames. Kanade-Lucas-Tomasi feature point tracker is used to detect and track moving humans, and then tracklets are grouped to form trajectories. At the mid-level, a Collective Interaction Filtering [4] is presented to identify groups by clustering trajectories. It is suitable for group detection in low, medium, and high crowds. At the high-level, the result of group detection as an input in group motion pattern mining to predict collectiveness, uniformity, stability, and conflict generic descriptors. Group motion pattern mining and prediction is inspired by [5] to predict user movements based on trajectory or tracklet. The trajectory pattern is found to be very valuable in learning the connections between moving objects [5]. Fig. 1 illustrates the proposed group motion pattern mining and prediction approach. Group motion pattern mining includes two main tasks that calculate group size (GS), evenly space (ES), and speed direction (SD) connectivity between members in group as well as create an adjacency matrix, and graph partitioning algorithm for group motion pattern mining. Finally, Kalman filtering is used to predict generic descriptors. The generic descriptors are used in crowd behavior analysis. The generic descriptors identified are represented by graph-based descriptors.

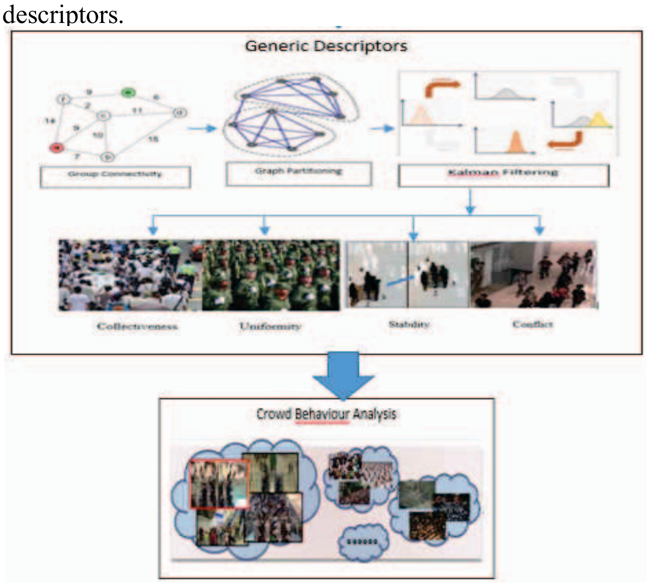


Fig. 1. Group motion pattern mining and prediction approach

III. DATASET

The CUHK Crowd Dataset provided the total 474 video clips with different crowded scenes are manually assigned into 8 classes [2] as shown in Fig. 2. A scene may include multiple video clips. These 8 classes are crowd behaviors that usually happen in the crowd video.

Class	Class Name	Total videos
1	Highly mixed pedestrian walking	15
2	Crowd walking following a main stream and well organized	153
3	Crowd walking following a main stream but poorly organized	72
4	Crowd merge	9
5	Crowd split	13
6	Crowd crossing in opposite directions	70
7	Intervened escalator traffic	21
8	Smooth escalator traffic	121

Fig. 2. List of crowd behavior classes

IV. EXPERIMENTAL RESULTS

The performance of the proposed graph-based descriptors is compared to holistic features [1], object-based features [6], group descriptors [2] and bilinear CD descriptors [3] to classify crowd video clips.

Leave-one-out evaluation protocol is used for crowd video classification. Average accuracy in confusion matrix over all video classes are the performance metrics employed in this experiment. For each experiment, each scene is selected as test set and training is performed on remaining scenes. If a video has various groups, average of the descriptor over groups is taken to represent crowd behavior for a video. Non-linear SVM with RBF-kernel classifier is employed for classification. Crowd video classification results are shown in Fig. 3 in confusion matrix. Fig. 3 (a) shows holistic features [1] and the average accuracy is 44%. Fig. 3 (b) shows object-based features [6] and the average accuracy is 41%. Fig. 3 (c) shows group descriptors [2] and the average accuracy is 70%. Fig. 3 (d) shows bilinear CD descriptors [3] and the average accuracy is 76%. Fig. 3 (e) shows the proposed graph-based descriptors and the average accuracy is 80%. The results show that the average accuracy of the proposed graph-based descriptors is higher compared to the others. The proposed graph-based descriptors performs better on the 1, 3, and 5 crowd behavior classes as shown in Fig. 4. The proposed graph-based descriptors benefits from not sensitive to the motion dynamics as shown in class 1. In class 3, the proposed graph-based descriptors also solve the occlusion caused by non-human moving object. Finally, the proposed graph-based descriptors advances average accuracy of class 5 by solving the correct track of each person involved in the occlusion after they split up from crowd merge.

V. CONCLUSION

This paper presents a proposed group motion pattern mining and prediction approach to identify collectiveness, uniformity, stability, and conflict generic descriptors based on interaction among people in group for crowd behavior classification. The effectiveness of the graph-based descriptors for crowd behavior analysis has been evaluated using CUHK Crowd Dataset and compared against other crowd features and descriptors. The experimental results have shown improved results. The complete trajectories are hard to be gained in crowded scenes due to occlusion. Therefore, we will use deep learning techniques to solve the generic description to describe crowd behavior for different crowded scenes accurately in future work.

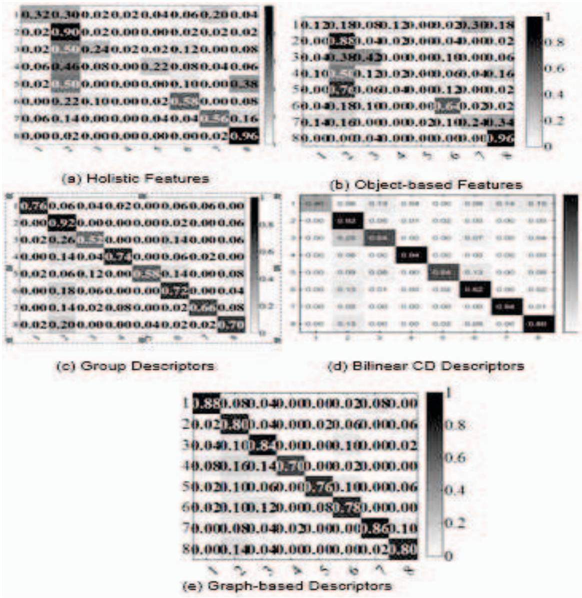


Fig. 3. Confusion Matrices of Crowd Video Classification

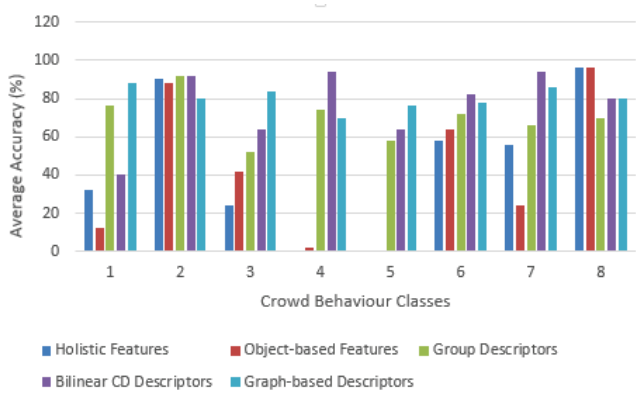


Fig. 4. The Quantitative Average Accuracy Comparison of the Proposed Graph-based Descriptors with Previous Works

ACKNOWLEDGMENT

This work is supported by the Fundamental Research Grant Scheme (FRGS) Malaysia (Project No.: 08-01-17-1918FR) and the Ministry of Science and Technology of Taiwan (Project No.: MOST-107-2218-E-027-018).

REFERENCES

[1] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," *2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009*, pp. 1446–1453, 2009.

[2] J. Shao, C. C. Loy, and X. Wang, "Learning Scene-Independent Group Descriptors for Crowd Understanding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 6, pp. 1290–1303, Jun. 2017.

[3] S. Wu, H. Su, H. Yang, S. Zheng, Y. Fan, and Q. Zhou, "Bilinear dynamics for crowd video analysis," *J. Vis. Commun. Image Represent.*, vol. 48, pp. 461–470, Oct. 2017.

[4] P. V. Wong, N. Mustapha, L. S. Affendey, and F. Khalid, "Collective Interaction Filtering Approach for Detection of Group in Diverse Crowded Scenes," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 2, pp. 912–928, Feb. 2019.

[5] M. Jalali, N. Mustapha, A. Mamat, and M. N. B. Sulaiman, "A new clustering approach based on graph partitioning for navigation patterns mining," in *2008 19th International Conference on Pattern Recognition*, 2008, vol. 4, no. 11, pp. 1–4.

[6] H. Idrees, N. Warner, and M. Shah, "Tracking in dense crowds using prominence and neighborhood motion concurrence," *Image Vis. Comput.*, vol. 32, no. 1, pp. 14–26, Jan. 2014.

Actived Edge Strength for Image Quality Assessment

Minjuan Gao
School of Electrical and Control Engineering
ShaanXi University of Science and Technology
Xi'an China
gaominjuan1984@163.com

Xuande Zhang
Shool of Electronic informaiton and Artifical Intelligence
ShaanXi University of Science and Technology
Xi'an China
love_truth@126.com

***Abstract*—These years has witnessed the success of deep learning methods in computer vision. The approximation capabilities of neural networks is partly responsible for these success, and active function is crucial for the approximation capability. Motivated by the success of deep learning, this paper presents an Actived Edge Strength Similarity (AESSIM) based image quality assessment algorithm. Numerical experiments on the public datasets indicates that AESSIM is quite competitive in assessing performance.**

Keywords—Image quality assessment, active function, edge strength

I. INTRODUCTION

Image quality assessment (IQA) has received extensive interest in image processing field in recent years, and there are hundreds of papers published each year devoted to this problem. The goal of IQA is to develop a mathematical model to mimic the function of human visual system (HVS) on evaluating image quality. According to the availability of the ground-truth images, IQA can be classified into full-reference (FR), reduced-reference (RR) and no-reference (NR) methods. This paper will focus on FR method.

The fundamental difficulty of IQA lies in the fact that HVS is an extremely complex nonlinear system that integrates sensing, understanding and discriminating functions and can perform multiple tasks, and the working mechanism of this system is far from clear to us at present. This means that IQA is a black-box-system modeling problem. By now, there are two categories of information can be exploited for modeling this system：1. The incomplete knowledge about HVS; 2. The input-output instance data (sample) generated by HVS when performing image quality evaluation tasks. Here, the input is the image pair composed of the reference image and the corresponding degraded image, and the output is the subjective evaluation score of the degraded image, that is, the mean opinion score (MOS). Exploitation of two categories of information naturally lead to two type of IQA methods: the first uses knowledge or assumptions about HVS to heuristically construct image quality models[1-6]; the second uses deep neural networks to model HVS and approximate the mapping from input to output, in which the network parameters are optimized according to data[7]. The first always refers to traditional methods and the second refers to deep learning methods. Naturally, two type of methods can shed light to each other. That is to say, there exist two research paths: 1. Integrate the knowledge about HVS into the deep neural network, and study knowledge-driven deep learning method; 2. Investigate the feature representation manner in deep learning (DL), and study the DL-inspired traditional methods. In other words, the first path is from traditional methods to deep learning methods, and the second path is from deep learning methods to traditional methods. The deep learning method has achieved great success in many tasks such as speech recognition, natural language processing, image recognition and classification, and even triggered a new round of artificial intelligence revolution. This makes the first path a mainstream trend in current IQA research. In this paper, we explore the second research path.

The active function plays a crucial role in DL method since the existence of it ensure the nonlinear approximation ability of the network. Motivated by the success of deep learning, we utilize active function to improve the Edge Strength SIMilarity (ESSIM) algorithm[5], and propose an Actived Edge Strength Similarity based image quality assessment algorithm. Numerical experiments on the public datasets indicates that the proposed algorithm is quite competitive in assessing performance. The rest of the paper is arranged as follows. Section II first describes the ESSIM algorithm, and then gives the definition of AESSIM. Section III presents numerical experiments; Section IV summarizes this paper.

II. ACTIVED EDGE STRENGTH BASED IMAGE QULAITY METRIC

A. Edge strength based metric

The construction of ESSIM is described below. The reference image is denoted as

$$f = \left[f_1 ,\ \cdots,\ f_i ,\ \cdots,\ f_N \right],$$

and its corresponding distorted image is denoted as

$$g = \left[g_1 ,\ \cdots,\ g_i ,\ \cdots,\ g_N \right],$$

in which i indexes the pixels and N denotes the total number of pixels. let ∂f_i^j denote the directional derivative at the i th pixel of reference image along the direction indexed by j . ESSIM only take horizontal, vertical, and diagonal changes into account, namely, $j = 1,2,3,4$. The edge-strength in the vertical-horizontal direction is defined as

$$E_i^{1,3}(f) = |\partial f_i^1 - \partial f_i^3|^p$$

and the edge-strength in the diagonal directions is similarly defined as

$$E_i^{2,4}(f) = |\partial f_i^2 - \partial f_i^4|^p ,$$

The total edge-strength around the i th pixel takes the following form

$$E(f,i) = \max(E_i^{1,3}(f), E_i^{2,4}(f)) .$$

To ensure that the edge-strength of f and g at each pixel can be compared in the same direction, the edge-strength of g at the i th pixel is defined as

TABLE I. THE ASSESSING PERFORMANCE OF EIGHT IQA METRICS, THE FIRST TWO BEST RESULTS ARE HIGHLIGHTED.

Dataset	Index	PSNR	SSIM	MS-SSIM	IW-SSIM	FSIM	ESSIM	GMSD	AESSIM
TID 2013	SROCC	0.6396	0.7417	0.7859	0.7779	0.8015	0.8005	**0.8045**	**0.8099**
	KROCC	0.4698	0.5588	0.6047	0.5977	0.6289	0.6284	**0.6331**	**0.6388**
TID 2008	SROCC	0.5531	0.7749	0.8542	0.8559	0.8805	0.8843	**0.8906**	**0.9027**
	KROCC	0.4027	0.5768	0.6568	0.6636	0.6946	0.7049	**0.7090**	**0.7242**
CSIQ	SROCC	0.8058	0.8756	0.9133	0.9213	0.9242	0.9326	**0.9571**	**0.9533**
	KROCC	0.6084	0.6907	0.7393	0.7529	0.7567	0.7686	**0.8122**	**0.8036**
Overall	SROCC	0.6390	0.7727	0.8266	0.8240	0.8447	0.8466	**0.8545**	**0.8606**
	KROCC	0.4709	0.5848	0.6416	0.6420	0.6689	0.6736	**0.6841**	**0.6905**

$$E(g,i)=\begin{cases}E_i^{1,3}(g) & \text{if } E(f,i)=E_i^{1,3}(f)\\ E_i^{2,4}(g) & \text{if } E(f,i)=E_i^{2,4}(f)\end{cases},$$

where $E_i^{1,3}(g)$ and $E_i^{2,4}(g)$ have the same form as $E_i^{1,3}(f)$ and $E_i^{2,4}(g)$.

Finally, ESSIM is defined as

$$ESSIM(f,g)=\frac{1}{N}\sum_{i=1}^{N}\frac{2E(f,i)E(g,i)+C}{\left(E(f,i)\right)^2+\left(E(g,i)\right)^2+C},$$

where C is a parameter.

B. Actived edge strength based metric

ESSIM contains parameters p and C , where p rescale the edge-strength and C adjust the similarity computation. Here we will employ sigmoid function to rescale the edge-strength and adjust the similarity computation, then these two parameters can be dislodged, namely, let $p=1$ and C be a very small positive real number used only to avoid a zero denominator in the proposed algorithm. Sigmoid function is selected since its mathematic properties are closely related to the characteristics of HVS. To enhance the approximation ability of the proposed algorithm, we introduce a parameter k to sigmoid function as the follow

$$S(x,\mathrm{k})=1/(1+\exp(-\mathrm{k}\cdot x)).$$

We first nonlinearly rescale the edge-strength $E(f,i)$ and $E(g,i)$ by using $S(x,\mathrm{k}_1)$ and compute the local similarity, then this similarity is adjusted by using $S(x,\mathrm{k}_2)$ and followed by a pooling. The final AESSIM metric have the following form

$$AESSIM(f,g)=\frac{1}{N}\sum_{i=1}^{N}S\left(\frac{2S(E(f,i),\mathrm{k}_1)S(E(g,i),\mathrm{k}_1)+C}{\left(S(E(f,i),\mathrm{k}_1)\right)^2+\left(S(E(g,i),\mathrm{k}_1)\right)^2+C},\mathrm{k}_2\right).$$

in which k_1 ,k_2 are parameters related to the shape of sigmoid function.

III. NUMERICAL EXPERIMENTS

The proposed AESSIM is evaluated on the following three publicly available datasets: TID2013[8], TID2008[9] and CSIQ[10]. The comparison algorithm includes PSNR, SSIM[1], MS-SSIM[2], IW-SSIM[3] FSIM[4], GMSD[6] and ESSIM[5]. The assessing performance of these algorithm is measured by Spearman rank-order correlation coefficient (SROCC) and Kendall rank-order correlation coefficient (KROCC). To ensure a fair comparison, the standard parameters configured by their authors are utilized for these comparing algorithms. For AESSIM, the parameter is empirically tuned on the subset of TID2008 composed of the first six reference images and the corresponding distorted images. we tried several parameters handily and choose the value that leading to the highest SROCC. As a result, the parameters are set as $\mathrm{k}_1=0.06$, $\mathrm{k}_2=16$. For color images, all IQA metric included for comparison are evaluated on the luminance component, and the luminance component is extracted by YIQ color space transformation. The SROCC and KROCC performance of eight comparing algorithms are listed in TABLE.I. We can observe that the proposed ASSIM can achieve the best performance on these datasets.

IV. CONCLUSION

Inspired by the success of DL, this paper proposed an AESSIM algorithm for image quality assessment, in which the sigmoid function is employed to adjust the edge-strength and similarity computation. Numerical experiments indicate that AESSIM can significantly outperform ESSIM in prediction accuracy, which imply the potential and feasibility to study the DL-inspired traditional methods.

REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli.“Image quality assessment: From error visibility to structural similarity,”IEEE Trans. Image Process., vol. 13, no. 4, pp. 600-612, Apr. 2004.

[2] Z. Wang, E. P. Simoncelli, and A. C. Bovik.“Multi-scale structural similarity for image quality assessment,”in Proc.IEEE Asilomar Conf. Signals, Syst. Comput., Nov. 2003, pp. 1398-1402.

[3] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” IEEE Trans. Image Process., vol. 20, no.

[4] L. Zhang, L. Zhang, X. Mou, and D. Zhang.“FSIM: A feature similarity index for image quality assessment,”IEEE Trans. Image Process., vol. 20, no. 12, pp. 2378-2386, Aug. 2011.

[5] X. Zhang, X. Feng,W. Wang , and W.Xue.“Edge strength similarity for image quality assessment,”IEEE Signal Process. Lett., vol. 20, no. 4, pp. 319-322, Jan.2013.

[6] W.Xue , L Zhang , X Mou , and A. C. Bovik.“Gradient magnitude similarity deviation: A highly efficient perceptual image quality index,”

IEEE Trans. Image Process., vol. 23, no. 2, pp. 684-695, Dec.2014.

[7] S Bosse ,D Maniry , T Wiegand , et al.“A Deep Neural Network for Image Quality Assessment,”in Proc. IEEE International Conf. on Image Process,Phoenix,pp.2381-8549,2016.

[8] N. Ponomarenko, L Jin , O Ieremeiev, et al.“Image database TID2013:peculiarities, results and perspectives.Signal,” Process Image Commun, vol. 30, pp. 57-77, 2015.

[9] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti.“TID2008—A database for evaluation of full-reference visual quality assessment metrics,”Adv. Modern Radioelectron., vol. 10, pp. 30-45, 2009.

[10] C. Larson and D. M. Chandler.“Categorical Image Quality (CSIQ) Database 2009,”. Available: http://vision.okstate

SRGNet: A GRU Based Feature Fusion Network for Image Denoising

Wenhao Wang[1], Cheng Pang[1], Zhenbing Liu[1], Rushi Lan[1], Xiaonan Luo[2]
[1] Guangxi Key Laboratory of Image and Graphic Intelligent Processing,
Guilin University of Electronic Technology, Guilin, China.
[2]National Local Joint Engineering Research Center of Satellite Navigation and Location Service,
Guilin University of Electronic Technology, Guilin, China.
Corresponding author: pangcheng3@126.com

Abstract—**Image denoising is an essential pretreatment for most of image processing pipelines, which has been extensively studied for several decades. Recently, convolutional neural networks with skip connections show promising performances on image denoising due to their discriminative feature modeling and utilizing of features from former layers. However, they only apply coarse feature fusion strategies like pixel-wise addition or concatenation which are insufficient in image denoising. In this work, we propose a novel network architecture to exploit finer feature fusion. Specifically, a module based on gate recurrent unit is introduced into the architecture, fusing features from different layers and adopting finer feature selection at the same time. Experiments on multiple challenging datasets show the effectiveness of the proposed network.**

Index Terms—**Image Denoising, Convolutional neural network, Residual Learning, Gate Recurrent Unit**

I. Introduction

Image denoising is adopted in mainstream image processing pipelines for effectively removing of noises and obtaining clean images. Thus, results of an image denoising algorithm have strong impacts on final outputs. In the past several decades, numerous denoising algorithms have been proposed to recover noise-free images, for example, BM3D [1], DnCNN [2], FFDNet [3] and DRCNN [4]. Among them, Convolutional Neural Network (CNN) based algorithms boost the performance significantly. Thanks to their bio-inspired structure, CNNs are able to build feature information from low-level to high-level progressively. However, the feed forward network architecture can't take full usage of features extracted by different layers. And the existing feature fusion techniques are usually simple pixel-wise addition or concatenation without selection strategy.

II. Proposed Method

The key novelties of the proposed method, namely SRGNet, are two folds: a specialized network architecture for image denoising and a new SRG block used for feature extraction and fusion. Our network incorporates multiple SRG blocks which substantially improve denoising performance. On the one hand, the squeeze-and-excitation (SE) [5] layer is adopted to model the weight of channels. On the other hand, the gate recurrent unit (GRU) [6] is introduced to apply gating mechanism to features.

Network Architecture: The proposed network architecture can be divided into three parts: an image feature extraction module, four SRG blocks and a reconstruction module (Fig. 1). A noisy image is firstly fed into the image feature extraction module implemented by a single convolution (Conv) layer, in which non-linear function is eliminated to avoid the loss of low-level information. Then, the output features are passed directly to the first SRG block for better feature abstraction. The GRU module within SRG block not only takes information from distinct layers, but also introduces finer feature selection. After that, three Conv layers reconstruct the estimated noise, serving as the reconstruction block. The clean image is finally obtained by subtracting the estimated noise from the input image. Given a noisy image $x = y + n$ corrupted by noise n, the above pipeline can be formulated as $\hat{y} = x - R(S_4(\ldots S_1(E(x))))$, where E and R denote the feature extraction layer and reconstruction block respectively, S_i stands for ith SRG block, and $\hat{y}$ is the estimation of clean image y.

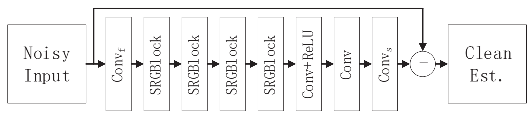


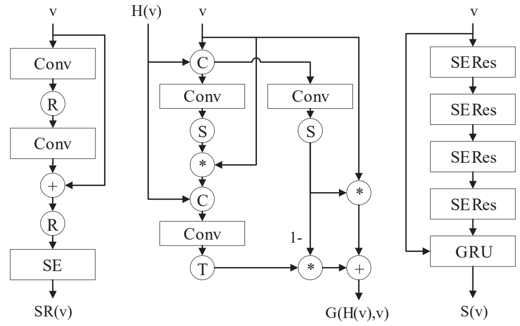
Fig. 1. Overall network architecture.

SRG Blocks: We enclose four SERes blocks in a SRG block (Fig. 2). Each SERes block is a residual block followed by a SE layer for strengthen feature expression via channel-wise recalibration. Given a feature v, the output of one SERes block is $SR(v) = SE(Res(v))$, where Res and SE stand for a residual block and a SE layer respectively.

The detail of SRG block shows in Fig. 2, in which circles with R, S, T, C, + and * stand for ReLU, Sigmoid, Tanh, concatenation, pixel-wisely addition and pixel-wisely multiplication respectively. Additionally, a gating mechanism is applied before output of this block via GRU. As illustrated in the Fig. 2(b), the input of SRG block is taken as previous hidden state, fusing with features from last SERes block to achieve

This work was partially supported by the National Natural Science Foundation of China (Nos. 61772149, 61962014, and U1701267), and Guangxi Science and Technology Project (Nos. AD18216004 and AD18281079).

finer feature selection. Specifically, the SRG block computes: $S(v) = G(H(v), v)$ and $H(v) = SR_4(\ldots SR_1(v))$.



(a) a SERes block; (b) a GRU; (c) a SRG block;

Fig. 2. Detail of a SRG block.

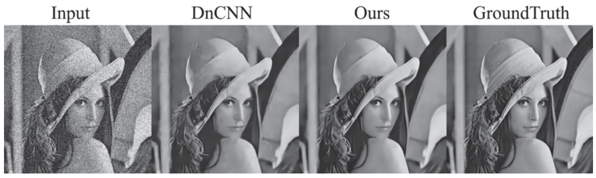


Fig. 3. Denoising comparisons of DnCNN and our method on Lena with noise level $\sigma = 50$.

TABLE I
AVERAGE PSNR(DB) SCORES OF DIFFERENT METHODS FOR GAUSSIAN DENOISING WITH NOISE LEVELS σ=15,25 AND 50 ON SET12, BSD68 AND URBAN100.

Dataset	σ	BM3D	DnCNN	FFDNet	SRGNet
Set12	15	32.37	32.86	32.75	33.03
	25	29.97	30.44	30.43	30.66
	50	26.72	27.18	27.32	27.51
BSD68	15	31.08	31.73	31.63	31.81
	25	28.57	29.23	29.19	29.33
	50	25.60	26.23	26.29	26.38
Urban100	15	32.34	32.67	32.42	33.09
	25	29.70	29.97	29.92	30.54
	50	25.94	26.28	26.52	27.01

TABLE II
PERFORMANCE CONTRIBUTION TO OUR NETWORK

Benchmark	Dataset	SRGNet	SRGNet_1	SRGNet_2
PSNR	Set12	27.51	27.42	27.19
	BSD68	26.38	26.33	26.20
	Urban100	27.01	26.78	26.30
SSIM	Set12	88.49%	88.32%	87.71%
	BSD68	83.02%	82.89%	82.37%
	Urban100	88.86%	88.38%	87.17%

III. EXPERIMENT

The dataset used for training consists of 400 images from BSD [7] and 500 images from DIV2K [8]. The images from DIV2K are down-sampled by factor 4 using bicubic. The input of our network is $y = x + n$ where y, x and n denote the noisy observation, clean image and additive white Gaussian noise (AWGN) at certain noise level respectively. These images are cropped to 96×96 patches for training the model. We train the network for 180 epochs using Adam optimization algorithm, with an initial learning rate of 0.001 which is decayed by 0.2 every 30 epochs. The number of output channels in all layers is set to be 64, except the last layer whose output is noise estimation. We test our network on AWGN corrupted images with zero mean and standard deviation $\sigma = 15, 25$ and 50 respectively on BSD68, Set12 and Urban100 datasets.

Fig. 3 illustrates the denoising result by DnCNN and SRGNet on Lena, an image from Set12, with noise level $\sigma = 50$. It can be seen that our model provides better result and less artifact than DnCNN.

The average PSNR scores of different methods on three test datasets are shown in Table I. As one can see, our SRGNet achieves the best results against all competitors. Specifically, our method improves the BM3D by 0.6 dB with $\sigma = 25$, and outperforms the DnCNN by 0.2 dB on Set12. For the more challenging BSD68, SRGNet achieves scores higher than the best of competitors by 0.08, 0.1 and 0.15 dB respectively. On Urban100, our model outperforms the competitors by 0.4 to 0.7 dB.

Table II illustrates the denoising results on BSD68 and Set12 dataset with the noise level of 50, using PSNR and SSIM benchmarks. The SRGNet_1 and SRGNet_2 are SRGNet without SE layer and GRU block respectively. As it is seen from the table, the feature fusion strategy with GRU block boosts the performance to the most extent due to its finer feature selection mechanism.

IV. CONCLUSION

In this paper, we propose a CNN-based network for image denoising, in which SRG block, a feature fusion strategy, is adopted to enhance the network performance. The gating mechanism is introduced by SRG block for finer feature selection instead of straightforward feature concatenation, improving the feature expression. In addition, SE layer is used for explicitly modeling weights of channels. According to the experiments on benchmark datasets, the proposed network achieves satisfactory visual and quantitative results.

REFERENCES

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug 2007.

[2] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, July 2017.

[3] K. Zhang, W. Zuo, and L. Zhang, “Ffdnet: Toward a fast and flexible solution for cnn-based image denoising,” *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[4] R. Lan, H. Zou, C. Pang, Y. Zhong, Z. Liu, and X. Luo, “Image denoising via deep residual convolutional neural networks,” *Signal, Image and Video Processing*, Aug 2019.

[5] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.

[6] J. Chung, Çaglar Gülçehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *ArXiv*, vol. abs/1412.3555, 2014.

[7] D. Martin, C. Fowlkes, D. Tal, J. Malik *et al.*, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics.” Iccv Vancouver:, 2001.

[8] E. Agustsson and R. Timofte, “Ntire 2017 challenge on single image super-resolution: Dataset and study,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 126–135.

A 11.75-bit Hybrid Sturdy MASH-21 Δ-Σ Modulator for Audio Applications

Ting-Hui Chang and Chia-Yu Yao

Department of Electrical Engineering, National Taiwan University of Science and Technology

ABSTRACT

In this paper, we present a hybrid sturdy-MASH-21 (SMASH-21) delta-sigma modulator (DSM) chip design. We employ differential opamps in the hybrid SMASH-21 DSM chip to reduce even-mode harmonics. The proposed circuit is realized in TSMC 0.18-μm CMOS technology. It consumes 2.58 mW of power. The sampling rate is 5.9976 MHz and the over-sampling ratio is 136. The achieved SNDR of the chip is 72.48 dB. The chip occupies an area of 1.68 mm^2.

1. INTRODUCTION

Many electronic systems require analog-to-digital converters (ADCs). Among various types of ADCs, delta-sigma modulator (DSM) ADCs are suitable for converting low-to-medium-bandwidth analog signals to digital signals with medium-to-high resolution. Let L denote the order of the DSM, Δ denote the step size of the quantizer, and OSR represent the over-sampling ratio that is defined as a half of the sampling rate divided by the signal bandwidth. Then the in-band quantization noise power is approximately $q_{rms}^2 = (\pi^{2L}\Delta^2)/[12(2L+1)\mathrm{OSR}^{2L+1}]$ [1]. Thus, increasing either the OSR or the order of the DSM can raise the SQNR. However, increasing the DSM order leads to a stability concern when a single-loop DSM structure is employed.

Hayashi *et al* [2] proposed a multi-stage noise-shaping (MASH) DSM structure to avoid handling the stability issue in a high-order single-loop DSM. An N-stage MASH DSM can eliminate the quantization noise produced in its first $N-1$ stages only when the DSM employs high accuracy integrators. In contrast, the sturdy MASH (SMASH) DSM has an advantage that its SQNR performance is less sensitive to the integrator impairments than that of its conventional MASH counterpart [3]. However, a two-stage SMASH structure cannot eliminate the quantization noise of its first stage. Thus, when the integrators used in a two-stage SMASH DSM possess sufficiently high accuracy, the SMASH DSM's SQNR performance is 3 dB worse than that of the conventional two-stage MASH structure.

We modify the SMASH structure by using both non-inverting integrator and inverting integrator and call it a hybrid SMASH (HSMASH) DSM. The two-stage HSMASH structure can eliminate the quantization noise in its first stage while retaining the insensitivity to the accuracy of the integrators used in the DSM. In an HSMASH-21 DSM chip design using TSMC 0.18-μm CMOS process, the measured SNDR reaches 72.48 dB, which approximates a 11.75-bit ENOB.

2. THE HSMASH-21 DSM ARCHITECTURE

Fig. 1 shows the proposed HSMASH-21 DSM structure, in which g_1, g_2, and g_3 are used to prevent all three integrators from overloading. The obtained output Y is derived as

$$Y = \frac{(H+1)zFH^2X + [zH-(H+1)]E_1 + E_2}{(zFH^2+H+1)(H+1)}. \quad (1)$$

When H is ideal, $zH = 1 + H$, and when the bandwidth of the LPF F is sufficiently large, we can obtain

$$Y \approx z^{-1}X + (1-z^{-1})^3E_2. \quad (2)$$

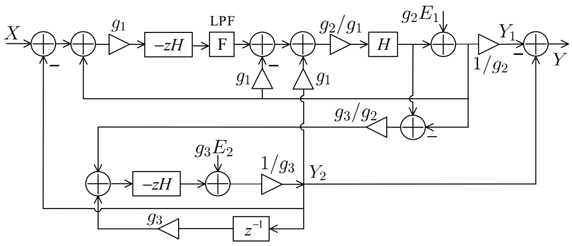


Fig. 1. Block diagram of the HSMASH-21 DSM.

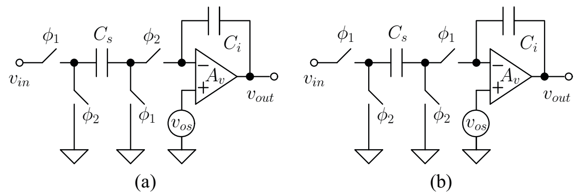


Fig. 2. (a) The non-inverting SC integrator. (b) The inverting SC integrator.

Comparing (2) with the SMASH DSM result given in [3], the first stage quantization noise (E_1) of the proposed HSMASH DSM is eliminated if the integrator realization is sufficiently accurate.

2.1. Switched-Capacitor Integrator Impairments

Fig.2(a) and Fig.2(b) show the switched-capacitor (SC) non-inverting integrator and the SC inverting integrator, respectively. A_v represents the finite gain of the opamp and v_{os} represents the offset voltage of the opamp. The z-domain input-output equations of the respective SC integrators are

$$V_{out,n} = V_{in}\frac{g\beta z^{-1}}{1-\alpha z^{-1}} + V_{os}\frac{g\beta}{1-\alpha z^{-1}}, \quad (3)$$

$$V_{out,i} = V_{in}\frac{-g\beta}{1-\alpha z^{-1}} + V_{os}\frac{g\beta}{1-\alpha z^{-1}}, \quad (4)$$

where $g = C_s/C_i$, $\alpha = (A_v+1)/(A_v+g+1)$, and $\beta = A_v/(A_v+g+1)$.

Using the simulation model presented in [4], we simulate both SMASH-21 and HSMASH-21 structures with various A_v and measure the output SNDR results. The target application of the modulator is to convert the audio signals. The simulation scenario is: The system bandwidth BW = 22.05 kHz, the OSR = 136, the input frequency $f_i = 2.5$ kHz, and the amplifier output saturates at ± 0.5 V. Both SMASH-21 and HSMASH-21 DSMs employ the same scaling factors $g_1 = 0.05, g_2 = 0.05$, and $g_3 = 0.025$. Fig. 3 shows the obtained SNDR vs. amplifier gain plot. When the amplifier gain is greater than 45 dB, the HSMASH-21 structure outperforms the SMASH-21 structure by 6 dB approximately, better than the predicted value by (2). The reason is that the same scaling factors have different effects on different DSMs. Apparently, their effects on the new HSMASH DSM are better.

3. HSMASH-21 CHIP DESIGN AND MEASUREMENTS

Fig. 4 shows the SC circuit schematic of the proposed HSMASH-21 DSM, in which the scaling factors in previous simulations

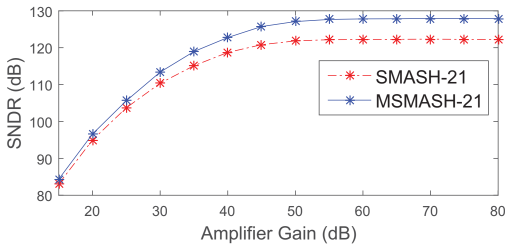


Fig. 3. The SNDR comparison for SMASH-21, and HSMASH-21 DSMs by simulations.

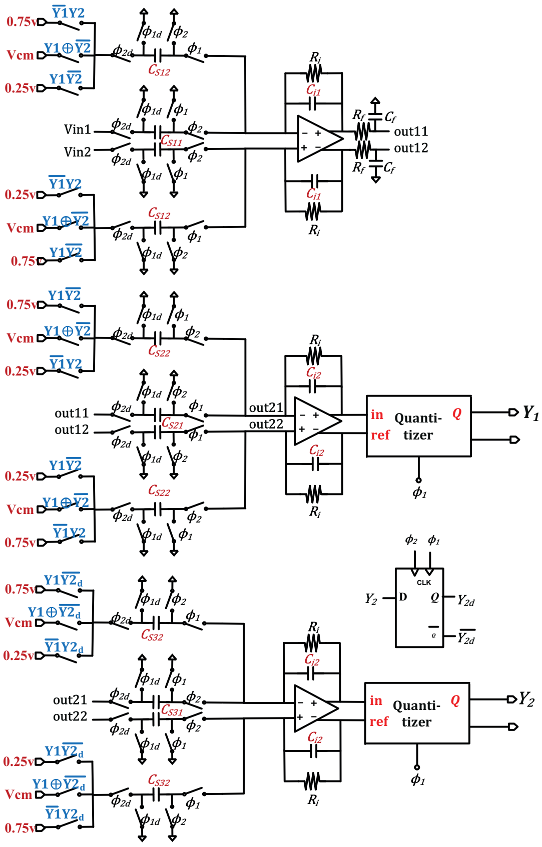
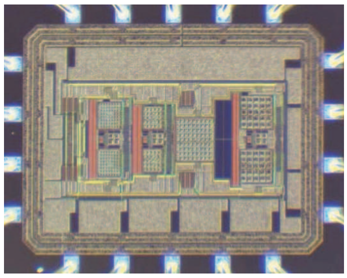


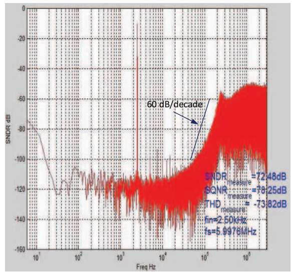
Fig. 4. SC circuit schematic of the HSMASH-21 DSM.

are employed. The supply voltage is 1 V and $V_{cm} = 0.5$ V. The HSMASH-21 DSM chip is realized in TSMC 0.18-μm CMOS technology. Fig. 5(a) shows the photo of the chip. The chip area is 1.68 mm^2. In measurements, the sampling frequency was set at 5.9976 MHz. The input sinusoidal signal frequency was first set at 2.5 kHz and the input amplitude was set at 0.42 V. We used the logic analyzer to read Y_1 and Y_2. We then calculated the spectrum of Y, from which we could estimate the output SNDR value. Fig. 5(b) shows the obtained spectrum. The output SNDR value is 72.48 dB. The effective number of bits (ENOB) is equal to 11.75 bits. Fig. 5(c) shows the plot of the output SNDR values when the input signal amplitude equals 0.1 V, 0.15 V, 0.2 V, 0.25 V, 0.3 V, 0.35 V, 0.4 V, 0.42 V, 0.45 V, and 0.5 V. The corresponding regression line is also shown in Fig. 5(c). We see that the DSM chip performance shows moderate linearity when the input signal amplitude is less than 0.42 V. The power consumption of the DSM chip is approximately 2.58 mW. Finally, we fixed the input amplitude at 0.42 V and scanned the input frequency from 500 Hz, 1 kHz, 2.5 kHz, to 5 kHz. Fig. 5(d) shows the corresponding measured SNDR values. We see that the measured peak-to-peak SNDR variation is approximately 1.14 dB.

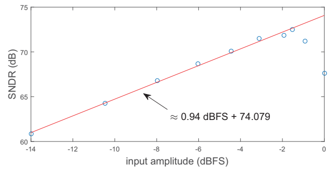
Table 1 shows the performance comparison with the bench-



(a)



(b)



(c)

Input frequency	SNDR (dB)
500 Hz	71.46
1 kHz	71.51
2.5 kHz	72.48
5 kHz	72.6

(d)

Fig. 5. (a) Photo of the HSMASH-21 DSM chip. (b) The obtained spectrum of Y when the input signal amplitude equals 0.42 V. (c) The obtained SNDR values vs. various input amplitudes. (d) The obtained SNDR values vs. various input frequencies.

Table 1. Performance comparison with the benchmark SMASH-22 DSM chip.

	CMOS tech.	Archi-tecture	Band-width	OSR	SNDR	Power	FOM
[5]	0.18 μm	SMASH-22	625 kHz	16	74 dB	3.2 mW	0.625
This work	0.18 μm	HSMASH-21	22.05 kHz	136	72.48 dB	2.58 mW	16.97

$$\mathrm{FOM} = \frac{\mathrm{Power}}{2^{\mathrm{ENOB}} \times 2 \times \mathrm{Bandwidth}} \mathrm{pJ/step}$$

mark SMASH-22 DSM chip [5]. The FOM of our chip is much higher than the FOM of the work in [5]. The main reasons are the power consumption and the total harmonic distortion (THD) of the two-stage opamp we employed in this work are too high.

4. CONCLUSION

In this paper, we proposed a novel HSMASH DSM architecture. When the amplifier gain is sufficiently high, the integrator is accurate enough and the proposed HSMASH DSM outperforms the SMASH one by 3 dB in SQNR. We designed a HSMASH-21 DSM chip using the TSMC 0.18-μm CMOS process. The measured SNDR value reached 72.48 dB. The DSM chip performance showed moderate linearity when the input signal amplitude varied. When the input signal frequency varied from 500 Hz to 5 kHz, the measured SNDR fluctuation was not greater than 1.14 dB.

5. REFERENCES

[1] R. Schreier and G. C. Temes, *Understandarding Delta-Sigma Data Converters*. Piscataway, NJ: IEEE Press, 2005.

[2] T. Hayashi, Y. Inabe, K. Uchimura, and A. Iwata, “A multistage delta-sigma modulator without double integration loop,” in *ISSCC Dig. Tech. Papers*, Feb. 1986, pp. 182–183.

[3] N. Maghari, S. Kwon, G. C. Temes, and U. Moon, “Sturdy MASH Δ-Σ modulator,” *Electron. Lett.*, vol. 42, no. 22, pp. 1269–1270, 26th Oct. 2006.

[4] C.-Y. Yao, Y.-H. Ho, W.-C. Hsia, and J.-J. Huang, “Simulating delta-sigma analog-to-digital converters with the opamp nonlinearity using the newton’s method,” in *IEEE Int. Symp. Circuits Syst.*, May 2015, pp. 537–540.

[5] N. Maghari, S. Kwon, and U.-K. Moon, “74 db sndr multiloop sturdy-mash delta-sigma modulator using 35 db openloop opamp gain,” *IEEE J. Solid-State Circuits*, vol. 44, no. 8, pp. 2212–2221, Aug. 2009.

Using Inversion-mode MOS Varactors and 3-port Inductor in 0.18-μm CMOS Voltage Controlled Oscillator

Ming-Xuan Li, Ching-Yi Jiang, Yu-Ying Pan, Hao-Hui Chen and Yao-Wen Hsu#

Department of Electronic Engineering, National Kaohsiung University of Science and Technology, No.2, Jhuoyue Rd., Nanzih Dist, Kaohsiung, Taiwan

hsuyw@nkust.edu.tw

***Abstract* —This paper presents a RF voltage controlled oscillator (VCO) using inversion-mode MOS varactors and 3-port inductors to achieve low power consumption, low phase noise, broad tuning range and minimized chip size. The proposed circuit architecture using body-biased technique operates from 4.3 to 5 GHz with 20.8% tuning range. The measured phase noise is less than -125.34 dBc at a displacement frequency of 1 MHz. The power consumption of this VCO is 25 mW when biased at 1.8 V. This VCO was implemented in standard TSMC 0.18-μm 1P6M process. The chip size is 0.476 mm^2 including the pads, which is only 63% comparing with an identical VCO using TSMC inductor model.**

***Index Terms*—CMOS, VCO, body-biased technique, 3-port inductor.**

I. Introduction

Voltage controlled oscillators (VCOs) are integrated as a core part of the phase-locked-loops (PLLs) or frequency synthesizers in microwave transceiver systems to provide signal sources for frequency conversion and carrier generation. The primary consideration of VCO designs include the tuning range, phase noise, power consumption and chip size. Recently, it has been shown that adopting PMOS in VCO core can achieve lower phase noise [1]. The use of MOS transistor varactors in RF VCO has been investigated in [2], which show a lower power consumption and a low phase noise than those by using diode varactors. In this paper, we adopt PMOS varactors as the part of the LC tank to decrease phase noise and refer to the architecture in [2] as our primary RF VCO design. Since the conventional inductor model has relatively large layout area, we will reconstruct the resonant inductor similar to those in [3] as a compact design instead of using the standard model. Another VCO chip using TSMC inductor model is also fabricated as a comparison.

II. Theoretical Fundamentals

The topology of our proposed VCO is illustrated in Fig. 1. This circuit is constructed by three parts: the transistors M1 and M2 as the differential pair serves as the core of the cross-coupled oscillator, LC tank using variable MOS capacitors (M3 to M10) and the buffer circuit. An MOS transistor with drain, source and bulk (D, S, B) connected together realizes a conventional MOS capacitor with capacitance value dependent on the voltage V_{BG} between bulk and gate. Instead of using the conventional MOS varactors, we apply the inversion-mode MOS (connect drain to source, apply bias voltage to bulk and connect gate to the inductor) which has a lower parasitic resistance than the pMOS varactors. Since the resonance frequency from the LC tank depends on the variable capacitance, we can achieve a wider tuning range.

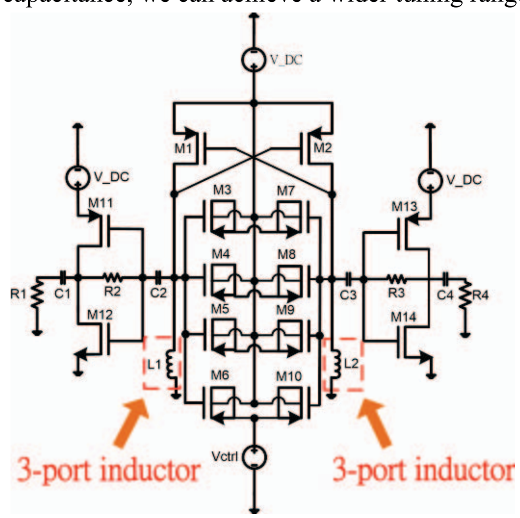


Fig. 1 The architecture of our proposed high-modulation VCO diagram using 3-port inductor

The RF inductors of our LC tank occupies the most area in VCOs on RFIC. To minimize the inductor area, we consider combining nearby inductors and applying the usage of the mutual inductors. The same inductance can be obtained by using the full-wave simulation software HFSS. This 3-port inductor using layers of Metal 5 (blue) and Metal 6 (red) has been shown in Fig. 2.

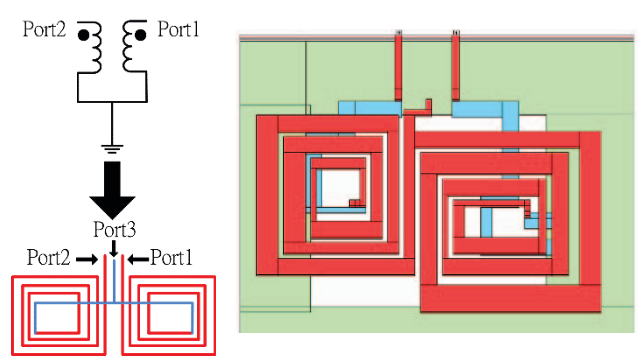
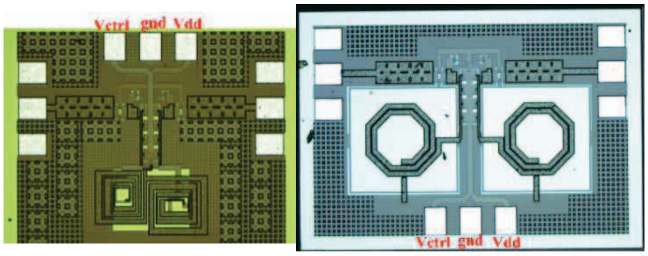


Fig.2 The implementation of our 3-port inductor in various layout layers using CMOS 0.18-µm process. (Port 3 is connecting to ground.)



(a) (b)

Fig. 3 Die photograph of our purposed VCO (a) using TSMC inductor model (643*850 μm^2) (b) using 3-port inductor (633*750 μm^2)

III. Fabrication and Measurements

The proposed high-modulation voltage controlled oscillators are implemented by CMOS 0.18-µm supplied by Taiwan Semiconductor Manufacturing Co., Ltd using TSMC inductor model (without 3-port inductor) and 3-port inductors. The die photo of the VCO using TSMC inductor model is presented in Fig. 3 (a). In addition, the die photo of our proposed VCO using 3-port inductors is shown in Fig. 3 (b). Both chips include DC and RF pads with the area of 0.547 and 0.475 mm^2 respectively.

Fig. 4 shows the simulated and measured oscillation frequency of the proposed voltage controlled oscillator. It is observed that the maximum output power at the design frequency is -2.9 dBm, and the overall trend is similar to the simulation.

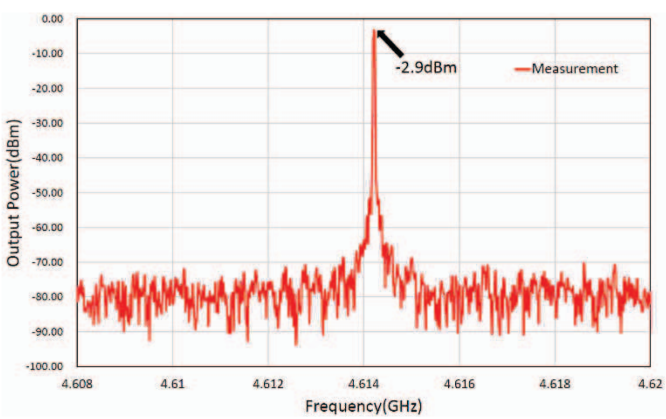


Fig. 4 The measured oscillation frequency from our proposed VCO using 3-port inductors.

Fig. 5 shows the comparison of the simulated and measured phase noise. The measured phase noise is less than -125 dBc at a displacement frequency of 1 MHz. Fig. 6 shows the comparison of the simulated and measured tuning range. The tuning range of the measurement is 4.3 GHz to 5.3 GHz up to 20.8%.

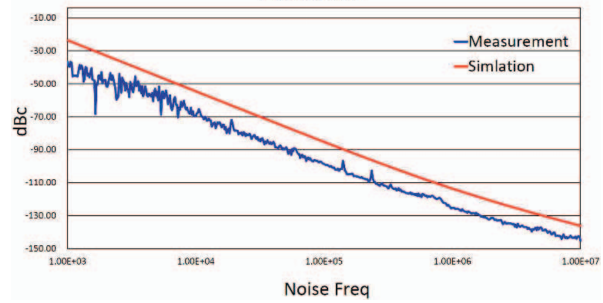


Fig. 5 Comparison of the simulated and measured phase noise from our proposed VCO using 3-port inductors.

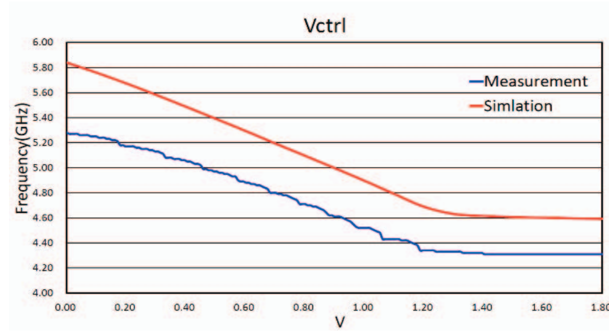


Fig. 6 Comparison of the simulated and measured tuning range from our proposed VCO using 3-port inductors.

IV. Conclusion

A RF voltage controlled oscillator (VCO) using inversion-mode MOS varactors and 3-port inductors is presented in this paper. The inversion-mode MOS varactors provide 20.8% tuning range from 4.3 to 5.3 GHz. Two VCOs are fabricated in CMOS 0.18-µm process for the same design architecture using TSMC inductor model or 3-port inductors for comparison. The chip size of VCO using 3-port inductors is 0.476 mm^2 including the pads, which is only 63% compact chip area comparing with another design using TSMC inductor model while maintaining comparable performance.

References

[1] M.-D. Tsai, Y.-H. Cho, and H. Wang, "A 5-GHz low phase noise differential colpitts CMOS VCO," IEEE Microwave and Wireless Components Letters, vol. 15, no. 5, 2005, pp. 327-329.

[2] P. Andreani and S. Mattisson, "On the use of MOS varactors in RF VCOs," IEEE Journal of Solid-State Circuits, vol. 35, no. 6, 2000, pp. 905-910.

[3] C.-Y. Yen, Y.-W. Hsu, H.-H. Chen, and Y.-Y. Pan, "A 0.1-28 GHz Differential Cascaded Distributed Amplifier in 0.18-µm CMOS Technology," in 2019 IEEE MTT-S International Wireless Symposium (IWS), 2019, pp. 1-3.

[4] H.-H. Lai, I.-S. Shen, and C. F. Jou, "Colpitts current-reused QVCO based on capacitor coupling," in Asia-Pacific Microwave Conference 2011, 2011, pp. 1638-1641.

[5] Y.-C. Tsai, Y.-S. Shen, and C. F. Jou, "A Low-Power Quadrature VCO Using Current-Reused Technique and Back-Gate Coupling," Piers Online, vol. 3, 2007, pp. 971-975.

[6] H.-M. Chen, Y.-D. Jhuang, J.-C. Lin, and S.-L. Jang, "A 5.6 GHz Balanced Colplitts QVCO With Back-gate Coupling Technique," in 2007 IEEE Conference on Electron Devices and Solid-State Circuits, 2007, pp. 965-967.

Capacitance Minimization Clock Synthesis with Blockage-Avoiding Hybrid-Structure Network

Chun-Wei Ho and Shao-Yun Fang
Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan
e-mail: syfang@mail.ntust.edu.tw

***Abstract*—Clock network design has become a great challenge since it accounts for a huge portion of chip power budget and and plays a crucial role determining circuit delay. Among different methods, clock mesh provides high robustness to process, voltage, and temperature (PVT) variations due to redundant paths. However, the mesh structure suffers from high power dissipation. By contrast, conventional clock tree structure is commonly used due to low power consumption, less routing resource usage. Nevertheless, a tree-based network is highly sensitive to PVT variations. In this paper, we propose to use the hybrid structure that combines tree-based and mesh-based structures for power and skew trade-off methodology. Experimental results suggest that hybrid-structured clock network can minimize the total capacitance under skew constrains.**

I. Introduction

As the technology process continues to shrink below deep nanometer nodes, clock networks synthesis faces several design issues such as manufacturing variation, power supply noise, and temperature. All these impacts on clock skew become more significant. Higher skew directly reduces the clocking frequency and will further influence the timing yield. In order to handle the clock skew to make sure the timing signals are under the frequency constraint, more clock resource and thus higher total capacitance is required, leading to higher capacitance usage and thus great power dissipation.

Currently, there are two categories of clock network synthesis techniques widely available: (1) tree-based clock synthesis [2], [3], [4], [5], [7], [8], and (2) mesh-based clock synthesis [6], [9]. A conventional clock tree is characterized by a tree-based topological structure from the clock root to all sinks within the chip boundary. Most of the sinks in the design share very few paths back to the clock source. The less use of routing resource usage, the less power consumption, and thus clock tree is commonly used as its simplicity of implementation and simulation. However, tree-based architecture is extremely sensitive to the process, voltage and temperature (PVT) variation, especially in high-performance chip design. In contrast, mesh-based clock network is another clock synthesis structure. Instead of a few sharing path of clock tree, mesh-based clock network fabricates a mesh structure as the sharing path among the clock sinks from the clock source, and a top-level tree drives the mesh structure. Therefore, clock mesh architecture provides robust tolerance to PVT variations. However, mesh architecture with lots of the sharing paths, as a result, will degrade the chip performance due to much higher total wire capacitance. Furthermore, with a large quantity of mesh nodes and unbalanced sink loading, clock mesh analysis and automation is difficult.

In this paper, we propose a clock network synthesis flow with a hybrid structure that combines tree-based and mesh-based structures for simultaneous power and skew optimization, as illustrated in Fig. 1. Loading balance is considered to optimize sink loadings among the lattices of the constructed mesh. Experimental results suggest that hybrid-structured clock network can minimize the total capacitance under skew constrains.

II. The Proposed Clock Tree Synthesis Flow

A. Algorithm Flow

Fig. 2 shows the clock synthesis flow for the hybrid-structured clock network. Given a set of sinks, we temporally determine the initial pitch size of the clock mesh with local skew distance, within which the clock skew of a pair of local sinks needs to be considered [10]. Then, if the maximum sink loading in a lattice is too large to be driven by the largest-sized buffer, we shrink the mesh size to meet the drivability of the buffer. A loading balance process is adopted to let the sink loading in each lattice more balanced. After that, mesh reduction is performed to eliminate empty lattices and lattices covered by obstacles. Having the constructed mesh, we build the local tree for every mesh lattice based on the DME algorithm [1], and the roots of local tree are regarded as the sinks of the top-level tree. As the top-level tree have been built, we handle the obstacle avoiding issue by adopting routing algorithm. In order to meet the slew-rate constraint, buffers are inserted in the top-level tree in the final stage. Due to the page limit, we detail the key steps highlighted in Fig. 2 in the following subsections.

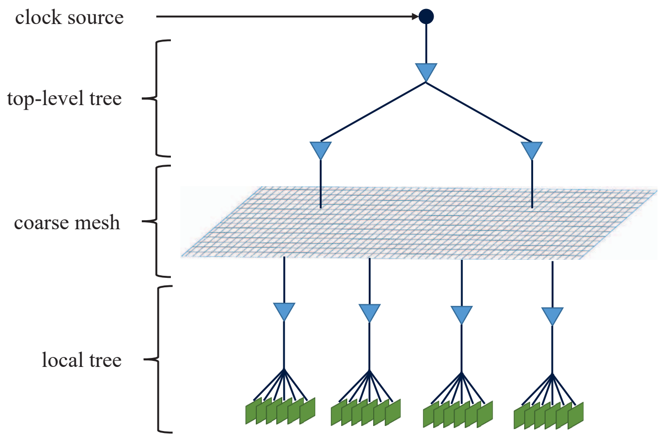


Fig. 1. An example of a hybrid structure clock network

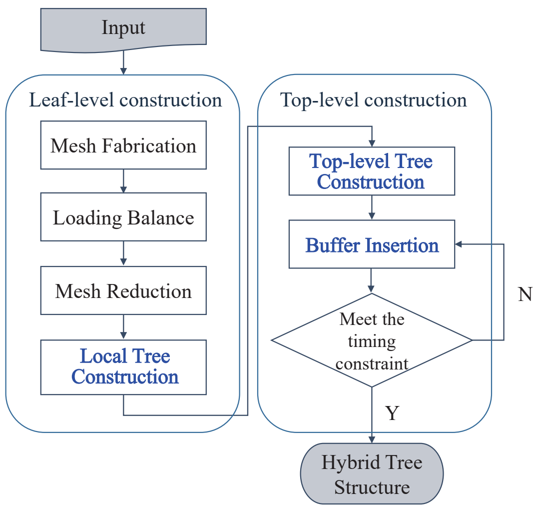


Fig. 2. The proposed algorithm flow.

B. Local Tree Construction

For the local tree construction, we first generate the tree topology. We use the concept of the Balanced Bipartition (BB) method which

This work was partially supported by Synopsys, TSMC, and MOST of Taiwan under Grant No. MOST 108-2636-E-011-002.

TABLE I

Benchmark	[9]			[6]			Ours		
	Skew	Cap	CPU	Skew	Cap	CPU	Skew	Cap	CPU
ispd10-01	7.23	1168104	675	6.27	453183	0.84	7.39	422261	0.51
ispd10-02	7.35	2099811	2140	7.36	867581	3.83	7.44	768352	2.23
ispd10-03	3.95	93965	21	4.06	94846	0.20	4.11	89828	0.42
ispd10-04	7.25	125333	22	6.75	99179	0.18	7.32	109316	0.64
ispd10-05	7.27	74084	10	5.19	74973	0.07	7.16	50615	0.23
ispd10-06	6.79	87390	46	6.76	98924	2.76	7.21	42678	0.24
ispd10-07	5.97	128351	27	5.96	138757	3.31	6.71	143723	1.2
ispd10-08	5.37	97421	18	4.87	110876	3.14	6.28	113103	1.26
Comp.	0.95	1.61	329.93	0.88	1.24	2.64	1.00	1.00	1.00

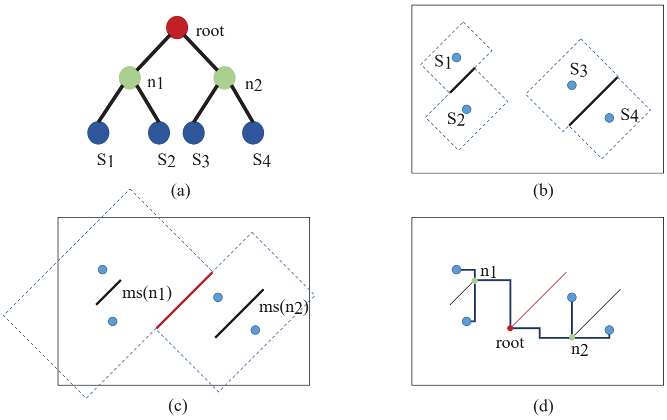


Fig. 3. The process of the DME algorithm.

is first presented [1]. The BB method is a top-down method, dividing the sink nodes recursively into two partitions with nearly equal total loading capacitance. After the tree topology is constructed, we adopt the deferred-merge embedding (DME) algorithm [1] to handle the clock tree routing. The DME algorithm adopts a two-phase process. A bottom-up phase constructs the candidate segment for the top-down phase. Once the tree of segments has been constructed, the exact embedding of internal nodes in the tree are chosen in top-down manner.

Fig. 3 illustrates the overall DME flow. In Fig. 3(a), the given topology which is generated from the BB method predefines the connection order of sinks, where s_1, s_2, s_3 and s_4 are sinks to be merged, n_1, n_2 are the nodes to be embedded, and $root$ is the clock source. In the Fig. 3(b), the bottom-up phase of DME starts with all sink locations. Each sink location is taken as a zero Manhattan arc. If two sinks have the same parent, a merging segments denoted as $ms(n_1)/ms(n_2)$ and represented as the solid line in Fig. 3(b), is determined. According to the bottom-up merging order suggested in Fig. 3(a), the merging segment $ms(root)$ is determined by the merging segment of $ms(n_1)$ and $ms(n_2)$, as shown in Fig. 3(c). After all the merging segments of nodes and sinks are determined in the bottom-up phase, the exact location of the internal nodes are embedded within these merging segments in the DME top-down phase, started with the root. The embedding process determines the tapping point locations by selecting tge nearest points to achieve the shortest wirelength tree, as shown in Fig. 3(d).

C. Top Tree Construction

At this stage, the local tree root is directly connected to the nearest mesh spine and it will be treated as the lowest-level node for the top-level tree construction. The capacitance of each node is the sum of the loading capacitances, which is computed by accumulating the sink loadings and routing wires in a local tree. An extended-DME algorithm builds up the merging segments in the bottom-up phase, and a top-down embedding phase determines the exact locations of all nodes and handle the obstacle-avoiding routing in the final stage by using the Dijkstras shortest-path algorithm.

D. Buffer Insertion

In order to meet the slew rate constraint, buffers are often needed to insert in the routing path. The slew rate is approximated by accumulating the capacitance from the latest inserted buffer. We adopt a top-down manner, from the clock source to sinks, to insert the buffers. Along a tree edge, once the accumulated capacitance considering wire capacitance and the upstream capacitance is too large to be driven by the largest buffer, we add a buffer on the location.

III. Experiment

Our hybrid-structured clock network was implemented in C++ language, and experiments were conducted on a Linux machine with 2.10GHz CPU and 58GB memory. The ISPD 2010 Clock Network Synthesis Contest benchmarks [10] and 45-nm technology nodes are used. Monte Carlo simulation provided in the evaluation script is performed to simulate the process variation.

Table I shows the experimental results on clock skew (Skew), capacitance usage (Cap), and the running time (CPU). The skew limitation is 7.5 ps for each circuit except for the third circuit whose skew limitation is 4.9 ps. Compared with [9] that uses mesh structure, the result of [9] has 5% lower clock skew, but it has 61% higher capacitance usage. Compared with [6] that also uses mesh structure, our clock skew is 12% larger since the tree structures in the lowest-level suffer from lower variation tolerance than pure mesh structures. However, mesh structures use more resource than tree structures, and thus the total capacitance of our clock network is reduced by 61% and 24% cimpared with [9] and [6]. On the other hand, our runtime is 329.93X and 2.64X faster over [9] and [6] approaches respectively. The results show that the proposed flow can minimize the total capacitance under skew constraints.

IV. Conclusion

In this thesis, in order to minimize the total capacitance in clock network, we use the hybrid-structured clock network on capacitance optimization. Experimental results have shown that hybrid-structured clock network have the superiority on capacitance minimization over pure mesh-structured clock networks.

References

[1] T. H. Chao, Y. C. Hsu, and J. M. Ho, “Zero skew clock routing with minimum wirelength,” *IEEE TCAS-II*, vol. 39, no. 11, pp. 799–813, 1992.
[2] J. Cong, A. Kahng, and G. Robins, “Matching based model for high performance clock routing,” *IEEE TCAD*, vol. 12, no. 8, pp. 1157-1169, 1993.
[3] S. Dhar, M. Franklin, and D. Wann, “Reduction of clock delay in VLSI structure,” in *Proc. ICCAD*, 1984.
[4] M. A. B. Jackson, A. Srinivasan, and E. S. Kuh “Clock routing for high performance ICs,” in *Proc. DAC*, 1990.
[5] A. Kahng, J. Cong, and G. Robins, “High-performance clock routing based on recursive geometric matching,” in *Proc. DAC*, 1991.
[6] X. W. Shih, H. C. Lee, K. H. Ho, and Y.W . Chang, “High variation tolerant obstacle avoiding clock mesh synthesis with symmetrical driving trees,” in *Proc. ICCAD*, 2010.
[7] R. S. Tsay, “Exact zero skew,” in *Proc. ICCAD*, 1991.
[8] A. Vittal and M. Marek-Sadowska, “Low-Power buffered block tree design,” in *Proc. ICCAD*, 1997.
[9] L. Xiao, Z. Xiao, and Z. Qian, “Local clock skew minimization using blockage-aware mixed tree-mesh clock network, in *Proc. ICCAD*, 2010.
[10] ISPD 2010 high performance clock network synthesis contest. Available: http://www.sigda.org/ispd/contests/10/ispd10cns.html

Latency Constraint-aware Scheduler for NVMe Solid State Drives

Cheng-Yu Li, Jih-Hsiang Cheng, and Ya-Shu Chen
Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan
{m10607417, m10707413, yschen}@mail.ntust.edu.tw

***Abstract*—Non-Volatile Memory Express (NVMe) solid-state drives (SSDs) provide high data transfer rates by the benefited from the multi-queue interface and high-speed bus. With a better tradeoff between power and performance, the NVMe SSD is becoming ideal storage for embedded systems. Although multiple dispatch queues supporting increased the data rates, the interference between applications affects the response time of applications which have latency constraints in the embedded systems. In this study, we propose a latency constraint-aware scheduler to manage the response time of applications while maintaining the throughput of the NVMe SSD.**

***Index Terms*—Non-Volatile Memory Express, Solid State Drives, I/O scheduling, Storage Management, Storage Performance**

I. Introduction

Non-Volatile Memory Express based solid-state drives (NVMe SSDs) have been applied the multi-queue interface between block derivers and high-speed PCIe-based bus SSDs. The new block device driver supports multiple application-level I/O request queues, and each queue is accessible to the host system and the SSD controller [1]. The default scheduler for multiple dispatch queues is round-robin fashion to maximize the total throughput. However, multiple dispatch queues supporting introduces a significant resource contention problem for multiple I/O requests accessing and results in significant response delayed for applications with latency constraints. Latency may be a critical consideration for certain applications in embedded systems. Different applications would have varying data rates and latency requirements. This raises a challenge to manage the response time of applications while maintaining the throughput of the NVMe SSD.

With the increased need for SSDs, researches explored I/O scheduling issues for SSDs to improve performance. Authors in [2] proposed a dispatching algorithm to solve the internal contention of SSDs. To reduce the access conflicts between I/O requests, authors in [3] introduced access conflict detection to the host I/O scheduler, and authors in [4] proposed to check possible conflicts of fetched requests before dispatching requests. Considering managing application response, authors in [5] proposed bank reservations to satisfy the service level of virtual machines on NVMe SSDs, and authors in [6] proposed interference aware scheduler to provide fairness of multiple applications. Different from previous studies, this study focused on the tradeoff between response and throughput for applications with latency constraints in the NVMe SSD.

This work was supported in part by MOST Program under Grant MOST 107-2221-E-011-028-MY3.

II. System model

This study explores scheduling issues on NVMe solid-state drives (SSDs). The system consists of a set of different workload applications. Each application APP_i has multiple read/write requests, and each request is denoted by operation type, logical sector number (LSN), and data size [7].

The considered NVMe SSD supports multiple dispatch queues, which can execute multiple I/O requests at the same time when the requests access different physical addresses. When multiple requests can be executed at the same time, the interference between requests increases the response of applications. With the increased bandwidth of the SSD, the SSD has been recently applied to embedded systems. Most applications executed in embedded systems have latency constraints. The latency constraints of applications might be violated when the traditional round-robin schedule [8] is applied. As shown in Fig. 1, by applying a round-robin scheduler, the request of $(R, 9, 1)$ from APP_1 is delayed by the request of $(W, 7, 3)$ from APP_2 on the same channel, and then APP_1 misses its latency constraint. This research is motivated by the above observation to explore an I/O scheduler to minimize the response time of latency constraint applications while maintaining the throughput of the system.

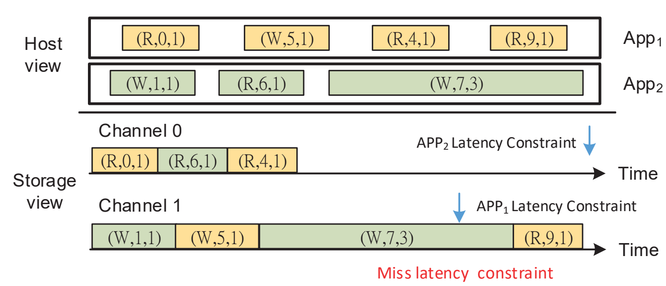


Fig. 1: Interference Between Applications

III. Approach

To manage the response time for applications with latency constraints, in this study, we proposed a latency constraint-aware scheduler (LCS). The idea is to set a higher priority to the application with the shorter latency constraint. However, to maintain the system throughput, instead of applying the earliest deadline first scheduler, the weighted round-robin is chosen to implement the latency constraint-aware scheduler. The higher priority application can get higher weight, and the weight among all applications is based on the ratio of the inversion of the latency constraint.

For example, when the latency of applications are respectively 100 ms and 200 ms, and then the weight of these two applications are 2 and 1. After determining the weight of applications, we schedule requests of the corresponding application based on the given weight under the weighted round-robin scheduler, and then serve the application with the higher priority first. As shown in Fig. 2, after applying LCS, we first serve the requests belong to APP_1 twice (i.e., $(R, 0, 1)$ and $(W, 5, 1)$), and then serve the requests belong to APP_2 once (i.e., $(W, 1, 1)$), and then repeat the procedures. Although the response of APP_2 is delayed by one read request, the response of APP_1 can be significantly shortened (e.g., compared to Fig. 1) and meets the latency constraint.

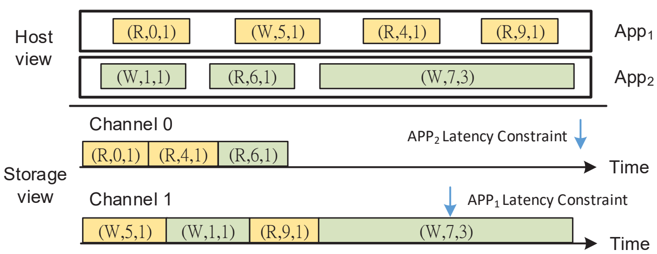


Fig. 2: Latency Constraint-aware Scheduler

IV. Performance Evaluation

An event-driven and trace-based driven simulator, SSDSim [9], is modified to support the multi-queue interface and is used to evaluate the proposed approach. A 128 GB SSD with 16 channels and each channel employed 4 chips, and page size is 4KB is considered in the experiment. The evaluated workloads are Systor1, Systor2, DAPDS, Financial2, and FIO. Systor1, Systor2, and DAPDS are collected from the SNIA IOTTA trace repository [10], and Financial2 is collected from OLTP applications running at financial institutions [11]. The synthetic workload FIO is generated by Flexible I/O Tester [12] for evaluating the random access pattern. We compared our proposed LCS with $Baseline$ which schedules requests based on round-robin.

We first consider read-intensive applications, including FIO, Financial2, and DAPDS, where the write ratio in these traces is between 10% and 30%. When FIO is defined as the shortest latency constraint in this experiment, the response time of FIO under LCS is 35% earlier than that of $Baseline$, while the response time of DAPDS with the longest latency constraint under LCS is only 14% later than that of $Baseline$. We then consider read-write applications, including FIO, Systor2, and Systor1, where the write ratio is between 30% and 60%. When the FIO is defined as the shortest latency constraint in the read-write application set, the response time of FIO under LCS is 59% earlier than that of $Baseline$, while the response time of Systor1 with the longest latency constraint under LCS is only 15% later than that of $Baseline$ when the garbage collection is not considered. The garbage collection might introduce a significant delay to shorter latency applications, so we will continuously explore this issue in future work.

V. Conclusion

Executing multiple applications in the NVMe SSD introduces longer interference which might significantly affect the quality of service of applications. We propose a latency constraint-aware scheduler to manage the response time of applications. The evaluation results revealed that the response of the applications with short latency constraints can be reduced by our proposed algorithm. For future research, we shall further explore the parallelism maximization by managing garbage collections and physical resource access conflicts.

References

[1] M. Bjørling, J. Axboe, D. Nellans, and P. Bonnet, "Linux block io: introducing multi-queue ssd access on multi-core systems," in *Proceedings of the International Systems and Storage Conference*. ACM, 2013, p. 22.
[2] M. H. Jo and W. W. Ro, "Dynamic load balancing of dispatch scheduling for solid state disks," *IEEE Transactions on Computers*, vol. 66, no. 6, pp. 1034–1047, 2016.
[3] C. Gao, L. Shi, M. Zhao, C. J. Xue, K. Wu, and E. H.-M. Sha, "Exploiting parallelism in i/o scheduling for access conflict minimization in flash-based solid state drives," in *Proceedings of the Symposium on Mass Storage Systems and Technologies (MSST)*. IEEE, 2014, pp. 1–11.
[4] T. Yang, P. Huang, W. Zhang, H. Wu, and L. Lin, "Cars: A multi-layer conflict-aware request scheduler for nvme ssds," in *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2019, pp. 1293–1296.
[5] S.-M. Huang and L.-P. Chang, "Providing slo compliance on nvme ssds through parallelism reservation," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 23, no. 3, pp. 28:1–28:26, 2018.
[6] A. Tavakkol, M. Sadrosadati, S. Ghose, J. Kim, Y. Luo, Y. Wang, N. M. Ghiasi, L. Orosa, J. Gómez-Luna, and O. Mutlu, "Flin: Enabling fairness and enhancing performance in modern nvme solid state drives," in *Proceedings of the ACM/IEEE Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2018, pp. 397–410.
[7] C. Gao, L. Shi, C. Ji, Y. Di, K. Wu, C. J. Xue, and E. H.-M. Sha, "Exploiting parallelism for access conflict minimization in flash-based solid state drives," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 1, pp. 168–181, 2017.
[8] "NVM Express," https://nvmexpress.org.
[9] Y. Hu, H. Jiang, D. Feng, L. Tian, H. Luo, and S. Zhang, "Performance impact and interplay of ssd parallelism through advanced commands, allocation strategy and data granularity," in *Proceedings of the International Conference on Supercomputing*. ACM, 2011, pp. 96–107.
[10] "Microsoft production server traces," http://iotta.snia.org/traces/list/BlockIO.
[11] "Umass trace repository," http://traces.cs.umass.edu/index.php/Storage/Storage.
[12] "Flexible i/o tester rev. 3.15," https://fio.readthedocs.io.

MMW Receiver Front-End for Noninvasive Glucose Measurement

Ming-Yu Yen, Fang-Yu Zhou, Wen-Ling Chang, and Hsiao-Chin Chen, *Member*, IEEE
Department of Electrical Engineering
National Taiwan University of Science and Technology
Taipei 106, Taiwan
M10107410@mail.ntust.edu.tw

***Abstract*—A millimeter wave (MMW) receiver front-end is designed and implemented using 90-nm CMOS technology for noninvasive glucose measurement. The LNA achieves the power gain of 19.43 dB at 29 GHz and the minimum noise figure of 6 dB at 29 GHz. The mixer performs frequency translation to convert the signals at 28-30 GHz to 10 MHz. The receiver front-end achieves the conversion gain of 14 dB and the P1dB of -23 dBm. The VCO delivers the frequency band of 28-30 GHz while achieving the FOM of -182.3 dBc/Hz.**

***Keywords*—Low noise amplifier, micromixer, non-invasive glucose sensing system, voltage control oscillator.**

I. Introduction

Recently, non-invasive glucose measurement has drawn more and more attention because the procedure of blood sample acquisition in conventional glucose measurement is painful and inconvenient [1].In a previous research, a strong correlation between the blood glucose levels and MMW absorptions has been shown using the ear of an anesthetized rat as the MMW transmission medium [2]. According to the experiment results, the MMW absorptions of rat ear under different glucose levels show the most difference over 28-30 GHz. In this work, 28-30 GHz MMW circuits such as the receiver front-end and VCO are implemented for a non-invasive glucose measurement system that is based on MMW absorption spectroscopy.

II. Circuits Design

The block diagram of the glucose measurement system is shown in Fig. 1. The glucose measurement system is formed from a transmitter and a receiver, along with antennas embedded in a special apparatus. The transmitter consists of a power amplifier (PA) and a voltage controlled oscillator (VCO), while the receiver consists of a low noise amplifier (LNA), a mixer, a VCO, a programmable gain amplifier (PGA), a low pass filter (LPF), a peak detector and a successive-approximate-register analog-to-digital converter (SAR-ADC).This work is focused on the design of the receiver front-end and VCO.

The schematic of the LNA is shown in Fig. 2. A three stage cascaded amplifier is adopted where the first stage is designed for low noise while the second stage and third stage for high gain [3]. Inductive degeneration is used in the first stage of the LNA to achieve input impedance and noise matching concurrently. The micro-mixer that follows the LNA exhibits an input impedance close to 50 ohm. Therefore, an output impedance matching network in the third stage is used to deliver an output impedance of 50 ohm to achieve the impedance match between the LNA and mixer.

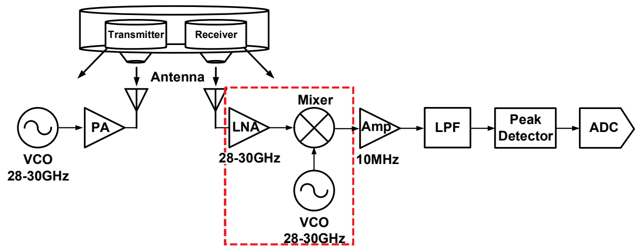


Fig. 1. Block diagram of the non-invasive glucose measurement system.

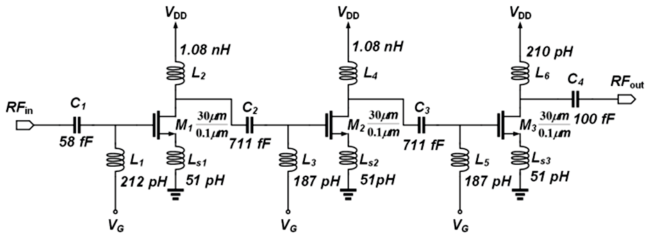


Fig. 2. Schematic of the three-stage LNA.

The schematic of the down-conversion micromixer is shown in Fig. 3. Notably, the impedance observed from the two branches at the input node can be approximated to $1/g_{m3}$ and $1/g_{m1}$. The admittance of Y_{CG} would decrease as the input voltage increases while the admittance of Y_{diode} would increase as the voltage increases. Therefore, the device selection and bias condition can be carefully performed for transistors M_1 and M_3 to minimize the variation in the input impedance at different input voltages to achieve good linearity.

The VCO is adopted to provide the local oscillator for the mixer to down convert the MMW input to 10MHz. The schematic of the VCO that uses the feedback LC tank, NMOS cross coupled pair and LC filter is shown in Fig. 4. A 2-bit coarse tuning capacitor array and a 3-bit fine tuning capacitor array are employed to cover the desired frequency range. Three varactors that uses different bias voltages are connected in parallel to minimize the variation in Kvco. The effective transconductance can be characterized to optimize the phase noise in this VCO topology so that the wide tuning range and low phase noise can be simultaneously achieved.

III. Mearsurement Results

All the circuits are fabricated in 90-nm CMOS process. The LNA consumes the power of 26.3 mW from the 1.2V supply. The measured power gain, input/output return losses and isolation of the LNA are shown in Fig. 5(a). The measured gain is 19.43 dB at 29 GHz with 3dB bandwidth of 2 GHz from 28-30 GHz.

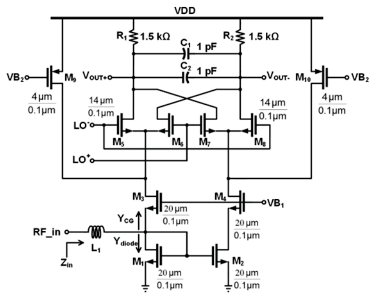


Fig. 3. Schematic of the down-conversion micromixer.

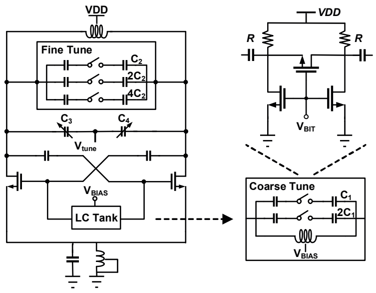


Fig. 4. Schematic of the VCO.

At 29 GHz, the input return loss is below -10 dB. The measured noise figure is presented in Fig. 5(b), where it reaches its minimum of 6 dB at 29 GHz. The measured P_{1dB} and IIP3 at 28 GHz are -18 dBm and -8.3 dBm, respectively. The receiver front-end consumes the power of 27.13 mW from the 1.2 V supply. The measurement results of the receiver front-end are shown in Fig. 6.The circuit delivers the voltage conversion gain of 14 dB at 29 GHz and P1dB of -23 dBm. During the measurement, the mixer is given with the LO power of 0 dBm for the maximum voltage conversion gain. The VCO consumes the power of 14.8 mW from the 1.2 V supply. Fig. 7 shows the measured VCO phase noises at the minimum and the maximum frequencies. The tuning range of the VCO is from 27.5 to 30.05 GHz. As shown in Fig. 7, the VCO achieves the phase noises of -125.2 and -118.1 dBc/Hz at 10-MHz offset, while operating at 27.5 and 30.05 GHz, respectively. The corresponding FOM is -182.3 dBc/Hz. For the noninvasive glucose measurement, the operating frequency of all circuits needs to cover 28~30 GHz. In Fig. 5(a), it can be observed that the I/O return losses of the LNA are greater than 10 dB at 28-30 GHz. The 28~30 GHz can also be covered by the tuning range of the VCO. So the performance of the receiver front-end circuits is sufficient. The chip photographs are shown in Fig. 8. The receiver front-end occupies the chip area of 1.25×0.9 mm^2 while the VCO occupies the chip area of 0.6×0.74 mm^2

IV. Conclusion

A 28-30 GHz receiver front-end for noninvasive glucose measurement is designed and implemented using 90-nm CMOS process. The receiver front-end achieves the power gain of 19.43 dB at 29 GHz and the minimum noise figure of 6 dB at 29 GHz. The VCO achieves the FOM of -182.3 dBc/Hz.

Acknowledgment

This work was supported by the Ministry of Science and Technology, Taiwan, under the contract number 108-2221-E-011-155-. We also appreciate National Chip Implementation Center of Taiwan for the technical support during the chip fabrication.

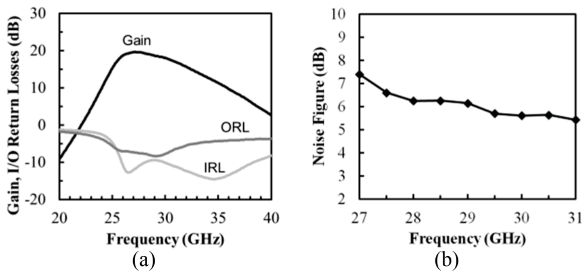


Fig. 5. Measurement results of LNA: (a) gain and I/O return losses (b) noise figure.

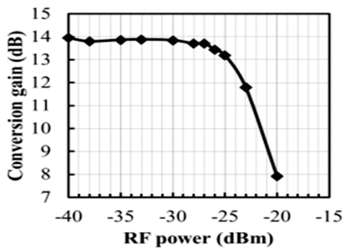


Fig. 6. Measured conversion gain of the receiver front-end

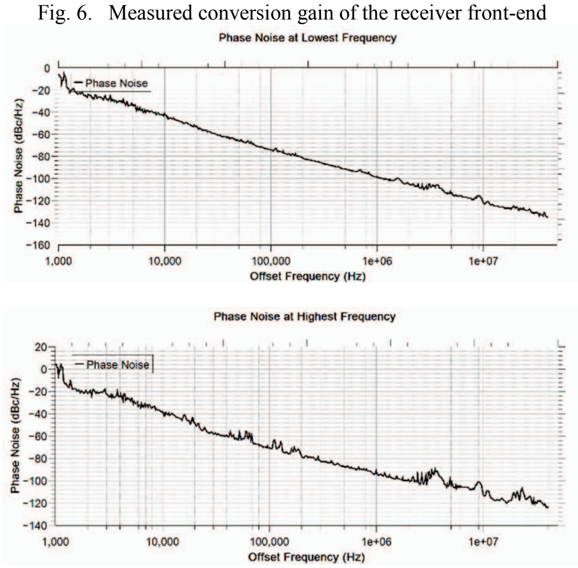
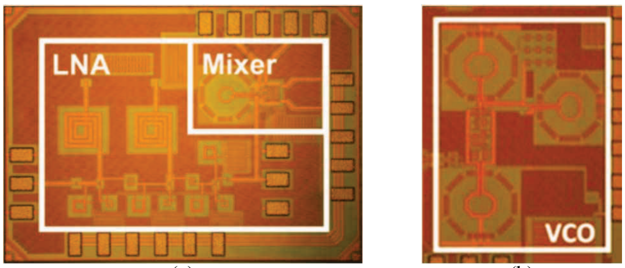


Fig. 7. Measuredphase noise of the VCO.



(a) (b)

Fig. 8. Chip photograph of: (a) The receiver front-end (b) The VCO.

References

[1] P. H. Siegel, A. Tang, G. Virbila, Y. Kim, M. C. F. Chang and V. Pikov, "Compact non-invasive millimeter-wave glucose sensor," *2015 40th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz)*, Hong Kong, 2015, pp. 1-3

[2] Alexeeva, N. V. & Arnold, M. A. Impact of tissue heterogeneity on noninvasive near-infrared glucose measurements in interstitial fluid of rat skin. *Journal of diabetes science and technology 4*, 1041–1054 (2010).

[3] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5-GHz CMOS low noise amplifier," in *IEEE Journal of Solid-State Circuits*, vol. 32, no. 5, pp. 745-759, May 1997.

2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)

Rehabilitation Seat Cushion System

Ming-Ta Ke
Bachelor Program in Interdisciplinary Studies
National Yunlin University of Science and Technology
Yunlin, Taiwan
kemingta@yuntech.edu.tw

Po-Cheng Su, Ya-Hsin Hsueh*, Yen-Chin Lin, Hsiang-Lung Huang, Yu-Jhang Wu, Jyun-Jhe Chen, Yi-Ting Zhong
Department of Electronic Engineering
National Yunlin University of Science and Technology
Yunlin, Taiwan
*hsuehyh@yuntech.edu.tw

***Abstract*—A rehabilitation seat cushion system is developed for wheelchair user in this work. The system includes a sitting posture tilt reminder, a slide warning and a real-time feedback rehabilitation training software. The sensor of the system can be battery-powered. This system can detect whether there is a tilt in the sitting position of the wheelchair user and also can give an immediate warning notice when the user slips off the wheelchair. When the user is leaning to left or right for more than a certain period of time, the system will send a warning signal to remind the user to adjust the sitting posture. In addition, a software is provided for users to conduct related training, where the training intensity can be increased to achieve better effects of rehabilitation.**

Keywords—Rehabilitation , Cushion, Sensor

I. INTRODUCTION

In Taiwan, the elderly has accounted for 14 % of the total population, which indicates that Taiwan has officially entered the stage of an aged society [1]. Along with population aging, the rise in average life expectancy leads to the increase in the number of people with disability or dementia. These elderly with disability or dementia would heavily reply on wheelchairs in their daily activities [2]. Moreover, people who are seriously injured in accidents like traffic accidents or slips may not be able to walk freely and would also require the use of wheelchairs for a long time.

The health condition and safety of these patients are highly concerned. Patients who have used wheelchairs for a long time could keep the same sitting posture for hours, which may raise the risk of developing pressure ulcers/sores [3]. Moreover, the stress on the bones would be increased and a bending moment could be caused to the intervertebral discs when a patient is accustomed to leaning to a single side while sitting [4]. The worst is that wheelchair users may sometimes slip off the wheelchairs, which can be quite dangerous especially when they are left alone. In order to solve the problems mentioned above, this work is aimed at improving the health condition and safety of wheelchair users. In this work, a rehabilitation seat cushion system is designed and implemented to upgrade the function of wheelchairs.

II. INTRODUCTION REHABILITATION SEAT CUSHION SYSTEM

The rehabilitation seat cushion system mainly contains two important functions. One of the functions is the detection of bad sitting postures, where two force sensing resistors (FSRs) are used as pressure sensors to detect whether there is a tilt in the sitting position of the wheelchair user. If the wheelchair user keeps sitting with a tilted posture for a long time, the system will turn on an LED to remind the user. In addition, another sensor is set behind the seat cushion to detect whether the wheelchair user has slipped off the wheelchair. The buzzer on the sensing device would make a sound to alert the caregiver when the patient slips off the wheelchair. The other function offered by the system is the rehabilitation function. A real-time feedback rehabilitation training software is designed for the wheelchair users to conduct rehabilitation exercises or balance training with the posture detection system in the seat cushion.

As previously mentioned, two pressure sensors are used in the system for bad sitting posture detection, where one is fixed on the left side and the other on the right side of the seat cushion. The FSR is a resistor whose resistance is inversely proportional to the pressure applied. The resistance is close to infinity as no pressure is applied and would decrease as the applying pressure increases. The bearing pressure is in the range of 10 to 10,000 grams (g).

The resistance change is obtained by measuring the difference between the resistance of a FSR and a fixed resistance. The resistance changes are then converted into voltages. Then, the voltages are delivered to the microprocessor for digitalization by the analog-to-digital converter (ADC).

The voltages associated with the two FSR sensors would be compared, which enables the system to tell whether the user sitting on the wheelchair leans to left or right. If a large voltage difference occurs and lasts for a while, the reminder LED lights will turn on. In addition, a third sensor is placed behind the seat cushion and connected to the buzzer. When the voltage associated with this sensor becomes nearly ground to indicate a zero pressure condition that the user may slip out of wheelchair, the buzzer will immediately make a sound as a warning generated by the system. The procedure of sitting posture detection is illustrated by the flow chart shown in Fig. 1.

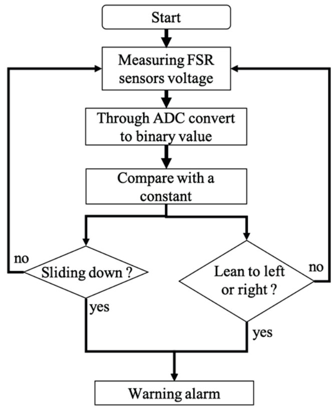


Fig. 1. Sensing system flow chart

In this system, all of FSR sensors and devices get 5V power from Universal Serial Bus (USB) port, Fig. 2.The procedure of real-time feedback rehabilitation training is illustrated by the flow chart shown in Fig. 3. After the digital output associated with the FSR sensor is obtained from the ADC, instead of being compared with the digital output associated with a fixed resistor, it is delivered to the computer via RS232. The real-time feedback rehabilitation training software is developed using Visual Basic (VB). The digital output associated with the FSR sensor would be used as the parameters of the rehabilitation training software so that the user can move the body toward either left or right to control the direction of the car in the training program.

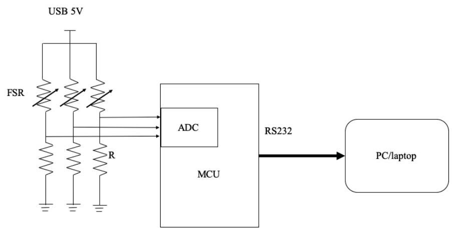


Fig. 2. The system diagram

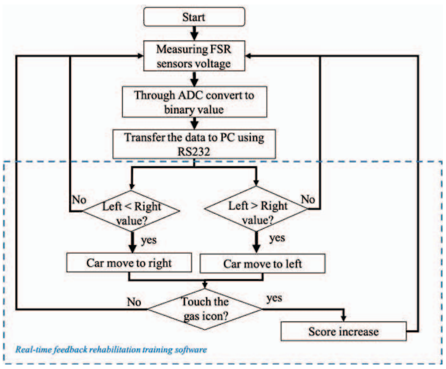


Fig. 3. Rehabilitation training system diagram

III. SYSTEM IMPLEMENTATION

The cushion system is battery-powered and can be directly mounted on a wheelchair. Users can select among different modes of the system according to their needs. The photo of the circuit board is shown in Fig. 4.

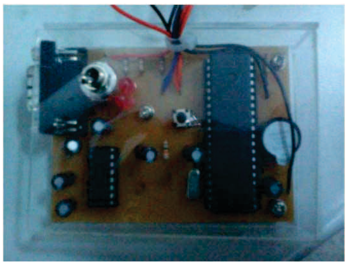


Fig. 4. Photo of the circuit board

As previously mentioned, the system provides a real-time feedback rehabilitation training software. Once the user starts up the training software, the pressure distribution associated with the sitting posture of the user would be first detected by the pressure sensor set on the cushion. The user can adjust optimize his sitting posture according to the obtained pressure distribution. The pressure distribution would also be used to adjust the control over the car direction while using the training software. The training screen was shown in Fig. 5. The obstacles in the training software would move from top to bottom and the user needs to use the body movement to control the moving direction of the car to avoid obstacles in the road. In the training software, the car speed can be increased to raise the training intensity because the user would need to make larger and quicker body movements. Through the training, the reaction and eye-body coordination of the user would be greatly improved not to mention the rehabilitation is so interesting. The pressure ulcers resulted from sitting in the wheelchair could be avoided if the user conducts the training frequently.

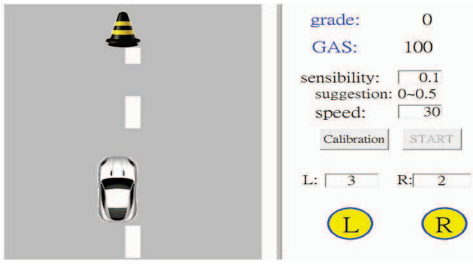


Fig. 5. Training screen of the software.

IV. CONCLUSION

A rehabilitation seat cushion system has been developed for sedentary and long-term wheelchair users. System functions, such as the reminder of bad sitting posture, alert of slipping and a real-time rehabilitation training software, have been implemented and verified. The system not only guarantees the safety of the user but also helps to improve the health condition of the user. In fact, the cushion system is also suitable for ordinary people who usually sit in the office for hours and may be lack of exercise.

REFERENCES

[1] Y. J. Wu, W.J. Liu, and C.H. Yuan, "A mobile-based barrier-free service transportation platform for people with disabilities," Computers in Human Behavior, Nov. 2018.

[2] C. L. Chang and K. M. Chen, "Physical and mental health status and their correlations among older wheelchair users with dementia in long-term care facilities," Quality of Life Research, vol. 27, no. 3, pp. 793-800, Mar. 2018.

[3] T. H. Bui, P. Lestriez, D. Pradon, K. Debray, and R. Taiar, "The Prevention of Pressure Ulcers: Biomechanical Modelization and Simulation of Human Seat Cushion Contributions," Proc. of the International Conference on Advances in Computational Mechanics, Lecture Notes in Mechanical Engineering, Singapore, pp. 1157-1170, 2017.

[4] C. J. Snijders, P. F. Hermans, R. Niesing, C.W. Spoor and R. Stoeckart, "The influence of slouching and lumbar support on iliolumbar ligaments, intervertebral discs and sacroiliac joints," Clinical Biomechanics, vol. 19, pp. 323-329, May 2004.

Automatic Spine Vertebra segmentation in CT images using Deep Learning

Ping-Cheng Wu[1], Teng-Yi Huang[1], Chun-Jung Juan[2,3,4]
1. Department of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan
2. Department of Medical Imaging, China Medical University Hsinchu Hospital, Hsinchu, Taiwan, Republic of China
3. Department of Radiology, School of Medicine, College of Medicine, China Medical University, Taichung, Taiwan, Republic of China
4. Department of Medical Imaging, China Medical University Hospital, Taichung, Taiwan, Republic of China
m10707317@mail.ntust.edu.tw

***Abstract*—In this study, we aimed to develop a fully automatic segmentation and classification system for spine vertebra regions. We used datasets provided by xVertSeg and trained a SegNet segmentation model. In our preliminary result, this model was able to identify the spine vertebra regions and levels.**

I. INTRODUCTION

In clinical applications of spine CT, analysis of individual vertebrae such as their shapes and structures could produce indicators for diagnosis of spinal pathologies. The recently advanced deep learning method could lead to a fully automatic pipeline of segmentation and classification of fractured vertebrae. In this study, we aimed to develop a pixel-wise classification of vertebra regions using the SegNet based on the convolutional neural network.

II. METHODS AND MATERIALS

A. Spine dataset

We obtained the spine data from xVertSeg provided by the Laboratory of Imaging Technologies, University of Ljubljana, Slovenia[1-3]. The xVertSeg dataset is built on the two previous challenges held in MICCAI 2014 and 2015. It consists of CT lumbar spine images acquired from 15 subjects. The corresponding masks of manually outlined vertebra regions are also provided. Also, the level of each lumbar vertebra is defined by two observers. The label numbers for the background, L1, L2, L3, L4, and L5 are 0, 200, 210, 220, 230, and 240, respectively.

B. SegNet structure

We used 2D SegNet structure for pixel-wised classifications. Figure 1 shows the illustration of the SegNet model. It is an end-to-end model. The left triangle is the encoder part extracting image features with convolution and max-pooling layers. The right triangle is the decoder part performing transpose convolution and up-sampling.

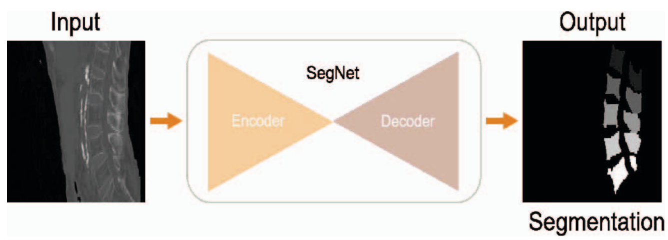


Figure 1 SegNet architecture

C. Pre-process and model training

The CT datasets are three-dimensional volumes. We totally extracted 14,336 2D sagittal images from the volumes. Among the images, there are 3,538 images identified as images, including vertebra labels. We performed several image pre-processing including image normalization to gray level values of 0-255, and resizing image matrix to 256 x 256. Of all the 15 patient datasets, we selected images of 12 patients (image003 to image014) as the training images and the rest images of 3 patients (image001, image002, and image015) as the testing images. We totally got 11,264 images for training and 3,072 images for testing. We first shuffled the images and then sequentially selected six images including vertebra labels as the input batch (256x256x6) for each model training step. We performed model training with the stacked batch for 60,000 steps. Then, we switched to another batch configuration. That is, we configured the batch into a different configuration, i.e., a 6-image batch with 3 images with vertebra labels and the other 3 images without vertebra labels. We kept model training with the second batch configuration for 45,000 steps and got the final SegNet model. During the training stage, we used image augmentations including random left-right and up-down flips, random image rotations (-15 to 15 degrees), and random image contrast adjustment (factor: 0.6 to 1.4).

D. Predictions and post-processing

For the testing datasets, we first extracted the 2D images and performed the same preprocessing steps as the training images. Then, the procedure fed the images into the trained network to obtain a pixel-by-pixel classification. We re-stacked all the obtained 2D segmentation images of one subject into a 3D volume. Finally, we used morphological labeling to collect 3D pixel clusters and removed clusters with less than 300 pixels.

III. Results

Table 1 displays the Dice coefficients of the predicted images before and after image post-processing. According to the results, the pixel-removal post-processing prominently improve the segmentation accuracy. In this preliminary result, the SegNet model was not able to properly classify the vertebra levels L1 to L3.

Table 1 summary of the prediction metrics

	ID	L1_dice_list	L2_dice_list	L3_dice_list	L4_dice_list	L5_dice_list	Seg_dice
Result	image001	0.225	0.332	0.44	0.763	0.667	0.651
	image002	0	0	0.066	0.369	0.712	0.561
	image015	0.347	0.6	0.626	0.557	0.718	0.731
	Average	0.191	0.311	0.377	0.563	0.699	0.647
Post process	image001	0.233	0.337	0.451	0.773	0.716	0.672
	image002	0	0	0.069	0.377	0.792	0.588
	image015	0.347	0.6	0.626	0.557	0.796	0.75
	Average	**0.193**	**0.312**	**0.382**	**0.569**	**0.768**	**0.669**

Figure 2 shows the predicted vertebra regions L1 to L5 of subject #15. The SegNet model not only produced the vertebral regions but also correctly classified the levels of L1 to L5.

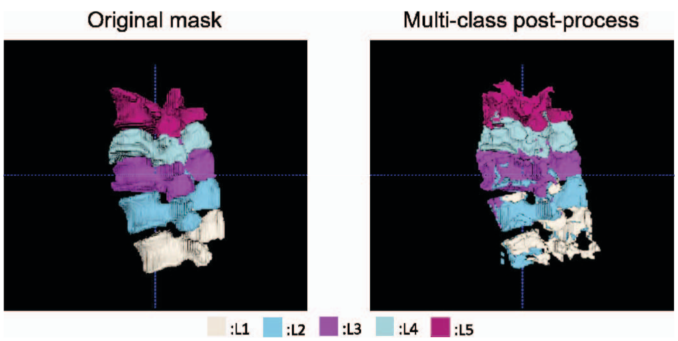


Figure 2 the true and the predicted label of the subject no. 15.

IV. Discussions and Conclusions

In this study, we attempted to develop an automatic segmentation method for spine vertebrae at the level of 1 to 5. We used the end-to-end SegNet model, which was initially developed for scene parsing applications, especially the unmanned cars. In the scene parsing applications, the SegNet model identifies the roads, buildings as well as traffic lights from videos recorded in vehicles. We adapted the model for the vertebra segmentation and identified the 5 levels of vertebra regions. In this preliminary result, the model properly classified the vertebra regions. However, the Dice coefficients of L1 to L3 were low (0.193 to 0.382). Optimizing the deep learning algorithms and increasing the training data size may improve the performance of segmentation and thus merit further investigations. In conclusion, the deep learning method could be practical in the application of vertebra segmentation and classification.

References

[1] Robert Korez, Bulat Ibragimov, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec, “A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation”, IEEE Transactions on Medical Imaging, 34(8):1649-1662, 2015

[2] Bulat Ibragimov, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec, “Shape representation for efficient landmark-based segmentation in 3D”, IEEE Transactions on Medical Imaging, 33(4):861-874, 2014

[3] Darko Štern, Vesna Njagulj, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec, “Quantitative vertebral morphometry based on parametric modeling of vertebral bodies in 3D”, Osteoporosis International, 24(4):1357-1368, 2013

[4] Vijay Badrinarayanan, Alex Kendall, Roberto Chipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Scene Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Jan 2017.

Trigeminal Neuralgia Alleviation on Demand with an CMOS SoC Using Current-mode Pulsed Radio-Frequency Stimulation

Hung-Wei Chiu , Kun-Ying Yeh, Shey-Shi Lu, Mu-Lien Lin

Hung-Wei Chiu
Department of Electronic Egineering
National Taipei University of Technology
Taipei Taiwan
hwchiu@ntut.edu.tw

Kun-Ying Yeh, Shey-Shi Lu
Graduate Institute of Electronics Engineering
National Taiwan University
Taipei, Taiwan
d98943028@ntu.edu.tw

Mu-Lien Lin
Department of Anesthesiology
Taipei City Hospital, Zhongxing Branch
Taipei, Taiwan
linmulien@yahoo.com.tw

***Abstract*—This paper presents a design of batteryless CMOS system-on-a-chip (SoC) with dual-mode programmable pulsed radio-frequency (PRF) stimulation, temperature detection, and wireless communication for trigeminal neuralgia alleviation on demand. The system is capable of providing square wave, sinusoid wave, and patterned stimulation waveform, and the stimulation parameters which include pulse frequency, amplitude, duty cycle, and pulse duration, can be wirelessly reconfigured by the external handheld device. The temperature sensor incorporated with analog front-end circuit can detect the temperature of body and SoC chip for the prevention of possible abnormal operation. All the information of stimulation parameters and temperature can be real-time monitored by a handheld device through wireless communication, which follows medical implanted communication system (MICS) standard. This SoC chip, whose dimension is 3.11 × 2.73 mm2, is fabricated in standard 0.35um CMOS process and mounted on a compact test board, whose size is 2.6 × 2.1 cm2, for delivering PRF stimulation in animal studies.**

Keywords*—*Trigeminal neuralgia, PRF, Pain relief

I. Introduction

Trigeminal neuralgia has been known as a “suicide disease”, is not life-threatening but very disabling and sufficiently painful for patients to commit a suicide [1-2]. This kind of extreme pain needs to be stopped or mitigated immediately when it occurs. Fig. 1 shows the application scenario of the proposed dual-mode PRF stimulation SoC. The device containing the PRF stimulation SoC chip will be implanted behind the ear and mandible to avoid interfering patient's daily activities, and the electrodes will be inserted close to the pain-causing nerve to ensure the effectiveness. This implantable device can achieve prompt pain alleviation by simply controlling a handheld device to initiate the PRF stimulation when TN occurs to the patient. Stimulation parameters are designed to be configurable for fitting the practical requirement of each TN patient. Although PRF stimulation is such a non-destructive and reliable option for TN treatment, it is known for its short-term effectiveness on pain alleviation. Being incorporated with an implantable device will make this shortcoming minimized and enable TN patients to relieve pain anytime anywhere by themselves.

II. SYSTEM ARCHITECTURE

Fig. 1 depicts the system block diagram of proposed PRF Stimulation SoC consisting of a radio-frequency to direct current (RF-DC) converter, a voltage limiter, regulators, a clock generator, a micro-controller unit (MCU), a temperature sensor, an analog front-end circuit, a 10-bit SAR ADC, an envelope detector, a low power transmitter, and a stimulation

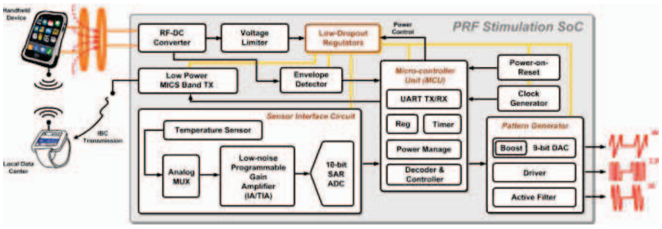


Fig. 1. System Architecture

pattern generator. Supply power comes from receiving 1MHz RF outside skin coupling, which is chosen by easy user alignment and good penetration depth. RF-DC circuit converts received RF power to dc voltage, and the subsequent voltage limiter limits the dc voltage level to a maximum of 5V, which can be regulated by regulators specified for each power domain. When system V_{DD} is available, 10MHz clock generator starts to work and thus enable the micro-controller unit. Temperature information delivered by the temperature sensor will go to low-noise PGA, and then be digitized by the 10-bit SAR ADC. These bio-information transmitted by 403MHz signals following medical implanted communication system (MICS) standard can be real-time monitored by a handheld device. Users can also define stimulation parameters wirelessly in the same way. Pattern generator is able to provide programmable output stimulation pattern in either current or voltage mode.

An instrumentation amplifier (IA) with large input impedance is responsible for handling the temperature signal and RF-DC output voltage. Traditional three op-amps IA suffers from limited input common mode range and insufficient CMRR due to mismatch of the feedback resistors. Further, because the sensing signals are at low frequencies in this system, the dominant in-band noises of analog front-end circuit are flicker noise and dc offset. Considering this point, the chopper technique along with rail-to-rail configuration is applied to the design of AFE so that it can achieve low noise while providing wide input dynamic range.The schematic of AFE, which comprises a rail-to-rail chopper instrumentation amplifier with resistive feedback and a third order Sallen-Key low-pass filter with 50 Hz bandwidth. The chopper IA is made of a rail-to-rail differential difference amplifier (RRDDA) with loop gain defined by the ratio of feedback resistors. Also, DDA-based architecture has two symmetric chopper switches in its input and hence reduces the residual offset caused by the spikes from input switches. The low-pass filter is added to eliminate the unwanted modulated noise and to further suppress the spikes from clock feedthrough and charge injection caused by the non-idealities of output chopper

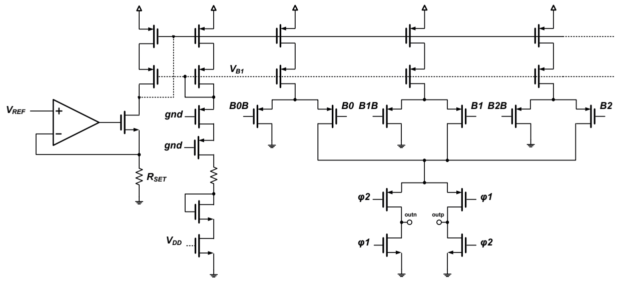


Fig. 2. Schematic of DAC

switches. To make the loop gain adjustable, an off-chip SMD resistor is used as the feedback resistor R_2 here.

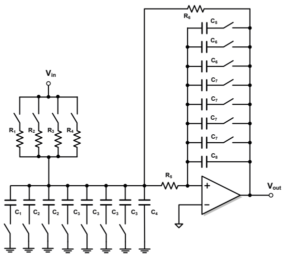
The pattern generator in this system provides both voltage-mode and current-mode waveform options to achieve the best effectiveness under well control. Voltage-mode square wave of PRF pattern is generated in MCU and connected to two basic inverter chains for double stimulating voltage and better driving capability. The inverting-based driver delivers square waves of constant voltage DV_{DD} (3.3V), and the voltage across the nerve is going to be positive or negative DV_{DD} alternatively [3]. Because of the uncertain impedance of the stimulated nerve, maximum output current has to be limited for safety. Thus two current-limit resistors are added in the last stage of inverter chain. The output stimulation current will be limited to 10 mA if the loading becomes unexpectedly low (several ohms).

According to previous experimental results, the PRF stimulation in sinusoid waveform will lead to less charge accumulation in neural ganglions comparing to that in square waveform. Hence a configurable active low-pass filter is added following the inverting-base driver to eliminate odd-order harmonics and deliver pure sinusoid wave for stimulation. Further, Chebyshev filter is chosen for have a steeper roll-off around the cutoff frequency. Two cutoff frequency options, 250 kHz and 500 kHz, are designed to be compatible with pulse frequencies the square wave might be. As shown in Fig. 3, multi-feedback filter topology is used, and thus it makes the pulse amplitude of sinusoid wave delivered by the filter able to be configured as 0.5, 0.667, 0.833, and $1V_{pp}$.

Ion exchange between cells is one of the most significant mechanisms for lots of reactions in creatures. For this reason, current mode stimulation is considered as a favorable method for biological analysis, and consequently a 9-bit current-steering DAC, as depicted in Fig. 2, is integrated in this system to provide current mode PRF stimulation. A typical folded-cascode OPamp is used to bring the reference voltage VREF comes from bandgap circuit (around 1.2 V) on the current path, and hence the reference current can be well defined as VREF divided by an external resistor RSET. The reference current is mirrored to current cells of 6-bit thermometer (8x8 array) and 3-bit binary combination to complete 512 steps. The unit current chosen is 3 μA, and the full-scale output current is about 1.5 mA. There are four switches on the output current path to change the output polarity, and thus the equivalent stimulation current could be doubled effectively.

III. MEASUREMENT

This current-mode stimulation chip is fabricated in a standard 0.35um CMOS process. Fig. 4 shows the die micrograph of



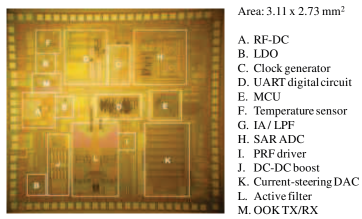


Fig. 3. Schematic of active filter Fig. 4 Chip micrograph

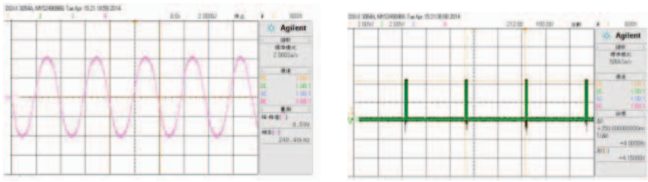


Fig. 5. Stimulation output waveform(a) sinusoid (b) square

the proposed trigeminal neuralgia alleviation chip whose size is 3.11 mm × 2.73 mm. Table I summarizes the measured performance of each block in this chip. Several specific power domains are designed and served by separate low-dropout regulators to balance the circuit performance and power consumption. Multiple options of stimulation pattern, whose aiming diseases are trigeminal neuralgia and epilepsy, are built in the system. The research about epileptic seizure control is realized lately, and thus trigeminal neuralgia is chosen for demonstration in the following animal study. The stimulation patterns in 50%, 25%, 10%, and 5% pulse duration are effectively generated by the generator. Fig. 5 shows measurement results in the modulations of other three stimulation parameters. 250KHz / 500KHz in pulse frequency, 2Hz / 4Hz in repetition rate, and 0.67Vpp / 1Vpp in pulse amplitude of modulations are also verified. All the PRF stimulation parameters in both current-mode and voltage-mode are successfully realized and generated, and all the options can be programmed wirelessly by the user's handheld device.

Table I. Basic PRF stimulation parameters and options

Technology	tsmc 0.35μm CMOS
Chip Area	3.11 x 2.73 mm²
Temperature Sensor	
Power Consumption	194.2μW @3V
R-Squared Value	0.999
Sense Range	-20 ~ 120 ℃
Analog Front-end	
Power Consumption	IA: 387μW @3V LPF: 130μW @3V
Closed-loop Gain	60 ~ 80 dB
ICMR	0 ~ 2.75V
CMRR	137 dB
Unit-gain BW	1.7 MHz
Input referred noise	70nV/√Hz @10Hz
Offset Voltage	31μV
Chopping Freq.	10 KHz
MCU	
Clock Rate	20MHz
Power Consumption	14.5mW @3V

SAR ADC	
Resolution	10 bit
Conversion Rate	100 KSPS
DNL	+0.62 / -0.74 LSB
INL	+0.60 / -0.58 LSB
ENOB	9.17 bit @49.6KHz
Power Consumption	2.1μW @1KSPS 20μW @10KSPS
Pattern Generator	
DAC Resolution	9 bit
DAC Output Range	0 ~ 1.54 mA
DNL	+0.3 / -0.3 LSB
INL	+0.6 / -0.6 LSB
Power Consumption	Voltage-mode: 2.05mW @7Kohm
OOK Transmitter	
Operating Freq.	403 MHz
Output Power	-12.9 dBm
Power Consumption	2.16mW @1.8V

REFERENCES

[1] C. J. Woolf and R. J. Mannion, "Neuropathic pain: Aetiology, symptoms, mechanisms, and management," Lancet, vol. 353, no. 9168, pp.1959–1964, Jun. 1999.

[2] H. Merskey, N. Bogduk, "Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms", Seattle: IASP Press; 1994; 59–71

[3] H. W. Chiu, M. L. Lin, C. W. Lin, I. H. Ho, W. T. Lin, P. H. Fang, Y. C. Lee, Y. R. Wen, and S. S. Lu, "Pain Control on Demand Based on Pulsed Radio-Frequency Stimulation of the Dorsal Root Ganglion Using a Batteryless Implantable CMOS SoC," IEEE Trans. Biom. Circuits Syst., vol. 4, pp. 350-359, Dec. 2010

Cervical Image Segmentation using U-Net Model

Yao Liu[1], Bing Bai[2], Hua-Ching Chen[3], Peizhong Liu[4] and Hsuan-Ming Feng[5]
National Quemoy Univisity[1,3,5], Huaqiao Univisity[2,4]
yowk0529@gmail.com[1],hmfeng@nqu.edu.tw[5]

Abstract--*In this paper, a pre-training system based on U-Net model is used to colposcopic image segmentation. This method is proposed to convert the full connection layer in the traditional convolution neural network into the convolution layer, and uses the deconvolution structure to realize the upsampling data, splits the feature graph on the channel dimension to understand the feature fusion, and adopts the skipping structure to combine the low layer and the high level feature. Experimental results showed that the proposed method based on U-Net model has the better effect in colposcopic cervical image segmentation.*

Keywords—*cervical cancer; colposcopic images; segmentation ;U-Net Model;*

I. INTRODUCTION

Cervical cancer remains an important cause of female mortality all over the world, especially in the low-income countries [1]. Relevant studies have shown that its pathological development has a long stage of precancerous lesions. Generally, clinicians can find lesions as early as possible at this stage and remove the injured tissues to prevent cervical cancer [2]. Colposcopy is a most common tool used in clinical screening of the cervical cancer. In the process of analysis of cervical images, the segmentation of images is the key step [3].

Due to the influence of bleeding and inflammation, there are various tissue or impurity pollution, and overlapping adhesion in cervical images [4]. Therefore, it is difficult to achieve the ideal segmentation effect by using conventional image segmentation methods, such as threshold method, boundary tracking way[5].In view of the above-mentioned characteristics of the cervical image, this paper presents an effective segmentation method.

II. METHOD

Image segmentation needs pixel level segmentation, which needs to preserve a lot of the position information of images. However, the pooling layer will affect the results of image segmentation, loses some spatial information in the process of up-sampling, and reduces the accuracy of image segmentation. One way to solve this problem is to use encoder-decoder structure in the network, in which the direct connection is added so that the decoder can repair the details more perfectly when the decoder is up-sampled. U-Net is developed on the basis of full convolution neural network, which mainly improves the upsampling part [6].The network structure of U-net is shown in **Figure 1**, in which the blue part represents the convolution layer with convolution kernel size of 3× 3 and the gray part of ReLU, represents replication and splicing with the pointed feature graph, the red part represents the pooled layer, the green part represents the convolution, and the last represents the convolution with the convolution kernel size of 1×1.Only convolution layer, pool layer and reverse convolution layer are used in the network structure to realize pixel segmentation and end-to-end image segmentation [7].

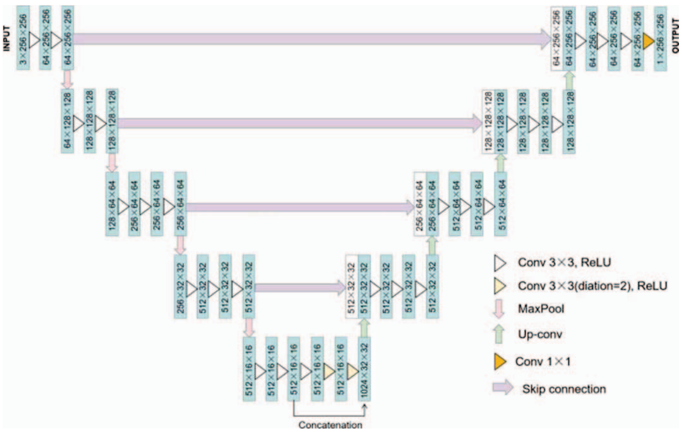


Figure 1: Network structure of U-net

U-net adopts the classical encoder-decoder structure. As shown in **Figure 1**, the encoder on the left can obtain the context information of the image in order to shrink the path, which usually adopts the classical image classification network which removes the full connection layer, which carries on the convolution kernel pooling operation to the original input picture, and the decoder on the right side can accurately locate the segmentation task at the pixel level for the expansion path [8].

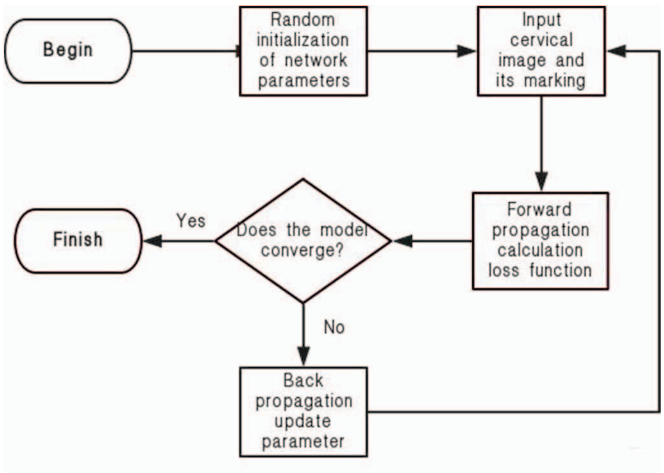


Figure 2: Flowchart of training model

The image and its corresponding marking image are input into U-net network, and then the iterative training cycle is carried out to calculate the loss function value. Iteratively update the network parameters to make the smaller loss function value. When the loss function value is reduced into a small one, the model is converged into stable state.

. In this experiment, the original U-net network is used, and the gradient drop algorithm with momentum parameters is determined to optimize the model. In here, the momentum value is set to 0. 99, and there is 0.01 attenuation; the parameters in the network are randomly initiated, the initial learning rate is set to 1e-5, and the attenuation coefficient of learning rate is set to 0.0005, the learning rate is updated every 10 iterations. When the loss function value no longer drops, the model approaches the appropriated state and the optimal parameters are saved.

III. EXPERIMENT RESULT AND DISCUSSION

A. *Simulation specification*

The cervical images in this experiment were firstly marked by colposcopy experts and 2500 labeled images are trained the U-Net model in training cycle. There are another 100 test images were proposed to verify the performance. Before the training procedure, all images are converted into the grayed domain for reducing the training cycle time. The training code executes at a workstation with NVIDIA 1660 GPU and 6GB memory. Training time takes about 20 hours at each training piece. The segmentation results are evaluated by using the performance index of average pixel accuracy (MPA), average intersection and (MIoU), frequency weight intersect ion (FWIoU).

In this experiment, the original U-net network is used, and the gradient drop algorithm with momentum parameters is determined to optimize the model. In here, the momentum value is set to 0. 99, and there is 0.01 attenuation; the parameters in the network are randomly initiated, the initial learning rate is set to 1e-5, and the attenuation coefficient of learning rate is set to 0.0005, the learning rate is updated every 10 iterations. When the loss function value no longer drops, the model approaches the appropriated state and the optimal parameters are saved.

B. *Results*

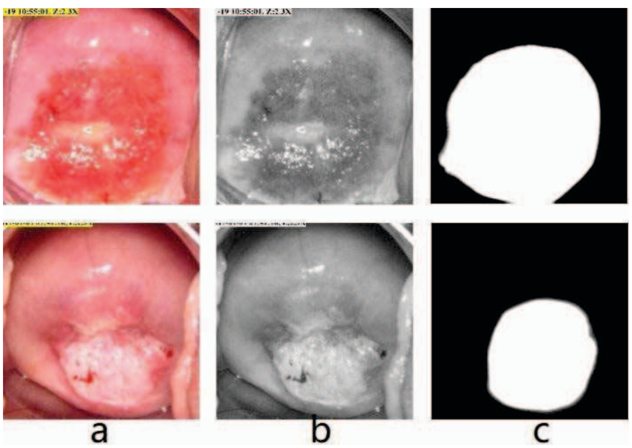


Figure3: Experimental Results: a. Original b.Grayscale c.After Segmentation

Table1 Evaluation results for three performance indices

Evaluation / Model	MPA	MIoU	FWIoU
U-Net	0.8736	0.7786	0.7895

After finishing the training cycle, cervical image segmentation model is gradually converged into the better system to recognize which image is a cancel one or not? A random test image is picked up to evaluate the performance with three indexes of the developed model. In every training cycle, the loss function value of the regulated model is appeared into the smaller and stable state for improving accuracy on the verification purpose.

IV. CONLUSIONS

In this paper, the neural network model of U-Net is used to segment the cervical image, which adopts the classical encoder-decoder structure, including contraction path and expansion path. The feature map is spliced into the smaller dimension of image channel and then fusing features, which further improves the accuracy results of image segmentation. The experimental results showed that the overall segmentation effect of cervical image is very good, and the phenomenon of over-segmentation and under-segmentation is relatively reduced. One disadvantage is that the segmentation accuracy of edge region is not enough. Generally speaking, this method has certain reference value for the diagnosis of cervical lesions. It offers the great significance to improve the accuracy of cervical screening.

ACKNOWLEDGMENT

This research was supported in part by the Ministry of Education of the Republic of China under deep learning module of teaching course.

REFERENCES

[1] R. P. Kauffman, S. J. Griffin, J. D. Lund, and P. E. Tullar, Current recommendations for cervical cancer screening: Do they render the annual pelvic examination obsolete, J. Med. Princ. Pract., vol. 22, no. 4, 2013, pp. 313–322.

[2] C. F. D. Control et al., Cervical cancer screening among women aged 18-30 years—United states, 2000-2010. J. MMWR Morbidity Mortality Weekly Rep., vol. 61, no. 51–52, 2013, pp. 1038.

[3] J.p. Fan, Research on image segmentation and identification of cervical cells in a doctoral thesis on biomedical engineering, Jinan University, 29-Apr-2010.

[4] W. H. Ren, J. D. Tian, and Y. D. Tang, Specular reflection separation with color-lines constraint. J. IEEE Transactions on Image Processing, vol. 26, no. 5, 2017, pp. 2327-2337.

[5] Z. Y. Hunan ,Image segmentation method based on threshold,City College School of Information and Electronics Engineering, 12 June, 2017.

[6] Q. An, M. Zheng, Image Research based on depth Learning, Automation and Instruments, third issue 2018, pp.115-117

[7] R.b. Olaf, F. h. Philipp, and Brox T.m., U-Net: Convolutional Networks for Biomedical Image Segmentation, 18 May, 2015.

[8] G.g. Li, Nuclear Image Segmentation method based on depth Learning, Harbin University of Technology, June 2018, pp.35.

Using QR Code Labels to Enhance OCR for Capturing Legacy Machines' Data

Suk-Ling Lai, Boon-Yaik Ooi
Faculty of Information and Communication Technology,
Universiti Tunku Abdul Rahman,
Kampar, Malaysia.
sukling97@1utar.my, ooiby@utar.edu.my

Yen-Lin Chen
Dept. of Computer Science and Information Engineering,
National Taipei University of Technology,
Taipei, Taiwan.
ylchen@csie.ntunt.edu.tw

Abstract – Despite the maturity of optical character recognition (OCR) technology, reading data from screens such as 7 segments and LCD screens possess some challenges such as identification, position and orientation for conventional OCR to be used directly for recognizing values from screens belongs to legacy manufacturing machines. Very often, these legacy machines do not have programmable interfaces and network connectivity thus the only way to capture its data is through optical means. This work, introduces the use of QR code (Quick Respond Code) labels as marker to guide the pre-processing steps of the capture image before using conventional OCR such as Tesseract to recognize the value in the image. From the QR code image, the system can detect if the capture image is tilted or slanted. This work shows that with such information the image can be processed and attain higher accuracy.

Keywords ---- Industry 4.0, Internet-of-Things, Legacy Manufacturing System, Optical Character Recognition

I. INTRODUCTION

Although we are now in the era of industrial 4.0, there are still a lot of factories who do not update to the era of industrial 4.0 due to some reasons. The most common reason will be the insufficient of capital to update their machines or system and their environment. Besides, there might be some operators who think that they have used to the old machines and solutions and do not want to adapt to the new technology, or perhaps some products can only produce by the proprietary solution which customization is not allowed. Furthermore, the legacy machines are all offline, therefore they might need people to monitor the machines, to observe the value on the LCD/LED display to avoid operation failures. Despite the advancement of OCR technology, the use of OCR for reading LCD from industrial machines are limited.

This study is partially sponsored by the Ministry of Science and Technology of the Republic of China under the contract No. MOST-107-2218-E-027-018 and also partially sponsored by Collaborative Research in Engineering, Science and Technology (CREST), Malaysia under the project No. URP1806.

Although the process is convenient for physical documents processing, it is not suitable for capturing readings from industrial machines, because due to the nature of OCR, tilted or slanted will still be processed regardless if the value is correct or not.

The proposed solution is to detect slanted or tilted images through the use of QR code before submitting the image for OCR to process. The QR code is also embedded with semantic information of the readings. For instance, location, labels, units and unique identification. In short, the proposed technique improves the confident of the values produced by contentional OCR through the use of QR code as a reference point.

II. THE PROPOSED SOLUTION

A. Detecting ROI (Region of Interest)

Tesseract is used as the OCR engine for this project. To eliminate human intervention and calibration, the system able to detect the ROI instead of processing the whole image captured. To know the position and size of the ROI, QR Codes are used for detecting the ROI. 4 points of the QR Code are enough to know the ROI, extract the x and y value of the points and the size and area of ROI is known and calculated. The ROI will be crop according to the size of the ROI. Since there will have flicker problem when capturing image from LCD/LED screen, the cropped image have to convert into monochrome. After processing the image and get the data, the image will be delete to save the memory space. Besides, the metadata of the image will be extract and upload to the cloud along with the data. The update of the cloud do not need human intervention, it will update to the cloud automatically.

B. Detect and reposition the tilted image and eliminate image capture from certain angle

The system could not detect and process the image if the image is tilted. With the existing of QR Code, the problem could be solve. When the top left and top right of

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

the QR code are not in a same Y location, meaning that the image is tilted. The degree of rotation is able to find out and hence the image can be rotate to the position that can perform the recognition task the best with the formula of:

$$c^2 = a^2 + b^2 - 2ab\cos\gamma \quad (2)$$

Other than detecting the ROI, there is a problem with the angle to take pictures. When the camera is set at certain angles from the screen, the result will be inaccurate as the camera is not able to 'see' the data from the screen. For example when 180 degree from the screen. To avoid from getting the inaccurate result, those images have to be ignored and not processing them. To let the system know the current angle of the camera, use the area ratio of the QR codes. When the camera is placed right in front of the screen, then ratio should be approximate to 1:1. Therefore, for every image that has the ratio less than 0.7, the image will be ignore and deleted. The area of the QR code can be calculated as:

$$Area\ of\ QR\ Code = (|X_1\text{-}X_2|) * (|Y_1\text{-}Y_2|) \quad (3)$$

The ratio of the QR codes can be calculated by:

$$Ratio\ of\ QR\ Codes = \frac{Area\ of\ QR\ Code\ 1}{Area\ of\ QR\ Code\ 2} \quad (4)$$

III. EXPERIMENTAL RESULT

The test in this work is conducted under 74.7Lux, by setting the Tesseract engine 30cm at various angles. Fig. 1 is the OCR accuracy against angle from the screen.

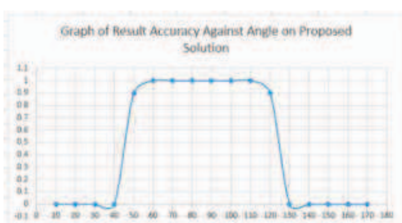


Fig.1: Graph of Result Accuracy against Angle on Proposed Solution

From Fig. 1 it shows that values are correctly recognized when the camera angle from the screen is between 60 and 110 degrees. This information is then mapped back to the QR code labels. As such, when QR code shows that the image taken is beyond the angle range, images will be discarded. We subsequently, tested the system with QR code using 40 images and obtained close to 100% accuracy.

IV. CONCLUSION AND FUTURE WORK

This project proposes the use of QR code labels to provide information regarding the image that being capture before feeding it for the OCR to process. As such, the propose solution can confidently reduce the errors of OCR. The future of this work is to convert the solution into a mobile application so that it allows human operator to capture the readings by use using smartphone cameras. This will reduces the chances of getting the readings wrongly form legacy manufacturing machines in factories.

REFERENCES

[1] Tekin, Ender; Coughlan, James M.; Shen, HuiYing, "Real-Time Detection and Reading of LED/LCD Displays for Visually Impaired Persons," pp. 491-496, 2011.

[2] J. Xie, "Optical Character Recognition Based on Least Square Support Vector Machine," pp. 626-629, 2009.

[3] S. Gleichman, B. Ophir, A. Geva, M. Marder, E. Barkan and E. Packer, "Detection and Segmentation of Antialiased Text in Screen Images," pp. 424-428, 2011.

[4] M. Cutter and R. Manduchi, "Towards Mobile OCR," *How To Take a Good Picture of a Document Without Sight,* pp. 75-84, 2015.

[5] S. Steven, "Screen OCR Super Easy To Use," [Online]. Available: https://easyscreenocr.com/how-to-convert-screenshot-to-text-with-a-free-ocr-software/. [Accessed 1 August 2018].

[6] E. Chan, "Hackernoon," 2015. [Online]. Available: https://hackernoon.com/optical-character-recognition-with-google-cloud-vision-api-255bb8241235. [Accessed 1 August 2018].

[7] T. L. V. C. O. Indiana, "DaVinci HD?OCR Video Magnifier," [Online]. Available: http://www.eyeassociates.com/davinci-hdocr-video-magnifier/. [Accessed 1 August 2018].

[8] "Enhance Vision," 2018. [Online]. Available: https://www.enhancedvision.com/downloads/user-manual/DaVinciPro/davinci-pro-user-manual-english.pdf. [Accessed 3 August 2018].

[9] L. Vincent, "Google Code Blog," [Online]. Available: https://web.archive.org/web/20061026075310/http://google-code-updates.blogspot.com/2006/08/announcing-tesseract-ocr.html . [Accessed 29 July 2018].

[10] R. Smith, D. Antonova and D.-S. Lee, "Adapting the Tesseract Open Source OCR Engine for Multilingual OCR," *Proceedings of the International Workshop on Multilingual OCR MOCR '09,* pp. 1:1-1:8, 2009.

[11] J. I. James, "Digital Forensic Science," 2017. [Online]. Available: https://dfir.science/2017/04/tesseract-ocr-extract-text-from-images.html. [Accessed 31 July 2018].

[12] B. U. Language Technologies Unit (Canolfan Bedwyr), "An Overview of the Tesseract OCR (Optical Character Recognition) engine, and its possible enhnacement or use in Wales in pre-competitive research stage," 2008. [Online]. Available: https://tesseract-ocr.repairfaq.org/downloads/saltcymru_document5.pdf. [Accessed 31 July 2018].

[13] H. Leung, "Tesseract vs Google Vision, Round 1, Fight! The Text extraction Wars!," 24 July 2017. [Online]. Available: https://www.linkedin.com/pulse/tesseract-vs-google-vision-round-1-fight-text-extraction-hanley-leung. [Accessed 4 August 2018].

Probabilistic Image Quality Assessment: An Economics Point of View

Guangtao Zhai
Institute of Image Communication and Information Processing,
Shanghai Jiao Tong University, Shanghai, China
zhaiguangtao@sjtu.edu.cn

Xiaoping Zhang
Department of Electrical and Computer Engineering,
Ryerson University, Toronto, Canada
xzhang@ee.ryerson.ca

***Abstract*—We analyze the problem of image quality assessment from economics point of view with two main arguments: 1) Visual contents is a kind of product (good/service) to be utilized by consumers and therefore the image quality assessment process can be modeled using the tools from economics, e.g the law of diminishing marginal utility/returns, the expected utility theory and the principle of risk aversion. 2) The subjective quality rating of a given visual signal is better depicted with a probability distribution than a deterministic value.**

***Index Terms*—Image quality assessment, Marginal utility, Diminishing returns, Expected Utility, Risk Aversion**

I. Introduction

In economics, the fundamental law of diminishing returns states that increasing a single factor of production, while holding other factors unchanged, will lead to decreased marginal (incremental) output of a production process. Nowadays, visual contents is a very common type of good (or service) which consumers are willing to pay. Deeming visual contents as a product, and its perceptual quality as the key factor, i.e. the utility, then we can arrive at the conclusion that keep increasing the perceptual quality, or Quality of Experience (QoE) for certain amount, will at some point bring back lower and lower returns. This observation is obviously true because higher QoE generally requires better equipments, more advanced processing techniques, bigger bandwidths, and therefore higher costs. But the return does not increase equally with the same amount of investment in improving the QoE. In other words, for service and content providers, the utility is a nonlinear function of visual quality. It is reasonable to assume there is a "minimum quality" requirement for the consumers to become willing to buy those visual contents for services. As the visual quality increases, higher price can be charged, but when quality reaches a certain level (e.g. Full HD), people will not spend extra for even better visual results (e.g. 4K HDR). And any risk aversion company will try to maximizie the "minimum quality" for higher profit.

We can also analyze the process of visual content consumption from the consumer point of view using the concept of marginal utility. Utility is an important concept in economics which is defined as the satisfaction for benefit derived from consuming a product, here the visual contents. So marginal utility is the increase of satisfaction or benefit gained from watching the visual contents at certain quality level. Gossen's first law, or the law of diminishing marginal utility, suggests that the amount of satisfaction of enjoying the visual content also decreases with the same level of visual quality improvement. The law of diminishing marginal utility is at the heart of the explanation of numerous economic phenomena. The "marginal decision rule" states that rational people consume the product at a quantity when the marginal utility is equal to the marginal cost. And this means that economically, for visual contents, people will not always pursue higher quality but prefer higher cost-effectiveness. The law of diminishing marginal utility is similar to the law of diminishing returns but with focus on the behavior of individuals.

The law of diminishing marginal utility is thought of being irrefutably true in nature, following the axiom of human action [1], and also finds its root in psychology, biology and even neuroscience [2]. A research discipline so called neuro-ecomonics is considered as a consilience of micro-level neuroscience and social-level human decision making theory [3]. Interestingly, recent research perceptual quality assessment also resort to biology and neuroscience for motivation, inspirations and explanations [4]–[6]. Also, tools that have been widely used in economics, e.g. catatrophic models, has been introduced to image quality assessment to explain the nonmonotonicity between subjective and objective ratings [7].

In mainstream economics, utility is generally assumed as quantifiable, that is, it can be measured and calculated. In economics and game theory, the most popular approach to the choice under uncertainty is the *expected utility* theory which is also a general solution to decision making under risk. Similarly, image quality assessment also aims to computed an estimate of perceptual quality of visual media. So for a product of visual contents, we can approximately equivalent its utility to perceptual quality. Then current image quality assessment approaches can be utilized for analysis from economic point of view. On the other hand, the research of utility can also lend some insights into the problem of perceptual quality assessment. Generally, the concept of utility in economics concerns the collective behavior of a group of people and therefore it is more appropriate to deem perceptual quality of given visual content as probabilistic rather than a single deterministic value, as practiced in current studies.

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

II. Analysis of Utility and Risk Functions

A. U_{con}: utility for consumers

As discussed, for the task perceptual quality assessment of visual products, the Mean Opinion Score (MOS) is equivalent of utility. Luckily, there exist a large number of image quality databases containing images with different type of distortions as well as their subjective quality scores. On the other hand, the law of diminishing marginal utility depends on a proper measure of cost. For images corrupted with JPEG compression artifacts, we can reasonably define the average bitrate (in bit per pixel, or bpp) as a natural proxy of cost because higher bit rate requires larger storage space and/or transmission bandwidth.

Note that the saturation effect of subjective quality vs. quality metrics are widely noticed in existing research. However, typically, saturation occurs at both very good and very bad quality conditions and are explained with Just Noticeable Difference (JND) theory (for high quality images) and masking effects (for low quality images) and etc. But the argument here is that if we deem quality as utility, and appropriately define the term of cost, then the law of diminishing marginal utility perfectly explains the one-side saturation phenomena in perceptual quality assessment.

B. U_{pro}: utility for providers

Now we continue our analysis of utility from the view points of content/service providers. As we have mentioned in the introductory section, higher visual quality does not always guarantee better return. And the providers generally tend to maximize the minimum quality rating from the consumers for more subscription. In other words, utility for content/service providers is a nonlinear function of quality.

If we write a provider's utility function as $U_{pro}(Q)$ then we can have the following two assumptions: First, the marginal utility is positive, i.e. $U'_{pro} > 0$. Second, the provider is risk averse, which means that the provider dislikes decreasing expected utility, i.e. $E[U(Q+\epsilon)] < U(Q)$, with ϵ a zero mean random variable. The risk aversion behavior also follows from Jensen's inequality, if we assume that U_{pro} is concave, $E[U(Q+\epsilon)] < U[E(Q+\epsilon)] = U[Q]$. In fact, the equivalence between concave utility function and risk aversion is a well-known fact in economics.

We can now write the composed utility function as $U(C) = U_{pro}(Q) = U_{pro}(U_{con}(C))$. Taking the second derivative of $U(C)$, we have $U''(C) = U''_{pro}(U_{con}(C))U'_{con}(C)^2 + U'_{pro}(U_{con}(C))U''_{con}(C)$. As we have discussed, since U_{con} and U_{pro} are concave, i.e. $U''_{con} < 0$ and $U''_{pro} < 0$, and the marginal utility is positive, i.e. $U'_{pro} > 0$, so we have $U''(C) < 0$. This suggests that the composed utility function for the provider is also concave in terms of the cost C, and therefore the provider is risk aversion in term of the cost.

C. Risk and Utility

Since that service providers are generally risk averters, which means they dislike zero-mean risks, a natural way to measure the level of risk aversion is to know how much they want to pay to eliminate a zero-mean risk ϵ. This quantity is known as the risk premium, π, which is defined as $E[U(C+\epsilon)] = U(C-\pi)$. Usually the risk premium is a complex function of ϵ, Q and $U(\cdot)$. If we consider a small risk, and apply Taylor expansion to the both sides of the above equation, then $E[U(C+\epsilon)] \simeq E[U(C) + \epsilon U'(C) + \frac{\epsilon^2}{2}U''(C)] = U(C) + \frac{\sigma_\epsilon^2}{2}U''(C)$, and $E[U(C-\pi)] \simeq E[U(C) - \pi U'(C) = U(C) - \pi U'(C)$. Then we can have $\pi \simeq \sigma_\epsilon^2/2A(C)$, where $A(C) = -U''(C)/U'(C)$ is the Arrow-Pratt measure of absolute risk aversion.

Mean-variance ($\sigma - \mu$) analysis theory assumes that the decision of the provider can be described by a preference function $V(\mu,\sigma)$, over mean and variance of the utility, or return of the cost, with the assumptions that $\partial V(\mu,\sigma)/\partial\mu > 0$, $\partial V(\mu,\sigma)/\partial\sigma < 0$ can be interpreted as positive marginal return and risk aversion if we deem variance as a measure of risk. The $\sigma - \mu$ assumption will greatly simplify the analysis and is therefore widely adopted in economics. However, with the largely simplified preference function, the $\sigma - \mu$ analysis and the expected utility theory are not necessarily equivalent unless we restrict the utility function to be quadratic or limit the return distribution to be normal [8]. Luckily, those two restrictive conditions are not difficult to meet in practice for the problem of visual quality assessment. The central limit theorem states that as the number of independent random variable added their sum tends towards a normal distribution. As for of perceptual quality assessment, an ideal subjective test scenario would be that a large number of independent observers watch the visual stimuli and average their ratings. So we have reason to assume a normal distribution for perceptual quality scores. Since mean and variance are sufficient statistics for normal distribution, the $\sigma - \mu$ analysis is optimal.

Acknowledgment

G. Zhai thanks the support from National Natural Science Foundation of China under grants 61831015 and 61527804.

References

[1] L. v. Mises, *Human Action, 4th ed.,*. Fox Wilkes, San Francisco, 1996.

[2] E. Wilson, *Consilience*. Knopf, New York, 1998.

[3] P. W. Glimcher and E. Fehr, Eds., *Neuroeconomics, Second Edition: Decision Making and the Brain*. Academic Press, 2013.

[4] G. Zhai, X. Min, and N. Liu, "Free-energy principle inspired visual quality assessment: An overview," *Digital Signal Processing*, vol. 91, pp. 11 – 20, 2019, quality Perception of Advanced Multimedia Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S105120041930020X

[5] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free energy principle," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 41–52, 2012.

[6] G. Zhai, "Recent advances in image quality assessment," in *Visual Signal Quality Assessment*. Springer International Publishing, 2015, pp. 73–97.

[7] G. Zhai and X. Wu, "On monotonicity of image quality metrics," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 1157–1160.

[8] J. W. Pratt, "Risk aversion in the small and in the large," *Econometrica*, vol. 32, no. 1-2, pp. 122–136, 1964.

Generative Adversarial Network-based Image Super-Resolution with a Novel Quality Loss

Xining Zhu, Lin Zhang*, Lijun Zhang, Xiao Liu, Ying Shen, and Shengjie Zhao

School of Software Engineering, Tongji University, Shanghai, China
Emails:{1731559, cslinzhang, 1632761, 1532787, yingshen, shengjiezhao}@tongji.edu.cn

***Abstract*—Single Image Super-Resolution (SISR) has been a very attractive research topic in recent years. Breakthroughs in SISR have been achieved due to deep learning and Generative Adversarial Networks (GANs). However, the generated image still suffers from undesired artifacts. In this paper, we propose a new method named GMGAN for SISR tasks. In this method, to generate images more in line with Human Vision System (HVS), we design a quality loss by integrating an IQA metric named Gradient Magnitude Similarity Deviation (GMSD). To our knowledge, it is the first time to truly integrate an IQA metric into SISR. Moreover, to overcome the instability of the original GAN, we use a variation of GANs named WGAN-GP. Experiments show that GMGAN with quality loss and WGAN-GP can generate visually appealing results and set a new state-of-art.**

***Index Terms*—Single Image Super-Resolution, Image Quality Assessment, Generative Adversarial Network**

I. Introduction

Single Image Super-Resolution (SISR) aims to recover a high-resolution image from a single low-resolution one. Despite the tremendous progress in SISR due to the fast development of deep learning and Generative Adversarial Networks (GANs) [1], there is still one problem: these learning-based methods commonly use per-pixel based loss function MSE to guide the training. Although MSE is easy to calculate and has clear physical meaning, it is connected poorly with Human Vision System (HVS). Thus it is urgent to find an alternative more in line with HVS to guide the training. Besides, learning-based methods facilitated with GANs suffer from training instability.

In this paper, to solve the above-mentioned problems, we propose a new method named Gradient Map Generative Adversarial Network (GMGAN). To overcome the flaw of MSE, we design a quality loss term from an IQA metric named Gradient Magnitude Similarity Deviation (GMSD) [2]. To stabilize the training, we use a variation of GANs named WGAN-GP [3].

II. GMGAN: A Generative Adversarial Network with a Novel Quality Loss

The architecture of the discriminator remains the same as SRGAN [4]. Compared with SRGAN, there are three main improvements of the generator. (1) To combine both advantages of the multi-level residual network and dense connections, we replace the original residual block with Residual-in-Residual Dense Block (RRDB) [5]. (2) To overcome the flaw of GAN, we replace it with WGAN-GP. (3) As suggested in EDSR, we remove BN layers to reduce computational complexity and memory usage. Moreover, it helps improving generalization ability.

Loss function is the optimization objective for learning-based SISR methods. In general, the loss function for the generator G is:

$$l_G = l_{MSE} + l_P + l_Q + l_{GA} \tag{1}$$

The loss function for the discriminator D is:

$$l_D = l_{DA} \tag{2}$$

Here, we introduce MSE loss l_{MSE} and perceptual loss l_P similar to SRGAN. Besides, we replace the original adversarial loss with its improved one l_A and propose the novel quality loss l_Q.

The quality loss l_Q is inspired by an FR-IQA metric named Gradient Magnitude Similarity Distortion (GMSD) [2]. Note that GMSD has a meaningful derivative, while many others do not, hence we cherry-pick it to form a loss term for training the network.

The calculation of GMSD is as follows: first the Prewitt filters along horizontal and vertical directions are defined as h_x and h_y. Then they are convolved with the generated image I_{SR} and the ground truth I_{HR}.

At location i, the gradient magnitudes of I_{SR} and I_{HR}, denoted as $m_{SR}(i)$ and $m_{HR}(i)$, are computed as follows:

$$m_{SR}(i) = \sqrt{(I_{SR} \otimes h_x)^2(i) + (I_{SR} \otimes h_y)^2(i)} \tag{3}$$

$$m_{HR}(i) = \sqrt{(I_{HR} \otimes h_x)^2(i) + (I_{HR} \otimes h_y)^2(i)} \tag{4}$$

where symbol "$\otimes$" denotes the convolution operation. To acquire the Local Quality Map (LQM) of I_{SR} and to reflect the local quality of each small patch in I_{SR} in pixel-wise manner, the gradient magnitude similarity (GMS) map is computed as follows:

$$GMS(i) = \frac{2m_{SR}(i) \cdot m_{HR}(i) + c}{m_{SR}^2(i) + m_{HR}^2(i) + c} \tag{5}$$

*Corresponding author: cslinzhang@tongji.edu.cn

978-1-7281-3038-5/19/$31.00 ©2019 IEEE

TABLE I
PUBLIC BENCHMARK TEST RESULTS (PI/PSNR(DB)). RED INDICATES THE BEST AND BLUE INDICATES THE SECOND BEST PERFORMANCE IN TERMS OF PI. ALL COMPARISON RESULTS ARE ACQUIRED FROM PUBLISHED RESULTS.

Datasets	Bicubic	LapSRN	EDSR+	EnhanceNet-E	EnhanceNet-PAT	SRGAN	GMGAN
Set5	7.33/28.42	6.48/31.54	5.99/32.62	6.05/31.74	2.93/28.56	3.35/32.05	3.25/30.02
Set14	6.97/26.10	5.96/28.19	5.50/28.94	5.25/28.42	3.02/25.77	2.88/28.49	2.77/26.37
BSD100	6.94/25.96	5.81/27.32	5.39/27.79	5.49/27.50	2.91/24.93	2.35/27.58	2.29/25.46

The positive constant c is used to keep the numerical stability. To estimate the image overall quality of the LQM, an average pooling strategy is applied to the GMS map as follows:

$$GMSM = \frac{1}{N}\sum_{i=1}^{N} GMS(i) \tag{6}$$

where N refers to the total number of pixels in I_{SR}. GMSM is short for Gradient Magnitude Similarity Mean. To reflect how the local quality degradation varies, the standard deviation of the GMS map is computed as follows:

$$GMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(GMS(i) - GMSM)^2} \tag{7}$$

The higher the GMSD score, the larger the distortion of I_{SR} and the poorer the image quality. Finally, the quality loss is defined as:

$$l_Q = GMSD((G_\theta(I_{LR})), (I_{HR})) \tag{8}$$

where $GMSD$ represents the whole calculation process of GMSD.

III. EXPERIMENTS

A. Training details

All the methods involved in our experiments were implemented with Pytorch. The experiments were conducted on a ubuntu workstation with a NVIDIA Titan Xp GPU.

As for datasets, DF2K dataset is used for training, while Set5, Set14 and BSD100 are used for evaluation. DF2K dataset is short for DIVIK + Flickr2K, consisting of 3450 high-quality images in total. Set5, Set14 and BSD100 are three public benchmark datasets.

B. Benchmark results

In this section, GMGAN is compared not only with PSNR-oriented methods, including LapSRN, EDSR+ and EnhanceNet-E, but also with perceptual-driven methods, including SRGAN and EnhanceNet-PAT.

Traditional IQA criteria like PSNR and SSIM can only reflect part of human perception, thus we attempt to find an alternative for them: PI, which combines two no-reference image quality metrics Ma and NIQE. In this setting, a lower PI indicates a better perceptual quality. PSNR is also provided for distortion reference.

Quantitive evaluation results of GMGAN on public benchmark datasets are provided in Table I. It can be observed that GMGAN performs the best or the second best, which indicates GMAGN can obtain comparatively the best perceptual quality among these methods.

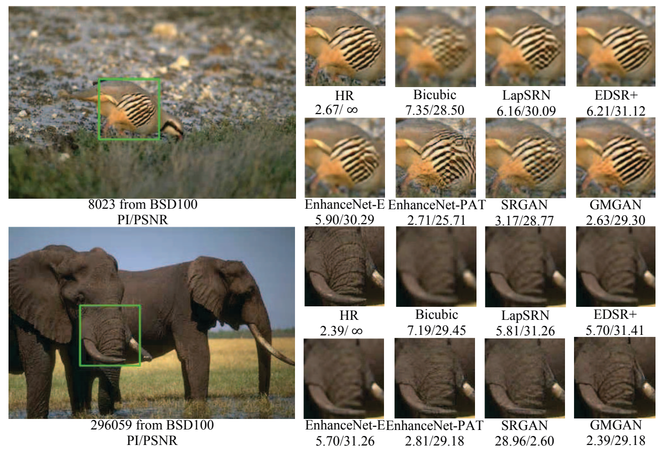


Fig. 1. Visual comparison for 4× SR on BSD100. GMGAN can generate photo-realistic details with less undesired artifacts.

Besides quantitive evaluation, qualitative results are also provided in Fig. 1. It can be inferred that GMGAN outperforms previous methods, reflecting in generating more natural textures and suffering from less undesired artifacts.

IV. CONCLUSIONS

In this paper, to address the problem of generating super-resolution images with more realistic textures and less unpleasant artifacts, we design a new method named GMGAN for SISR. In GMGAN, we integrate an IQA metric named GMSD into the loss function and replace GAN with WGAN-GP to stabilize the training. Experiments show that GMGAN can produce images with more natural textures and suffer from less unpleasant artifacts.

REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *NIPS*, 2014, pp. 2672–2680.

[2] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, 2014.

[3] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *NIPS*, 2017, pp. 5767–5777.

[4] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *CVPR*, 2017, pp. 105–114.

[5] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, "Esrgan: Enhanced super-resolution generative adversarial networks," in *ECCV*, 2018, pp. 63–79.

A Heart Rate Monitoring and Activities Recognition System for Badminton Training

Yu-Wen Huang, Shu-Hua Chang, Yu-Chih Tsai and Yuan-Hsiang Lin*
Department of Electronic and Computer Engineering, Taiwan Building Technology Center
National Taiwan University of Science and Technology
Taipei, Taiwan
Email: linyh@mail.ntust.edu.tw

***Abstract*--In this study, we developed an activities recognition system with a heart rate measurement (HRM) module for badminton training. Combining machine learning, data analysis, mobile computing, and wearable devices technology, we expect that through modern technologies, the specific badminton training efficiency can be increased and the performance in the competitions can be further improved.**

I. Introduction

The analysis of badminton movement is roughly divided into image processing or inertial sensors. For example, M. S. Salim et al. used image processing to analyze the movement trajectory [1] and the movement variety. The application of inertial sensors, such as the literature [2]–[3], require to install the sensors on the player or the racket, through the motion change amount to measure the joint angle and recognize activities. However, the current methods are still inconvenient to use and lack of high accuracy. And there is no analysis of integrated heart rhythm signals. Therefore, in this study, we use the accelerometer and gyroscope in inertial sensors as the motion sensing device, and use sequential minimal optimization (SMO) [4] as the classifier. Through the software developed and Weka data exploration [5], to recognize the activities of badminton players. We also integrated the HRM module into this system, which can provide the user a scientific and quantitative reference for training planning.

II. Methods

A. System architecture

The system architecture is as shown in Fig. 1, which consists of HRM module, motion sensing device and the data analyzing unit. The motion sensing device will receive the heart rate data from the HRM module, capture the data of the inertial sensor and detect the striking spots, and the data will be transmitted to the data analyzing unit (a computer or a smart phone) via Bluetooth transmission to be processed, including signal processing and real-time activities recognition.

B. Heart Rate Measurement Module

A soft heart rate band with the HRM module is wearing on the user's chest to measure the ECG signal as an index of the exercise intensity. Thus the user can understand his/her exercise intensity and further adjust the exercise efficiency. The hardware design and algorithms implement are shown in our previous work [6].

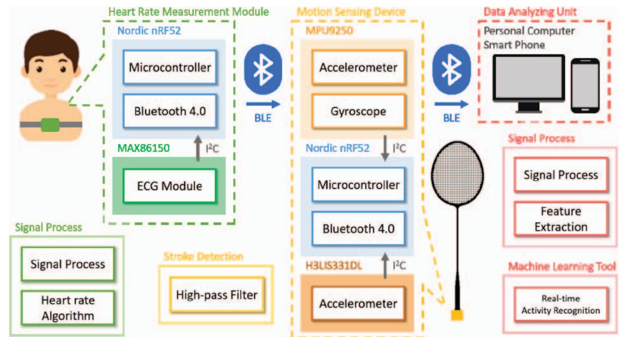


Fig. 1. System architecture.

In this study, we use MAX86150 as the heart rate monitor sensor module and the Nordic nRF52 as our microcontroller, which is embedded Bluetooth Low Energy (BLE) interface.

C. Motion sensing device

The motion sensing device is installed at the bottom of the racket handle. The accelerometer and gyroscope are used as the motion sensors. We use the Nordic nRF52 as our microcontroller. The striking spots data, sensors data, as well as the heart rate data from the HRM module will be transmitted to the data analyzing unit via BLE.

D. Data analyzing unit

The interface of the smart phone is divided into 3 parts: heart rate monitoring, data displaying, and activities recognition. We use the SMO algorithm to recognize 7 stroke activities including cut, drive, lob, long, net play, rush, and smash to help the users learn his/her own activities and further adjust the striking postures. In the smart phone version, the smart phone is off-line at first and a personalized model is built via the personal computer. The training data is written in the format of the ARFF file. Using the Java function provided by Weka to read the file, the classifier model can be built. After saving the model to the smart phone end, the function of activities recognition can be performed.

III. Results

A. Heart rate measurement

The subject was asked to wear the commercial heart rate band (Polar H7) and the HRM module developed in this study at the same time. The measuring took 30 seconds, and the data was compared every 2 seconds. There are 5 subjects in total, and the mean absolute error is 1.21 bpm.

TABLE I
ACTIVITIES RECOGNITION RESULTS

True \ Classifier	Cut	Drive	Lob	Long	Netplay	Rush	Smash
Cut	20	0	0	0	0	0	0
Drive	0	15	2	3	0	0	0
Lob	0	0	20	0	0	0	0
Long	0	1	0	17	0	0	2
Netplay	0	0	0	0	20	0	0
Rush	0	0	0	0	0	20	0
Smash	0	0	0	4	0	0	16
The Average Accuracy: 91.42%							

B. *Stroke detection and activities recognition*

In order to verify the accuracy of the stroke detection and the activities recognition algorithms. The subject was asked to do 7 stroke activities; each activity has 20 strokes. Therefore, 140 strokes are made in total. The average accuracy is 100% for stroke detection. The average accuracy of activities recognition is 91.42% and shown in TABLE I.

C. *Data record for a real practice*

The subject was asked to practice for ten minutes wearing the HRM module and the motion sensing device at the same time. As the heart rate change shown in Fig. 2, the subject's heart rate climbed from 75 bpm to 130 bpm after 5 minutes' training, ranked as medium exercise intensity in ages under 25, and the heart rate maintained at about 130 bpm after 5 minutes.

Combining the data of Fig. 3, it is inferred that the subject finished the initial warming up after about 60 strokes. As ordinary people, enough exercise effect can be achieved if the user keeps the same exercise intensity. The user can speed up the training tempo to fit the competition intensity accordingly.

As the stroke activities detection shown in Fig. 4, the subject did less smashes and rushes while hit more lobs and clears to the backcourt and tended to be a defensive player. Combining the previous two data, the user can judge if the current tactic consumes too much physical strength or fights with the other side for too many strokes.

IV. CONCLUSION

In this study, we developed an activities recognition system with a HRM module for badminton training. This system can detect the strokes and recognize the stroke activities in real time, providing the user a scientific and quantitative reference for training planning or tactical use during the game. Instant heart rate monitoring can be used as an indicator of exercise intensity and let the users adjust the training intensity according to the current heart rate. It can avoid overtraining and improve the efficiency of the badminton training to improve the performance of the game.

ACKNOWLEDGMENT

This work was supported in part by a research grant from the Ministry of Science and Technology under Grant MOST 107-2627-H-009-001-, MOST 108-2627-H-009-001-, and financially supported by the Taiwan Building Technology Center from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan.

REFERENCES

[1] M. S. Salim, H. N. Lim, M. S. M. Salim and M. Y. Baharuddin, "Motion analysis of arm movement during badminton smash," in *Proc. Biomedical Engineering and Sciences, 2010 IEEE EMBS Conference*, Dec. 2010, pp. 111–114.

[2] Z. Wang, M. Guo and C. Zhao, "Badminton stroke recognition based on body sensor networks," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 5, pp. 769-775, Oct. 2016.

[3] M. I. Rusydi, M. Sasaki, M. H. Sucipto, Zaini, and N. Windasari, "Local euler angle pattern recognition for smash and backhand in badminton based on arm position," in *Proc. Applied Human Factors and Ergonomics and the Affiliated Conferences, 6th International Conference*, Jul. 2015, pp. 898–903.

[4] J. C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," *Technical Report MSR-TR-98-14, Microsoft Research*, Apr. 1998.

[5] I. H. Witten, E. Frank, and M. A. Hall, "Data Mining: Practical Machine Learning Tools and Techniques," in 2011.

[6] E. A. P. J. Prawiro, C. I. Yeh, N. K. Chou, M. W. Lee, and Y. H. Lin, "Integrated wearable system for monitoring heart rate and step during physical activity," *Mobile Information Systems*, vol. 2016, pp. 1–10, Apr. 2016.

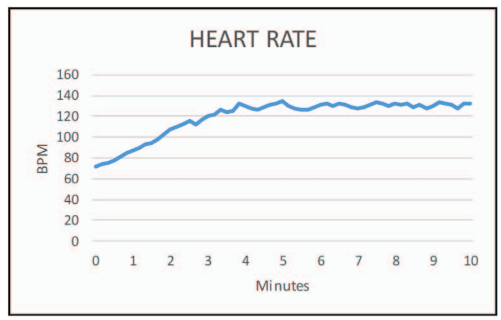


Fig. 2. Heart rate record for 10 minutes' practice.

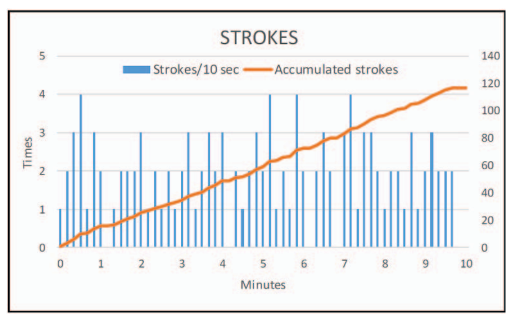


Fig. 3. Strokes/10sec and accumulated strokes for 10 minutes' practice.

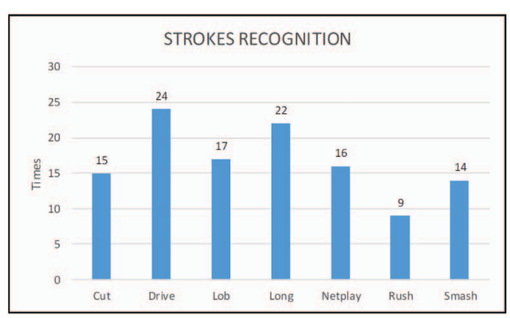


Fig. 4. Stroke activities statistics for 10 minutes' practice.

Author Index *sorted by last name

A

B

C

G

L

M

O

Q

R

S

U

W

Z