

Spontaneous Facial Micro-expression Analysis using Spatiotemporal Local Radon-based Binary Pattern

Xiaohua Huang

Center for Machine Vision and Signal Analysis
University of Oulu, Finland
Email: xiaohua.huang@oulu.fi

Guoying Zhao*

1. School of Information and Technology
Northwest University, Xi'an, China
2. Center for Machine Vision and Signal Analysis
University of Oulu, Finland
Email: Email: guoying.zhao@oulu.fi

Abstract—Micro-expressions are difficult to be observed by human beings due to its low intensity and short duration. Recently, several works have been developed to resolve the problems of micro-expression recognition caused by subtle intensity and short duration. One of them, Local binary pattern from three orthogonal planes (LBP-TOP) is primarily used to recognize micro-expression from the video recorded by high-speed camera. Several variances of LBP-TOP have also been developed to promisingly improve the performance of LBP-TOP for micro-expression recognition. However, these variances of LBP-TOP including LBP-TOP cannot well extract the subtle movements of micro-expression so that they have the low performance. In this paper, we propose spontaneous local radon-based binary pattern to analyze micro-expressions with subtle facial movements. Firstly, it extracts the sparse information by using robust principal component analysis since micro-expression data are sparse in both temporal and spatial domains caused by short duration and low intensity. These sparse information can provide much motion information to dynamic feature descriptor. Furthermore, it employs radon transform to obtain the shape features from three orthogonal planes, as radon transform is robustness to the same histogram distribution of two images. Finally, one-dimensional LBP is employed in these shape features for constructing the spatiotemporal features for micro-expression video. Intensive experiments are conducted on two available published micro-expression databases including SMIC and CASME2 databases for evaluating the performance of the proposed method. Experimental results demonstrate that the proposed method achieves promising performance in micro-expression recognition.

I. INTRODUCTION

Facial expressions of emotion, especially those known as micro-expressions, amongst nonverbal behaviors like gestures and voice, have received increasing attention in recent years [1]. A micro-expression is subtle and involuntary facial expression commonly occurring in high-stakes situations, where people have something to lose or gain. The importance of micro-expression study has become apparent in many potential applications, such as in the security field. However, micro-expression are quite distinguish from regular facial expression,

since the micro-expression generally remain less than 0.2 seconds and are very subtle since people are trying to control and repress [2], [3]. Comparing with regular facial expression, they are very hard for people to detect and recognize. In order to improve the ability of people to recognize micro-expressions, researchers in psychology made contributions to train people to better recognize micro-expressions using training tools. However, even with these training tools, the ability of people reading micro-expression achieves only around 40% in recognizing micro-expressions [4]. On the other hand, manual reading micro-expression takes people lots of time to analyze micro-expressions. Therefore, there is a great need for a high-quality micro-expression recognition system on assisting people to accurately recognize them.

As spontaneous micro-expressions can reveal genuine emotions in high-stake situations that people try to conceal their emotion, spontaneous facial micro-expression has been firstly and automatically examined by [5], [6]. As we know, there are many appearance features developed for facial expression recognition. Motivated by the research in facial expression recognition, researchers attempted to exploit the existed appearance features for micro-expression recognition. Amongst appearance-based facial expression representation, local binary pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings. As far as we know, the first automatic micro-expression recognition system is developed by [6], which is based on the LBP, due to its efficiency and performance in facial expression recognition [7]. The LBP-TOP is further implemented by [8], [9] to provide the baseline result in several micro-expression

* Corresponding author

This work was supported by the Academy of Finland, Tekes Fidioprogram, the strategic Funds of the University of Oulu, Finland, the Infotech Oulu. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

databases, such as the second version of Chinese Academy of Sciences Micro-expression (CASME2) database [9]. Since then, LBP and its variants have been employed as the feature descriptors for micro-expression recognition in many other studies. For example, Davison *et al.* [10] exploited LBP-TOP to investigate whether micro-facial movement sequences can be differentiated from neutral face sequences.

As observed in [8] [6] [9], the primary works only exploited the pixel-level in an image, resulting in the unstable of micro-expression recognition. Some works attempted to develop LBP-TOP on multi-layer architecture deriving from an image. In [11] Ruiz-Hernandez and Pietikäinen firstly exploited second order Gaussian jet, and then re-parametrize LBP-TOP on these jets. Wang *et al.* [12] proposed to use robust principal component analysis (RPCA) to extract the subtle motion as the high-level layer, in which micro-expression subtle change can be well described by RPCA. Additionally, based on this high-level layer, Local Spatiotemporal Directional Features is proposed to extract the spatiotemporal features for micro-expressions. Besides multi-layer methods, some researchers aim toward making LBP-TOP more efficient and robust. For example, recent work in [13] used six intersection points for LBP-TOP to reduce redundant information and achieve a promising performance on micro-expression recognition. As well, some static-image-based binary pattern is extent to spatiotemporal domain for micro-expression recognition. Centralized Binary Patterns is extend to temporal domain by using LBP-TOP framework for micro-expression recognition [14]. Riesz wavelet transform is implemented by Oh *et al.* [15] to obtain multi-scale monogenic wavelets for micro-expression recognition. Completed local quantized pattern [16], which exploited magnitude and orientation as additional source of LBP, has been shown a considerable result on texture classification. Its temporal extension is further proposed to extract the flexible and completed feature information for micro-expression recognition [17]. Another orientation related work is proposed by Li *et al.* [18] to improve LBP-TOP by combining LBP, histograms of oriented gradient (HOG) and histograms of image gradient (HIGO). Zhang *et al.* [19] extract the movement of micro-expression by combining optical flow orientation histogram and LBP-TOP.

It is observed that the above-mentioned approaches based on the framework of LBP-TOP mostly consider the dynamic texture information from three orthogonal planes to characterize the spatiotemporal feature of face image. However, these existed methods used in micro-expression recognition only consider texture information to represent face images, but ignore the shape information of face images. As suggested in [20], the shape information can provide useful for facial expressions complementary to appearance features. The work in [21] also demonstrated that the shape information can enhance the discriminative ability of LBP for distinguishing two different-class image but with the same LBP feature distribution. Therefore, it is useful to micro-expression recognition by exploiting shape information of micro-expression images for dynamic texture information.

Radon transform are classical methods for pattern analysis that is widely used in face tracking [22], [23], since they extract face shape properties from face images and increase discrimination of images. Houam *et al.* [21] exploited a simple but effective radon transform, namely integral projection, to enhance the discriminative ability of LBP. It firstly is invariant to a number of image transformations like scale and translation. It is also highly robust to white noise. Then it preserves the principle of locality of pixels and sufficient information in the process of projection. Recently, Huang *et al.* proposed to use integral projection to describe the shape of micro-expression and then combine LBP to boost the performance of micro-expression analysis [24], [25]. Their works demonstrated that image projection techniques can enhance the robustness and performance of LBP. Therefore, we propose a new spatiotemporal method based on radon transform to improve the performance of micro-expression recognition, in which radon transform can provide the shape property of facial images.

To explain the concepts of our approach, the paper is organized as follows. In Section II, we explain our method for exploring the new spatiotemporal features based on radon transform for micro-expressions. The results of applying our method for recognizing micro-expressions are provided in Section III. Finally we summarize our findings in Section IV.

II. PROPOSED METHODOLOGY

A. Extract subtle motion

Due to short duration and low intensity of micro-expression, the micro-expression data are sparse in both temporal and spatial domains [26]. For a frame of micro-expression \mathbf{I} , it can be decomposed into subject information \mathbf{S} and the subtle motion information \mathbf{M} , which is implicitly represented as $\mathbf{I} = \mathbf{S} + \mathbf{M}$. Robust principal component analysis (RPCA) was proposed to extract the sparse information and has been attracting much attention in face recognition [27], [28] and micro-expression [12], [25]. In this paper, we use RPCA to extract the sparse motion information \mathbf{M} of micro-expression. Based on these sparse motion information, we further exploit the spatiotemporal feature descriptor for micro-expression recognition.

B. Spatiotemporal local Radon-based binary pattern (STRBP)

The radon transform [29] is the projection of the image intensity along a radial line oriented at specific angle. It is based on the parametrization of straight lines and the evaluation of integrals of an image along these lines. Due to inherent properties of Radon transform, it is a useful tool to capture the directional features of an image. The radon transform has been extensively used to extract the local features in edge detection, texture and fingerprint classification and image retrieval in computer tomography. Mathematically, let $\mathbf{M}(x, y)$ be the

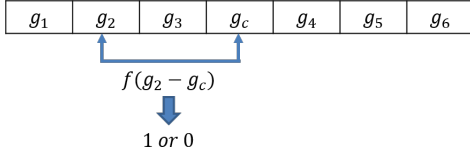


Fig. 1: Performing the binary label on the central element of a vector with 6 neighbors.

intensity of a pixel at location (x, y) in an subtle image \mathbf{M} , the random transformation of $\mathbf{M}(x, y)$ is defined as:

$$\mathbf{F}(\theta, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{M}(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy, \quad (1)$$

where δ is a Dirac function, θ is a projection angle and s is the threshold value.

Our empirical experience on [24] shows that radon transformation with θ and $\theta + \pi/2$ can better extract the subtle change of signals from two orthogonal directions. In this work, we focus on the above two cases. Besides sampling signal along Equation 1, the other signal can be collected by:

$$\mathbf{F}(\theta + \pi/2, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{M}(x, y) \delta(-x \sin \theta + y \cos \theta - s) dx dy, \quad (2)$$

where we empirically set s as 0.

The low-frequency components contribute to the global description, while the high-frequency components contribute to the fine details. As we know, facial expression features consists of the directional low-frequency components. It also means that signals in low spatial frequency band plays a dominant role in facial expression recognition. In essential, radon transform preserves the variations in pixel intensities. This process improves the spatial frequency components in a direction for facial expression. Therefore, while features are extracted using Radon transform, the variations in the spatial frequency are not only preserved but boosted also.

Radon projection preserves the directional low-frequency component, but they are insufficient to describe the micro-expression accurately, due to the low intensity of micro-expression. Amongst all appearance features, local binary pattern has been shown to perform a promising performance on facial expression recognition [30]. On the other hand, the work in [21] demonstrated that LBP enhanced by shape information can distinguish an image with different shape from those with the same LBP feature distributions. As its efficient computing, we implement the theoretical framework of LBP on radon transformation. It aims to derive from the classical descriptor used to characterize the low-intensity changes of micro-expression images.

Given $\mathbf{F}(\theta, s)$, we established a linear mask of size W , that sequentially scans the radon projection $\mathbf{F}(\theta, s)$ with one element step. W is an odd number ensuring the symmetric neighborhood of each element, such as $W = 7$. As exemplified in Fig. 1, for each signal point g_c , the number of its involved neighbors g_w is $W - 1$. Furthermore, each code is calculated by thresholding the neighborhood values against the central

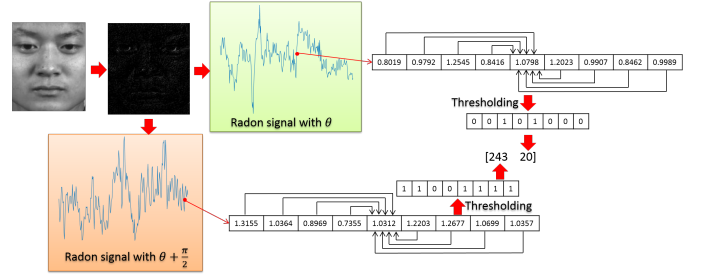


Fig. 2: Procedure of calculating the binary vector from two radon signals for micro-expression image.

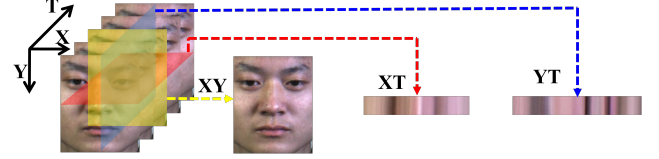


Fig. 3: An example that shows how to get the three orthogonal planes.

element. The neighbors will be assigned the value 1 if they are greater than or equal to the current element and 0 otherwise. In specific, the comparative function is defined as followed:

$$f_w = \begin{cases} 1 & \text{if } g_w - g_c \geq 0 \\ 0 & \text{if } g_w - g_c < 0, \end{cases} \quad (3)$$

where $w = 1, \dots, W$. The each binary vector for $\mathbf{F}(\theta, s)$ is represented as $\mathbf{H} = [f_1, f_2, \dots, f_W]$.

Basically, for (x, y) , we obtain the two binary vectors \mathbf{H}_1 and \mathbf{H}_2 for $\mathbf{F}(\theta, 0)$ and $\mathbf{F}(\theta + \pi/2, 0)$, respectively. Given \mathbf{H}_1 or \mathbf{H}_2 , its binary element of the resulting vector is multiplied by a weight depending on its position. This can be summarized as

$$h_W = \sum_w f_w 2^w, \quad (4)$$

where the result of \mathbf{H}_1 or \mathbf{H}_2 is a decimal. The histogram $\bar{\mathbf{H}}_1$ or $\bar{\mathbf{H}}_2$ is finally computed, in which it represents the frequency of each number occurring. We can concatenated the histograms $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ into one feature vector $\bar{\mathbf{H}} = [\bar{\mathbf{H}}_1 \bar{\mathbf{H}}_2]$ to describe an image. The example is illustrated in Fig. 2.

An operator based on co-occurrences of local binary patterns on three orthogonal planes (LBP-TOP) was introduced in many applications such as dynamic texture classification and facial expression recognition [7]. It considers three orthogonal planes of a video and concatenates LBP co-occurrence statistics in these three directions. Specifically, the XY plane appearance information, while the XT plane gives a visual impression of one row changing in time and YT describes the motion of one column in temporal space. The procedure of getting three orthogonal planes is shown in Fig. 3. Following LBP-TOP, we furthermore extend Radon-based binary pattern into three orthogonal planes of video sequence. For each plane, we firstly use radon transformation to obtain the radon

projections along θ and $\theta + \pi/2$ by using Equations 1 and 2. Furthermore, we implemented the comparative function to get the binary codes over θ and $\theta + \pi/2$. Finally, we concatenate the histograms from three orthogonal planes, namely \bar{H}_{XY} , \bar{H}_{XT} , \bar{H}_{YT} , respectively, into one feature vector \bar{G} . In such a representation, a micro-expression video is encoded by an appearance (XY-LBP) and two spatial temporal (XT-LBP and YT-LBP) co-occurrence statistics.

III. EXPERIMENTS

A. Database description and experimental setup

For evaluating STRBP, we conduct the experiments on SMIC [8] and CASME2 [9] databases for micro-expression recognition.

The SMIC database consists of 16 subjects with 164 spontaneous micro-expressions, which were recorded in a controlled scenario using 100 fps camera with resolution of 640×480 . 164 spontaneous facial micro-expressions are categorized into positive (51 samples), negative (70 samples) and surprise (43 samples) classes.

The CASME2 database includes 247 spontaneous facial micro-expressions recorded by a 200 fps camera and spatial resolution with 640×480 pixel size. In this database, they elicited participants' facial expressions in a well-controlled laboratory environment and proper illumination. These samples are coded with the onset and offset frames, as well as tagged with AUs and emotion. There are 5 classes of the micro-expressions in this database: happiness (32 samples), surprise (25 samples), disgust (64 samples), repression (27 samples) and others (99 samples).

We firstly use active shape model (ASM) [31] to obtain the 68 facial landmarks for facial images of video sequence, and align them to a canonical frame. We implement RPCA to obtain the subtle information of micro-expressions for SMIC and CASME2 databases. Furthermore, we employ the temporal interpolation method (TIM) [32] to normalize these subtle information into the same frames length (10 frames). For SMIC database, we crop facial images into 170×139 and divide them into 4×7 blocks. For CASME2 database, the face images are cropped to 308×257 pixel size, and are divided into 6×1 blocks. We used leave-one-subject-out cross validation protocol in the experiments, where the videos from one subject are used for testing, the rest for training. We use Chi-Square Kernel based support vector machine (SVM) [33] for the classification.

B. Parameter Evaluation

As previously mentioned in Section II, it is seen that the projection angle θ of radon transform and the size W of linear mask are two key parameters determining the performance of STRBP. In the following, we will discuss the influence of two important parameters θ and W on two databases. In this scenario, we firstly examine the recognition rate affected by θ , where θ varies from 0° to 90° with an interval of 15° . Then, we evaluate the performance of W across the sets of four linear masks that have the size of 3, 5, 7 and 9, respectively.

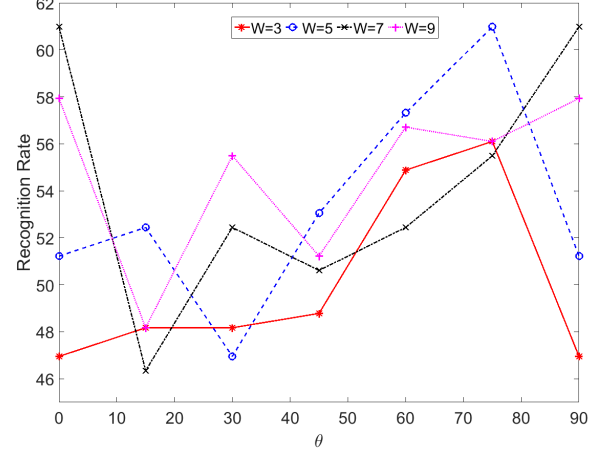


Fig. 4: Parameter influence of W and θ on SMIC database.

Figs. 4 and 5 show the influence of θ and W to STRBP on two micro-expression databases. As seen from both figures, in the case of $W = 3$ and $W = 5$, θ can lead to the increasing performance when it is increased. However, when W becomes bigger, the increasing θ cannot make the significant result comparing with the result using $\theta = 0$. It is seen that STRBP achieves the best result when θ equals to 0 with the bigger W . It demonstrates that the signals sampled along horizontal and vertical directions can provide the most information to describing the structure of micro-expression images.

Furthermore, we remain θ as 0° . For SMIC2, the accuracies of STRBP are 46.95%, 51.22%, 60.98% and 57.93% for $W = 3, 5, 7, 9$, respectively. For CASME2, we obtained the accuracies of STRBP as 50.2%, 56.68%, 64.37%, and 52.23% for $W = 3, 5, 7, 9$, respectively. It is interesting to see that the largest mask size ($W = 9$) cannot give the best results on two databases. It may be explained by that the largest size may provide redundant information and noise to one-dimensional local binary pattern. In contrast, when W is 7, STRBP obtains the promising results of 60.98% and 64.37% for SMIC and CASME2 databases, respectively.

C. Algorithm comparison

1) *SMIC*: For SMIC database, we compare STLBP-IP with the commonly used spatiotemporal features [7], [13] and the state-of-the-art methods on SMIC database. In our implementation on [7], [13], we used temporal interpolation method (TIM) to normalize each video into 10 frames. We employ to use spatiotemporal feature descriptor (LBP-TOP and LBP-SIP), to 4×7 facial blocks. For convenience, we name them as TIM10+LBP-TOP and TIM10+LBP-SIP, respectively.

The comparison results are reported in Table I. From Table I, it is seen that STRBP increases the accuracy of 12.2% for micro-expression recognition. Although STLBP-IP used radon transform with $\theta = 0$ and LBP on temporal domain for micro-expression recognition, STRBP obtains better performance than STLBP-IP. As well, STRBP is more simple and efficient than STLBP-IP. Amongst the comparative algorithms, it is

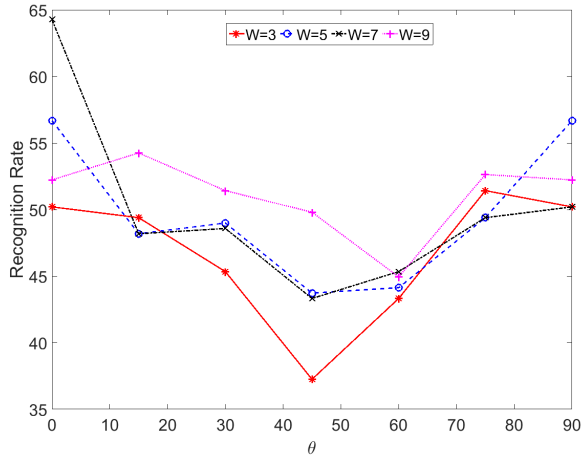


Fig. 5: Parameter influence of W and θ on CASME2 database.

observed that HIGO-TOP is the best one. But our proposed method STRBP outperforms to HIGO-TOP. On the other hand, HIGH-TOP has many controlled parameters while STRBP depends on two ones.

TABLE I: Performance comparison with the state-of-the-art methods on SMIC database.

Methods	Recognition rate (%)
Baseline [8]	48.78
TIM10+LBP-TOP [7]	48.17
TIM10+LBP-SIP [13]	44.51
HOG-TOP [18]	57.93
HIGO-TOP [18]	59.15
OSW-LBP-TOP [34]	53.05
OS+Wiener filter [35]	53.56
STLBP-IP [24]	57.93
STRBP	60.98

Finally, we present the confusion matrix for LBP-TOP+SVM(linear), STLBP-IP and STRBP in Fig. 6. From Fig. 6, it is found that STRBP improves the accuracies of each class comparing with LBP-TOP. As well, it also boosts the ability of recognizing negative micro-expressions comparing with STLBP-IP.

2) *CASME2 dataset*: We conduct the comparative experiments of STRBP with the baseline algorithm [9], LBP-TOP [7], LBP-SIP [13] and three new spatiotemporal feature descriptors [17], [18], [25]. According to experimental setup of [9], we set the radius as 3 for LBP-TOP and divide facial images into 5×5 blocks. For classification, we use linear-kernel based SVM [33]. For convenience, we name this method as Baseline. Secondly, we implement the framework of [8] based on LBP-TOP [7] and LBP-SIP [13] as a comparison. Features are extracted on 8×8 facial blocks.

Table II shows the comparative results of STRBP with the state-of-the-art spatiotemporal feature descriptors, where the results of STCLQP [17], HIGO-TOP [18], HOG-TOP [18] and STLBP-RIP [25] are directly extracted from their works. It is

	Negative	Positive	Surprise
Negative	42.86	30.00	27.14
Positive	29.41	58.82	11.76
Surprise	30.23	11.63	58.14

(a)

	Negative	Positive	Surprise
Negative	57.14	28.57	14.29
Positive	31.37	62.75	5.88
Surprise	34.88	11.63	53.48

(b)

	Negative	Positive	Surprise
Negative	65.71	21.43	12.86
Positive	35.29	60.78	3.92
Surprise	37.21	9.30	53.49

(c)

Fig. 6: The confusion matrix of (a) LBP-TOP+SVM(linear) [8], (b) STLBP-IP [24] and (c) STRBP for three micro-expression categorizations on SMIC database.

observed that STRBP significantly outperforms the LBP-TOP. Comparing with LBP-SIP, STRBP increases the recognition rate by 24.29%, because STRBP preserves the shape information of micro-expression images by using radon transform. Amongst the state-of-the-art algorithms, it is seen that STLBP-RIP achieves the considerable recognition rate of 62.75%. Although STLBP-RIP and STRBP used RPCA as the first step to extract the subtle information, STRBP obtains the better results than STLBP-RIP. Different from STLBP-RIP, STRBP directly extracted spatial and temporal information from radon

signals while STLBP-RIP used LBP as the post-preprocessing step for micro-expression videos. It also demonstrates that STRBP has a efficient computation but effective performance.

TABLE II: Comparison under micro-expression recognition rate on CASME2 database. * means that we directly took the results from their works.

Methods	Block Number	Recognition rate (%)
Baseline [9]	5×5	38.87
LBP-TOP [7]	8×8	39.68
LBP-SIP [13]	8×8	40.08
STCLQP* [17]	8×8	58.39
HIGO-TOP* [18]	8×8	55.87
HOG-TOP* [18]	8×8	57.49
STLBP-IP [24]	8×9	59.51
STLBP-RIP [25]	6×1	62.75
STRBP	6×1	64.37

IV. CONCLUSION

This paper have demonstrated that the spatiotemporal local Radon-based binary pattern (STRBP) obtains the considerable performance on two available facial micro-expression databases. Specifically, the simple signal extraction method based on radon transform preserves the low-frequency components of micro-expression images while the similar way of one dimensional LBP can further enhance these components by using texture information. Additionally, we observed that radon transform based on horizontal and vertical direction can well characterize the structure of micro-expression images. As a result, through the extensive experiment analysis, we obtain a simple but effective spatiotemporal feature descriptor for micro-expression recognition.

REFERENCES

- [1] G. Warren, E. Schertler, and P. Bull, "Detecting deception from emotional and unemotional cues," *Journal of Nonverbal Behavior*, vol. 33, no. 1, pp. 59–69, 2009.
- [2] P. Ekman and W. Friesen, "Detecting deception from emotional and unemotional cues," *Psychiatry*, vol. 32, no. 1, pp. 88–106, 1969.
- [3] X. Shen, Q. Wu, and X. Fu, "Effects of the duration of expressions on the recognition of microexpressions," *Journal of Zhejiang University Science B*, vol. 3, no. 13, pp. 221–230, 2012.
- [4] M. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan, "I see how you feel: Training laypeople and professionals to recognize fleeting emotions," in *International Communication Association*, 2009.
- [5] W. Yan, Q. Wu, Y. Liu, S. Wang, and X. Fu, "Casm database: A dataset of spontaneous micro-expressions collected from neutralized faces," in *Proc. AFGR*, 2013, pp. 1–7.
- [6] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen, "Recognising spontaneous facial micro-expressions," in *Proc. ICCV*, 2011, pp. 1449–1456.
- [7] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary pattern with an application to facial expressions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 915–928, 2009.
- [8] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikäinen, "A spontaneous micro-expression database: Inducement, collection and baseline," in *Proc. AFGR*, 2013.
- [9] W. Yan, X. Li, S. Wang, G. Zhao, Y. Liu, Y. Chen, and X. Fu, "CASME II: An improved spontaneous micro-expression database and the baseline evaluation," *PLOS ONE*, vol. 9, no. 1, pp. 1–8, 2014.
- [10] A. Davison, M. Yap, N. Costen, K. Tan, C. Lansley, and D. Leightley, "Micro-facial movements: an investigation on spatio-temporal descriptors," in *ECCV workshop on Spontaneous Behavior Analysis*, 2014.
- [11] J. Ruiz-Hernandez and M. Pietikäinen, "Encoding local binary patterns using re-parameterization of the second order gaussian jet," in *Proc. AFGR*, 2013.
- [12] S. Wang, W. Yan, G. Zhao, and X. Fu, "Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features," in *ECCV workshop on Spontaneous Behavior Analysis*, 2014.
- [13] Y. Wang, J. See, R. Phan, and Y. Oh, "LBP with six interaction points: Reducing redundant information in LBP-TOP for micro-expression recognition," in *Proc. ACCV*, 2014.
- [14] Y. Guo, C. Xue, Y. Wang, and M. Yu, "Micro-expression recognition based on cbp-top feature with elm," *International Journal for Light and Electro Optics*, vol. 127, no. 4, p. 2404, 2009.
- [15] X. He, D. Cai, and P. Niyogi, "Monogenic riesz wavelet representation for micro-expression recognition," in *IEEE International Conference on Digital Signal Processing*, 2015, pp. 1237–1241.
- [16] X. Huang, G. Zhao, X. Hong, M. Pietikäinen, and W. Zheng, "Texture description with completed local quantized patterns," in *Proc. SCIA*, 2013, pp. 1–10.
- [17] X. Huang, G. Zhao, X. Hong, W. Zheng, and M. Pietikäinen, "Spontaneous facial micro-expression analysis using spatiotemporal completed local quantized patterns," *Neurocomputing*, no. 175, pp. 564–578, 2016.
- [18] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikäinen, "Towards reading hidden emotions: a comparative study of spontaneous micro-expression spotting and recognition methods," *IEEE Trans. on Affective Computing*, no. 99, pp. –, 2016.
- [19] S. Zhang, B. Feng, Z. Chen, and X. Huang, "Micro-expression recognition by aggregating local spatio-temporal patterns," in *Proc. MMM*, 2017, pp. 638–648.
- [20] S. Jain, C. Hu, and J. Aggarwal, "Facial expression recognition with temporal modeling of shapes," in *Proc. ICCV*, 2011, pp. 1642–1649.
- [21] L. Houam, A. Hafiane, A. Boukrouche, E. Lespessailles, and R. Jennane, "One dimensional local binary pattern for bone texture characterization," *Pattern Analysis and Applications*, vol. 17, pp. 179–193, 2014.
- [22] G. Mateos, "Refining face tracking with integral projection," in *Proc. AVBPA*, 2003, pp. 360–368.
- [23] G. Mateos, A. Ruiz-García, and P. Lopez-de Teruel, "Human face processing with 1.5d model," in *Proc. AMFG*, 2007, pp. 220–234.
- [24] X. Huang, S.-J. Wang, G. Zhao, and M. Pietikäinen, "Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection," in *Proc. ICCV Workshop*, 2015, pp. 1–9.
- [25] X. Huang, S. Wang, X. Liu, G. Zhao, and M. Pietikäinen, "Discriminative spatiotemporal local binary pattern with revisited integral projection for spontaneous facial micro-expression recognition," *IEEE Transactions on Affective Computing*, pp. 1–13, 2017.
- [26] W. Yan, Q. W. an J. Liang, Y. Chen, and X. Fu, "How fast are the leaked facial expressions: the duration of micro-expressions," *Journal of Nonverbal Behavior*, vol. 37, no. 4, pp. 217–230, 2013.
- [27] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: exact recovery of corrupted low-rank matrices via context optimization," in *Proc. NIPS*, 2009, pp. 2080–2088.
- [28] L. Wang and H. Cheng, "Robust principal component analysis for sparse face recognition," in *Proc. ICICIP*, 2013.
- [29] S. Deans, *The Radon transform and some of its applications*. John Wiley and Sons.
- [30] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary pattern: a comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.
- [31] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, pp. 38–59, 1995.
- [32] Z. Zhou, G. Zhao, Y. Guo, and M. Pietikäinen, "An image-based visual speech animation system," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 10, pp. 1420–1432, 2012.
- [33] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.
- [34] S. Liong, J. See, R. Phan, A. Nego, Y. Oh, and K. Wong, "Subtle expression recognition using optical strain weighted features," in *Proc. ACCV workshop*, 2014, pp. 644–657.
- [35] S. Liong, R. Phan, J. See, Y. Oh, and K. Wong, "Optical strain based recognition of subtle emotions," in *Proc. ISAPCS*, 2014, pp. 180–184.