IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 17, 2021, accepted March 7, 2021, date of publication March 10, 2021, date of current version March 19, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3065125

Social Density Monitoring Toward Selective Cleaning by Human Support Robot With 3D Based Perception System

ANH VU LE[1,2], BALAKRISHNAN RAMALINGAM[1], BRAULIO FÉLIX GÓMEZ[1], RAJESH ELARA MOHAN[1], TRAN HOANG QUANG MINH[2], AND VINU SIVANANTHAM[1]

[1]ROAR Laboratory, Engineering Product Development, Singapore University of Technology and Design, Singapore 487372
[2]Optoelectronics Research Group, Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam

Corresponding author: Tran Hoang Quang Minh (tranhhoangquangminh@tdtu.edu.vn)

This work was supported in part by the National Robotics Programme through Robotics Enabling Capabilities and Technologies under Project 192 25 00051, in part by the National Robotics Programme through Robot Domain Specific under Project 192 22 00058, and in part by the Agency for Science, Technology and Research.

ABSTRACT Monitoring the safe social distancing then conducting efficient sterilization in potentially crowded public places are necessary but challenging especially during the COVID-19 pandemic. This work presents the 3D human space-based surveillance system enabling selective cleaning framework. To this end, the proposed AI-assisted perception techniques is deployed on Toyota Human Support Robot (HSR) equipped with autonomous navigation, Lidar, and RGBD vision sensor. The human density mapping represented as heatmap was constructed to identify areas with the level being likely the risks for interactions. The surveillance framework adopts the 3D human joints tracking technique and the accumulated asymmetrical Gaussian distribution scheme modeling the human location, size, and direction to quantify human density. The HSR generates the human density map as a grid-based heatmap to perform the safe human distance monitoring task while navigating autonomously inside the pre-built map. Then, the cleaning robot uses the levels of the generated heatmap to sterilize by the selective scheme. The experiment was tested in public places, including food court and wet market. The proposed framework performance analyzed with standard performance metrics in various map sizes spares about 19 % of the disinfection time and 15 % of the disinfection liquid usage, respectively.

INDEX TERMS COVID-19, human space, social distance, cleaning robotics, human support robot.

I. INTRODUCTION

The recent outbreak of COVID-19 has caused a pandemic alert around the world. It has now globally affected almost all the continents, infecting more than 82 million people and 1,79 death reports (30 December 2020). According to World Health Organisation (2020), physical distancing and routine sterilizing are the effective ways to slow down the spread of the virus because when people maintain safe social distancing and avoid physical contact, the chances of transmitting the virus from one person to another reduces significantly [1]. The absence of an approved vaccine for this disease urges the need to minimize the spread of the contagious and lethal virus. Social distancing measures proved the efficient in reducing the risk of the local spread of COVID-19. Tobías and Saez *et al.* [2] statistics reports indicate that COVID 19 spread has considerably reduced in Spain and Italy after implementing safe social distancing measures.

In a safe distancing measure, the physical distancing of at least 1m is essential in crowded zone [1]. Globally, several practices are implemented to ensure safe social distancing and avoid crowd gathering in busy places. For example, several countries are deploying Safe Distancing Officers (SDO) in public places to monitor the people, safe distance sign mark, limiting the number of people in workplaces, and restricting to large gathering [1], [3]. However, monitoring and tracking safe social distancing in potentially crowded places such as public transit stations and industrial settings such as factory workplaces, dormitories, schools, and shopping malls are quite challenging. Moreover, monitoring safe

The associate editor coordinating the review of this manuscript and approving it for publication was Ehsan Asadi.

distance measures is a manual process that can bring the surveillance staff in close proximity with people affected by COVID-19.

One significant aspect that has been studied widely in social safety for spatial interaction is the idea of personal space, or proxemics [4]–[8]. According to [4], based on different types of interaction and relationships between people, people maintain different culturally defined interpersonal distances. Hall differentiated four different zones between the interaction distance as follows: Public interaction: Public speeches in-crowd, more than 4 m away. Social interaction: business meetings, 1–4 m. Personal interaction: friendly interaction distance gap of arm's length, 0.5–1 m. Intimate interaction: about 0.5 m apart.

In this paper, we specifically focus on the zone of personal space as it is a culturally defined zone of ''spatial insulation'' that people maintain around themselves and others [9]. Research work in [10], has described personal space for the case of both people approaching each other and standing in line is asymmetric [11]. Also [12] discusses personal space is not a constant as it is dependent on individual attributes such as volume, age, gender, and direction of interaction.

Automate the monitoring of social distancing measures is a viable solution. Some effort worldwide that were implemented on an ad-hoc framework to enforce social distancing rules. Some industries recently use lightweight wearable devices that employ ultra-wideband (UWB) technology to measure people's distance automatically. It alerts them immediately if they come closer than the required distance [13]. Some countries have adopted ubiquitous technologies, such as Wi-fi, cellular, GNSS positioning (localization) systems to monitor and alert the social distance in public and crowded areas [1], [14]. Recently, many countries worldwide have used the drones, IoT, and AI-assisted techniques to monitor the human density, predict and alert the safe distance breach in crowded areas in indoor and outdoor [3]. However, these techniques have numerous limitations and have poor performance in dynamic and complicated indoor environments. With the advanced high speed and accuacy devices, the captured 3D object [15] have been applied in suveliance applicaitons for quality assessmen. Moreover, the exiting CCTV system [16]–[18] consisting of monocular RGB cameras is not cover the whole the workspace and CCTV is hard to detect the 3D human attributes. In the works of [19], the authors have addressed the distance-time encounter patterns in a crowd that allows the fixed surveillance system to identify social groups, such as families, by imposing adaptive thresholds on the distance-time contact patterns. On other hand, in the works of [20], an artificial intelligence based social distancing surveillance system is present to detect distances between human and warning them can slow down the spread of the deadly disease. The work presented the four essential ethical factors of surveillance system of: keep the privacy, not target the particular detected human, no human supervisor, and open-source. The work of [21] has proposed a deep learning technique to track the social distance by an fixed overhead perception system. The system integrated transfer learning of the YOLOv3 an open source deep learning object detection framework with an overhead human data set in video sequences. The proposed system simply uses the Euclidean distance of detected bounding box centroids to estimate the distances of pairwise people. Then to estimate social distance violations between people, an approximation of physical distance to pixel are set as fixed threshold. However, the quantification in term of 3d interaction between human space and the utilization of human interaction are not considered in the mentioned references.

In this context, service robots are a viable candidate for monitoring safe distance measure. Robotics and several other autonomous technologies in AGV vehicles have made great strides in fighting the COVID-19 pandemic [22], [23]. The service robot uses the AI-assisted technology to deliver the medicine to covid 19 patients, safe entry check-in body temperature measurement, sanitize the infected area, frequent cleaning of high touchpoints like hospital walls, floor, and door handle [16], [24], [25]. Hence, by considering the advantage of service robots and AI-assisted technology, as well as the flexible navigation in the vast and complex environments such as shopping mall, food court, wet market, resents robot system gradually replace human in the tedious jobs. In the monitoring, surveillance tasks and cooperate conveniently through the comment operation system with another service robot such as cleaning robot are the recent trends.

In this article, based on the literature survey on the human space and multi human interaction, we propose a human safe distance monitoring technique using Toyota Human Support Robot (HSR) and AI-assisted 3D computer vision framework. The computer vision framework was built with a modified Openpose 3D human tracking algorithm, depth image fusion technique, Gaussian heat map scheme, and uses the RGBD vision sensor data. The entire framework built on top of Robot Operating System (ROS) [26] and tested in real-time with HSR Toyota robot [27] deployed in crowded public areas of Singapore, including food court and wet market. The service robot navigates to clear the waypoints around the mapped indoor area and performs the SDO tasks that include detecting people's clusters, space between the humans, human interaction pose, safe distance measure, and raising warning alerts commuters when violating the safe distance rule.

The main contributions of this paper are threefold. (1) the design and development of safe social distance surveillance with collaborative multi-robot cleaning system, (2) to develop and implement a vision-based AI perception algorithm for the robot to closely generate a heatmap based on the 3D human interaction model, (3) to develop and test an adaptive velocity behaviour model for the multi-robot cleaning systems to clean the environment efficiently based on the generated heat map. The efficiency of the proposed selective cleaning system was assessed with standard performance metrics, and results are reported.

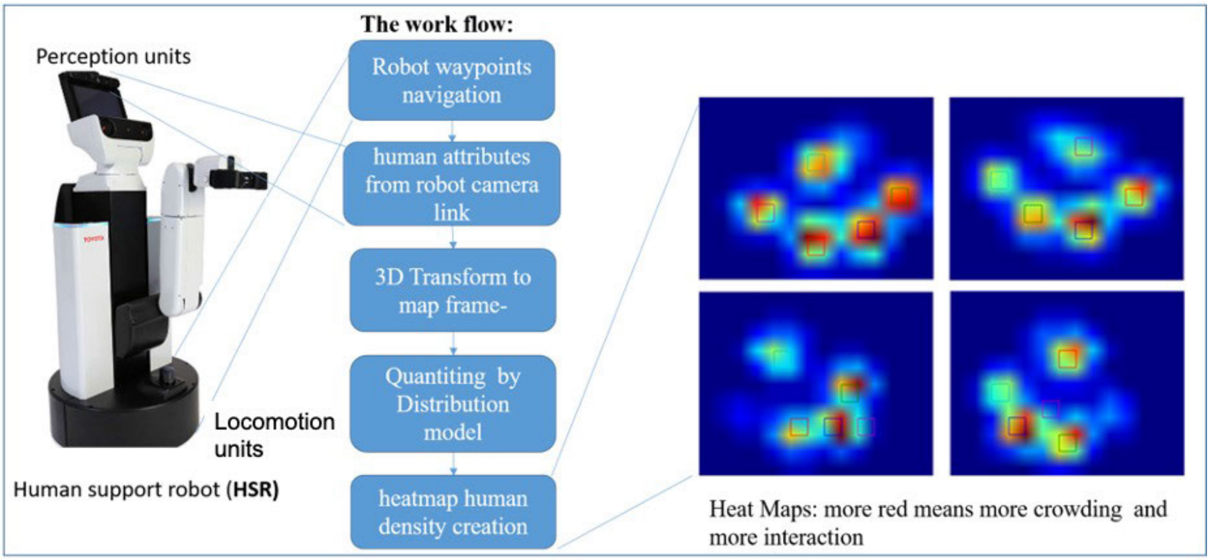

FIGURE 1. The framework for human density mapping based on human attributes.

The paper is organized as follows. The context of the application is introduced in Section 2. The methodology of the proposed robot is detailed in Section 3. In Section 4, the HSR platform is presented. The optimal human attributes estimation methods and Social Distancing and Density heat map are validated in Section 5 and Section 6, respectively. The conclusion, together with potential future works, is explored in the last Section 7.

II. CONTEXT OF APPLICATION

Figure 1 depicts the workflow used in the proposed framework to map the human density based on detected human attributes by service robot. The output heatmap generated is a distribution that highlights the level of interaction between humans. Based on this heatmap distribution, the system identifies locations with the level of human interactions in a pre-built map of the environment. The user can set a threshold to determine the level of interaction and issue a warning alert whenever safe distancing measures are violated. Moreover, the system helps deploy area sterilization by cleaning robot systems that activate adaptive cleaning based on the area heatmap levels. The interactive monitoring system has been deployed for trials evaluations at testbed wet-market in Singapore.

III. METHODOLOGY

This software framework is developed on a Toyota HSR robot equipped with an onboard AI-embedded perception system. Unlike traditional approaches that only monitor safe physical distancing based on a person's position from a fixed camera system, this study defines the degree of human-to-human interaction by quantifying social interaction using an asymmetrical Gaussian distribution with the shape derived from human attributes. Note that the theoretical background for social interaction is based on the survey work of [4], which state that human interaction tends toward the human direction, and space occupied. This is fundamental to motivate us to quaintly the human by human space of 3D location, volume and facing direction. The perception unit of HSR outputs the human attributes includes the 3D positioning of human joints in the map, the volume of space occupied by the detected person, and the person's movement direction.

To quantify human-to-human interactions, we propose a distribution kernel with direction and magnitude proportional to the detected and tracked a person's identity and plotted on the map concerning the person's position. To this end, we extract the color image data from the RGB-depth camera, and human joints are detected and marked by the AI-based Openpose algorithm [28]. The 3D depth information from the camera frame is used to estimate the person's joint positions, and the position is then converted to the map frame to be tracked while it is in the field of view of HSR camera. We divide the pre-built map into grid cells and plot 3D directional distributions of tracked human positions over the grid cells.

The asymmetrical Gaussian distribution has its peak value set at the location where humans are detected and spread gradually along the human volume and moving direction of each person. After plotting all the human positions, each cell value in the map is updated by accumulating the human distribution values during detection. Figure 2 shows the proposed framework models in terms of quantity the interaction of three persons at the same distance but own different attributes then interaction between them differently.

IV. HSR ROBOT PLATFORM

A. OVERVIEW OF HSR SYSTEM ARCHITECTURE

The Human Support Robot (HSR), as shown in Figure 3, is the research platform developed by Toyota Ltd. Co. HSR has been implemented in multiple applications such as

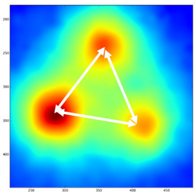

FIGURE 2. Proof of concept where 3 humans stay same distances and with difference attributes so that interaction differently represented by the level of color of areas in constructed heatmap.

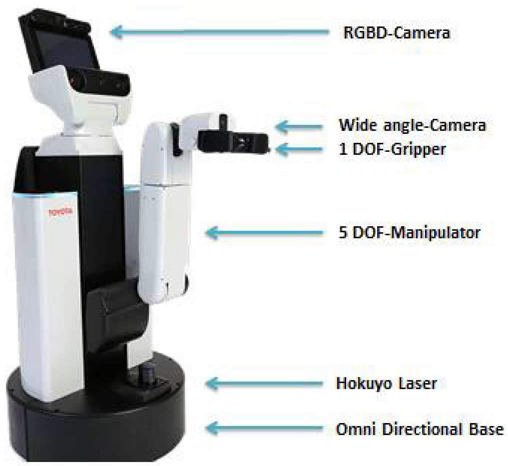

FIGURE 3. Human Support Robot (HSR).

cleaning and inspection [29]. The HSR platform is equipped with sensors like 2D Lidar, IMU, ASUS Xtion RGBD camera, stereo vision camera, and embed main processing unit with dedicated GPU necessary to support the autonomy AI-based perception of the robot platform. This research uses the data from the ASUS Xtion RGBD camera that extracts RGB color and depth information around the robot. The RGBD camera has a resolution of 1280 $\times$ 1024 and runs at 30 frames per second. Hokuyo URG-04LX 2D LiDAR installed in the HSR base enables the simultaneous localization and mapping to build a map of the environment. Besides, pneumatic bumpers are installed at the robot base to provide an emergency stop to avoid any possible collision. A dual system including Intel Core i7 CPU (16GB RAM, 256GB SSD) and NVIDIA Jetson TK1 embedded GPU board are the central processing units used alongside the robot. Robot control and motion planning require proper communication between software algorithms and hardware modules. This communication framework is enabled through the ROS environment by ROS nodes and topics in Linux Ubuntu-based system.

B. HSR COVERS AREA BY DEFINED WAYPOINTS NAVIGATION

The service robots are the active research and gradually to be the commercial applications such as building maintenance [30]–[32], road maintenance [33], [34], and path tracking [35], [36] and service robot [37], [38]. The HSR robot the research service platform developed by Toyota [29] is set up in a public wet-market and food-court in a neighborhood community area in Tampeniss Town, Singapore. This location is chosen as it is a highly busy environment with more crowd gathering. So it is crucial to monitor safe social distancing and regularly sterilize to avoid the spreading of the virus.

Hector Simultaneous Localization and Mapping (SLAM) [39] is implemented in HSR to map continuously from the environment using a Hokuyo laser to locate its correct position. Within the generated market map, the HSR can reach to its destination efficiently and quickly. The generated map is shown in Figure 4(left) and the market is shown in Figure 4(right).

FIGURE 4. Generated market map & Market environment.

To maximize area coverage, a set of coordinates in the map environment are generated in a zigzag pattern. These generated coordinates are given to the robot one by one as sub-goals for the robot to manoeuvre. The waypoints are assigned manually to HSR move based after the map is built. HSR generates a trajectory connecting the coordinates using a local trajectory planner algorithm and follows the trajectory to generate the heatmap by simultaneously inspecting the human density in the market environment. Based on the feedback from laser data and wheel odometry information, HSR will move to the designated coordinates one by one. During the movement, HSR can make proper decisions from collected data for obstacle detection and avoidance.

V. HUMAN ATTRIBUTES ESTIMATION METHODS

A. HUMAN JOINT DETECTION IN 2D COLOR IMAGE

HSR uses an inbuilt Asus Xtion Camera that provides the RGB and depth data. This data is used as input for OpenPose [28] that gives us the skeleton of the detected humans in every frame, as shown in Figure 5. This dependency gives us every joint of the human body in a two-dimensional array of 25 rows and 3 columns where every row is a skeleton's joint. Note that we consider the (x,y) plane is the 2d navigation, (x,y) plane is the camera world coordinate, y is the distance and z direction models the object height. The first two columns are the X and Z pixel position of the joint in RGB image, the third column shows the percentage of accuracy.

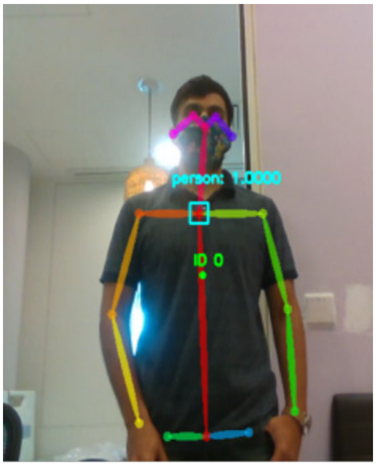

FIGURE 5. Human skeleton by Openpose.

By taking the average of accuracy from every joint, we filter the false detections.

1) 3D LOCATION ESTIMATION

After estimating the 2D pixel coordinates of joints from the RGB camera frame using the object detection algorithm, the system will derive the world frame corresponding to the 3D position obtained from depth data. Since the RealSense D435 of the perception system is a stereo camera with both an RGB image sensor and depth image sensor, they are of the same resolution. Since each pixel in the RGB image corresponds to the pixel in the Depth image, this enables us even to map noisy objects and localize their pixel coordinates concerning depth value. When the depth images are subscribed, the image will be filtered via an adaptive directional filter [33] to remove noise. After filtering the depth image, the depth value of the noisy human objects, y_o, can be derived by using the Equation 1.

$$y_o = \frac{\sum_{\Omega} y_i}{\Omega} \tag{1}$$

where Ω represents the sum of pixels within a 10×10 window, and y_i represents the value of the depth in the pixel. The human objects in real-world coordinates will be estimated by converting the X and Z pixel coordinate system to the x and z world coordinate system. This is done through camera calibration techniques to calculate the intrinsic and extrinsic matrix values of the camera. Camera calibration method estimates the focal length, f_x, and f_z, of the RealSense RGB-D sensor and the optical center, c_x, and c_z, of the RealSense RGB-D sensor in its respective x and y coordinate to get the intrinsic camera matrix,I. The RealSense RGB-D sensor in-built functions also provide us with the translation vector, K, and a Q rotation vector. The extrinsic matrix, S, can be derived from $S = [QK]$. We can convert from the pixel coordinate system to the world coordinate system as Equation 2.

$$p = I * [QK] * P = I * S * P. \tag{2}$$

where P is the world coordinates, p is the pixel coordinates.

2) 3D JOINT TRACKING

The DeepSORT is an improved version of the Simple Online and Real-time Tracking (SORT) algorithm. The DeepSORT tracking framework was build using the hypothesis tracking technique with Kalman filtering algorithm and DL based association metric approach (Deep SORT). Further, the Hungarian algorithm was utilized to resolve the uncertainty between the estimated Kalman state and the newly received measuring value. The tracking algorithm uses the appearance data to improve the performance of Deepsort [40], [41].

In this work, the 3D joint coordinates of human detection with the aligned RGB and depth frames are fed into the modified DeepSORT network for tracking the human movements. Then we retrained by transferred learning technique the original DeepSORT object tracking algorithm which is initially designed for 2D object tracking to track the detected 3D joints. According to bounding box coordinates and object appearance, deep sort assigns an id for each human detection and performs the tracking in 3D camera frame coordinate.

B. TRANSFORMATION TO MAP FRAME AND HUMAN VOLUME

When HSR navigates, the SLAM algorithm maintains a relative position from the map frame to the robot base link frame. Sine color and depth image is aligned by Realsense camera, we crop the area of 10×10 at center of the neck's position of every detected human in the depth images. By appling the mean filter for the cropped area, we get the distance with noise is filtered out from the camera to the human's neck, thus avoiding the probability of taking a None value in-depth data array. The frame transformation operations are done in ROS to estimate an object's position in the real world relative to the sensor component used on the robot. The transformation is then translated from the sensor frame to the Base Link, the robot's centre. The base link frame is the base point of reference on the robot to locate the humans and link their position to the Map transform frame that is the origin of the World Space where we handle on ROS. We estimate the distance between the shoulders, the height of the human, and the hips' distance from the OpenPose output data. Then we apply Equation 3 for the calculations.

$$H_s = D_S \times H_h \times D_h \tag{3}$$

where H_s is human size, D_s is the Distance Shoulders and H_h is human height and D_h is the Distance Hip.

Base on [4], to tun the parameters the values σ_h,σ_s and σ_r of asymmetrical Gaussian distribution adaptively, we take the referenced values as the adult human dimensions: 150cm for the height, 40 cm as the distance between the two shoulders, and 30cm as the distance between hips. The values will be used as a quotient with the actual measures of the detected human. The results will be multiplied to σ_h,σ_s and σ_r as a scalar value. Derived parameters give us a variation of size in the distribution proportional to the current size of the detected humans.

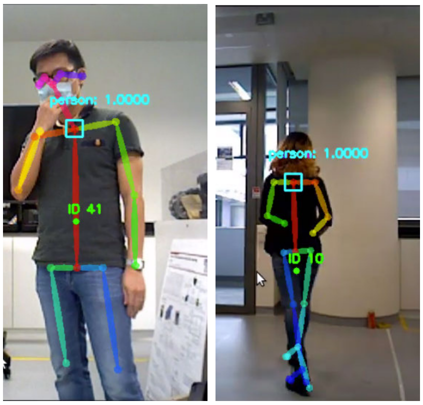

FIGURE 6. In the left image a person facing directly the camera with the nose tracked by OpenPose, in the right a human facing contrary the camera without nose joint.

C. HUMAN DIRECTION AND FACING

The human direction is the vector with the θ direction is set to orthogonal with the vector linking left and right shoulder joints. Once we estimate the human left and right shoulder joints' position, we will find the angle of the vectors formed by those two joints. To deal with the situations of human facing upward and backward the camera, specifically, we classify the detected human direction into front and rear facing cases. Then the formula as in Equation 4 is used to find the direction for the case of existing the detected human nose joint and in Equation 5 for the case of non-existing the detected human nose joint. given the left and right shoulders have the tracked codinates $x_{ls}, y_{ls}, z_{ls}, x_{rs}, y_{rs}, z_{rs}$, the orientation value gives the angle of the line in a clockwise direction. Nevertheless, this angle can be wrong sometimes because the shoulders can only be detected in an interval of 180 degrees. To enhance the accuracy, with the known camera angle value concerning the tracked 3D human joints, we can estimate whether this human is facing the camera.

OpenPose tries to detect every joint in the human body. If it is impossible to detect the joint by the probability mean lower than the threshold, it will fill with an empty value such as null value for every column of the respective joint. This behavior is essential because if the human face a contrary direction to the camera, there will not be any joint-related with the face on the joint's array. Taking that as a reference, we can know if the human is facing toward or backward, so we modify the formula as follows, taking into account that 0 degrees are facing contrary to the camera:

If Joint 0 (Nose) exist:

$$\theta = (atan2(|y_{ls} - y_{rs}|, |x_{ls} - x_{rs}|) \times 180/\pi) + 180 \quad (4)$$

else:

$$\theta = atan2(|y_{ls} - y_{rs}|, |x_{ls} - x_{rs}|) \times 180/\pi \quad (5)$$

VI. SOCIAL DISTANCING AND DENSITY HEAT MAP

A. DISTRIBUTION KERNEL

An asymmetric Gaussian integral function using for social robot interaction in [42] was deployed to calculate the distribution of the detected humans. We quantify the shape of space human occupying by adjusting the sigma parameters of this function. Equation 6 shows a 1-dimensional Gaussian distribution with σ:

$$f(x) = exp^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

The two dimension Gaussian distribution as shown in 7 is the center at $(x_0; y_0)$, and the variance are represented by σ_x and σ_y:

$$f(x, y) = exp^{-(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2})} \quad (7)$$

Typically, a 2-dimensional Gaussian distribution is symmetric along the x and y axes. However, to formulate a distribution function for personal space between people, a 2-dimensional asymmetric Gaussian function is necessary; this can be done with a shared σ_x and differing σ_y values.

ALGORITHM 1: Asymmetric Gaussian Based-Heatmap Generation

1 : *find* : $\alpha \leftarrow atan2(y - y_c, x - x_c) - \theta + \pi/2$
2 : $Normalize(\alpha)$
3 : $a \leftarrow (cos\theta)^2/(2\sigma^2) + (sin\theta)^2/(2\sigma_s^2)$
4 : $b \leftarrow sin(2\theta)/(4\sigma^2) - sin(2\theta)^2/(4\sigma_s^2)$
5 : $c \leftarrow (sin\theta)^2/(2\sigma^2) + (cos\theta)^2/(2\sigma_s^2)$
6 : $return(exp(-(a(x-x_c)^2 + 2b(x-x_c)(y-y_c) + c(y-y_c)^2))$

Algorithm 1 explains the computation of an arbitrarily rotated asymmetric Gaussian function at a detected human location $(x; y)$. The following notations are used: θ is the direction of the distribution taken from the human's estimated orientation. σ_h variance in θ direction. σ_s variance to the sides ($\theta +- \pi/2$ direction). σ_r variance to the back (-θ direction). Those parameters σ_h, σ_s, σ_r are set to be proportional with the detected human volume. Lines 1, 2, and 3 calculate the normalized angle of the human facing σ_s direction. This means α points along the side of the function, and $0 < \alpha < \pi$. The two 2D Gaussian functions in Line 3 will be used for the point of interest, $(x; y)$. In the case of $\alpha = 0$, the point of interest is located directly to the side of the function center and depends only on σ_s. Figure 7 displays some views of an Asymmetric Gaussian cost function as shown in Equation 8. This function has a rotation of $\theta = \pi/6$, is centered at (0; 0) and has as variances $\sigma_h = 2{:}0$, $\sigma_s = 4{=}3$, and $\sigma_r = 1$. The maximum cost of this function is 1 in the center of the distribution.

$$f(x, y) = exp^{-(a(x-x_c)^2 + 2b(x-x_c)(y-y_c) + c(y-y_c)^2)} \quad (8)$$

The Figure 8 presents the example the human detected at the origin (0,0) facing upward by $\theta = \pi$ rads, $\sigma_s = 13.33$, $\sigma_r = 10.00$, $\sigma_h = 20.00$

B. HEAT MAP GENERATION

The heat map generation consists of following steps:

1) OpenPose that detect the humans and give the joints of the skeleton.

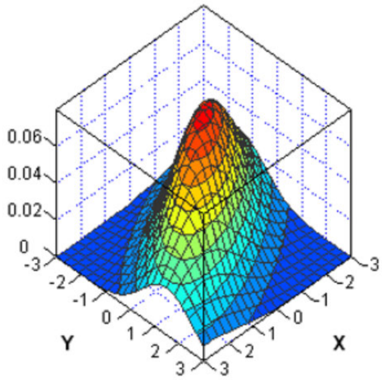

FIGURE 7. Asymmetric Gaussian Distribution views located at (0, 0), oriented by $\theta = \pi/6$ rad, and with variances $\sigma_h = 2$, $\sigma_s = 4$, and $\sigma_r = 1$.

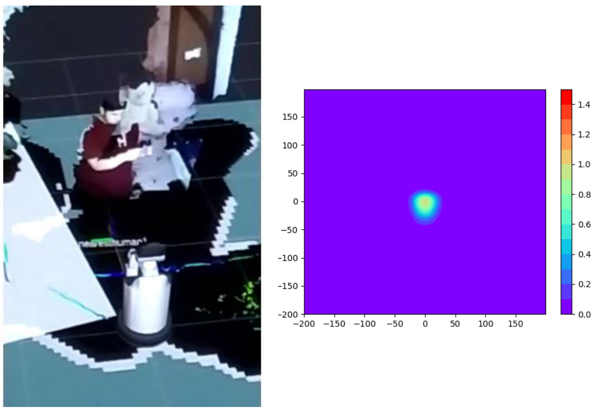

FIGURE 8. Output of the Aasymmetric Gaussian integral function for one detected human at origin and facing upward.

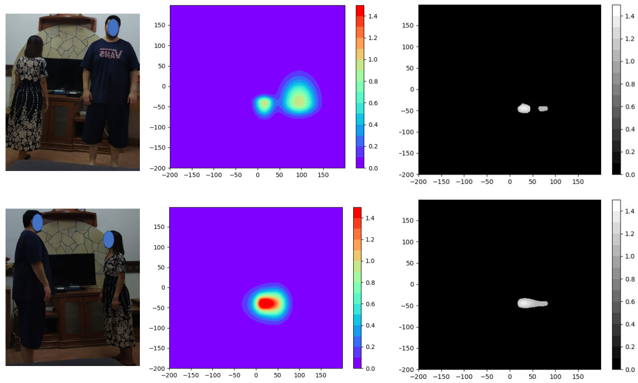

FIGURE 9. Two persons facing opposite directions & Two persons facing each other.

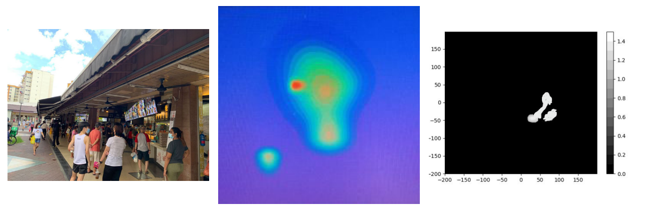

FIGURE 10. Four persons detected by system.

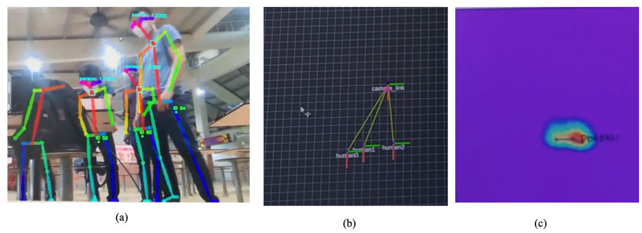

FIGURE 11. tracking group of humans and remove the outliner. (a) three detected humans and one outliner, (b) tracking the humans' location on ROS Rviz, (c) heatmap of three humans.

2) The Detection node that will take OpenPose output and calculate location, volume, and orientation

3) The Distribution Kernel and plotter Node that calculates the intensity of distribution using the asymmetric Gaussian formula and displays the results as a 2D heatmap.

Specifically, the map is divided into a grid base workspace with each cell equals 0.1×0.1 m. Once the distribution's raw data is calculated, we take it and plot it in a color mesh plot. For this process, we need to set up an interval as a reference, so the intensity of distribution is represented as a variation of color. The interval is chosen between the min value of intensity currently registered and the max value by default. This scale is updated every frame, so if there is an increment in the intensity, it will update the scale in real-time. An interval can be set up as an argument of the Heatmap generator node in ROS.

When two points overlap in the calculation of the two detected persons, the algorithm adds the intensity values of all asymmetric Gaussian distributions for to form the level of human interaction in the considered cell of the grid map. The detected human's id with the grid cell within the map frame where the human presents is tracked by HSR. During the robot navigation, if the detected person has the same id and staying in the same grid cell, the distribution of this human will not be re-plotted. the new distribution is plotted if the new id is detected or the tracked id moves to a different cell.

VII. EXPERIMENTAL RESULTS IN REAL ROBOT

To evaluate the proposed software algorithm's performance to monitor social safe distancing measures in public places, we implemented the software packages on HSR through the ROS framework. Principally, HSR runs through the ROS interface, and ROS architecture is flexible to implement different tools or dependencies in the form of packages and model utilities. OpenPose uses the RGB-D input to detect every human; this happens in the Detection node. Once we detect a human, we publish all the joints of every detected human, the location and orientation, and volume. To this end, we use a costumed Message on ROS that we publish under the topic. In parallel with the generated distribution kernels during robot flows the trajectory, the plotter Node will accumulate all values belonged to each cell of the grid map to generate the final heatmap.

In the first experimental section, we experimented with the proposed framework by considering several different interaction scenarios. Figure 9 demonstrates the interaction

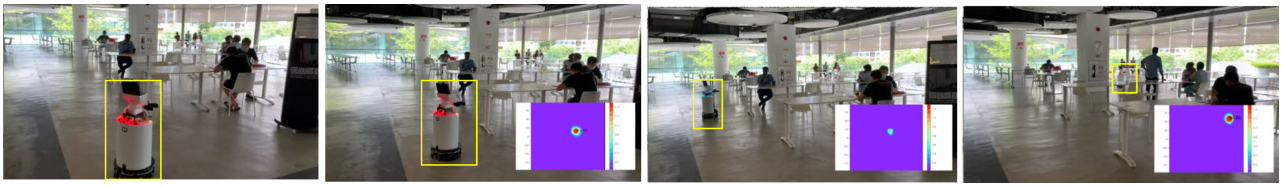

FIGURE 12. Robot navigation with real-time density mapping.

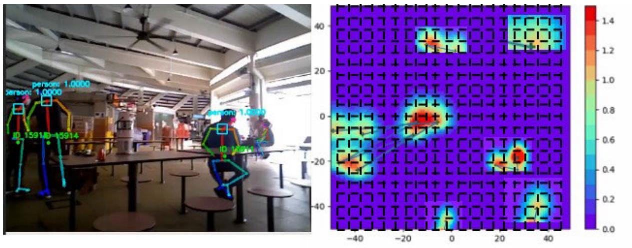

FIGURE 13. Heat map for Testbed Market environment.

scenarios between two persons maintaining a distance of 1.5 meters. Figure 9 shows a scenario where two persons face in different directions. Figure 9 b shows the scenario where two standing persons face opposite directions. One can observe from the distribution plot, two persons facing opposite directions generate the least risk area. On the other hand, the second scenario demonstrates a high risk in the distribution plot when two persons face each other.

The third scenarios consider a group of persons with different attributes and locations. The output results are shown in Figure 10 and Figure 11, where we can observe that the pedestrians maintaining a safe distance between each other have their distribution spread in the green region. On the other hand, when people come closer and face each other, their distribution changes gradually from blue to red spectrum. Meaning that they are interacting with other people and are not maintaining a safe distance. One can observe from the results that four persons with different attributes and locations. The persons who stay far away have the distribution with greener color. On the other hand, the persons who stay in the middle of two big volume persons and face toward them have the redder color; this indicates that these persons have a greater chance of interacting with other persons.

Based on the distribution's value, we can select the predefined threshold value to identify the area in the map with the higher risk of human interaction, as shown in Figure 12 and Fugue 13. Depending on this input threshold value, the HSR robot can trigger the alarm warning when it patrols around the working environments.

In the second experimental section, we validated the efficiency and performance of proposed heatmap in terms of selective disinfection efficiency by deploying the robot system at a public food court in Tampines, Singapore. Firstly, we mapped the environment using occupancy grid-based mapping. Depending on the human interaction activity in the environment, a heat map is constructed on the grid cells by the proposed framework, as shown in section 4. The heatmap

FIGURE 14. The sterilization robot in tested foodcourt.

TABLE 1. Numerical spent time and liquid of regular and selective cleaning schemes.

Map	Method	Total Time minutes	Total liquid liters
Map 1	Regular cleaning	20.23	1.01
(about 100sqm)	Selevtive cleaning	16.36	0.85
Map 2	Regular cleaning	33.93	1.42
(about 150sqm)	Selevtive cleaning	27.07	1.15
Map 3	Regular cleaning	41.91	1.58
(about 200sqm)	Selevtive cleaning	35.15	1.39

distribution result is used as input to a disinfection robot to decide disinfecting areas having level distributions on the map need to be cleaned.

The selective disinfection algorithm estimates the time to sterilize the environment in proportion to the heatmap distribution values. Specifically, the higher the heatmap's intensity, the longer the time will be taken to disinfect the region. In the case of a conventional disinfection robot, the main focus would only be on maximizing the area coverage. So during the cleaning process, the conventional disinfection robot spends equal time in all the map regions. Thus, the disinfection amount savings between the proposed method and conventional robotic disinfection will become a criterion for evaluation.

To conduct the comparison evaluation between regular uniform cleaning and selective cleaning, one conventional sterilization robot as shown in Figure 14 is first deployed to pray the cleaning liquid while following a defined zigzag trajectory, and the times it takes to cover the sub-maps of the wet market as of Figure 4(left) are recorded for all trials. This robot with a liquid sterilization system is deployed to cover the 100, 150, and 200 sqm areas with humans' presenting. Then, the total time to cover each testbed submap is divided by its size to find the time interval the robot has stayed at each grid cell. The found time is set to the time that the sterilization robot with selective cleaning methods remains at the highest heat cell (the cell with the reddest color inside the build heatmap). The retention time of cells will decrease gradually with the degrees of the cell heat maps.

Travel time and disinfection liquid spent were recorded during the 5 trials for each tesbed map. Table 1 describes the experiment's comparison averaged results on testbed layouts during the trials. The average per-time disinfectant solution offered by the selective method with heatmap can save about 19 %, 20 %, and 16 % the spent time and 15%, 19%, and 12% of disinfection liquid, respectively.

VIII. CONCLUSION

COVID-19 is the third pandemic of the 21st century. COVID-19 pandemic is easily spread by people in close proximity, especially in crowds with mobile individuals (e.g., food court, wet market). The proposed social distant monitoring and selective sterilization strategy have validated efficiency in the real public environment. The proposed system is the initial works to deploy an adaptive multi-robot cleaning strategy based on coverage path planning that works in synergy with the human interaction heat map generated by safe social distance monitoring systems

Our future works will focus on: redesigning the long-term autonomy framework intensity, implement the autonomous path generation to re-clean the part of the surface concerning the generated heatmap, working on the optimization algorithm to control the generated heatmap to reduce the running time usage.Since the deploying the system at a public food court in Tampines, Singapore requires the particular setups so that the comparisons between the surveillance systems will be also considered as the future works.

ACKNOWLEDGMENT

The authors would like to thank Jia Yin for insightful comments and discussion.

REFERENCES

[1] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact," *IEEE Access*, vol. 8, pp. 90225–90265, 2020.

[2] M. Saez, A. Tobias, D. Varga, and M. A. Barceló, "Effectiveness of the measures to flatten the epidemic curve of COVID-19. The case of Spain," *Sci. Total Environ.*, vol. 727, Jul. 2020, Art. no. 138761.

[3] M. Gupta, M. Abdelsalam, and S. Mittal, "Enabling and enforcing social distancing measures using smart city and its infrastructures: A COVID-19 use case," 2020, *arXiv:2004.09246*. [Online]. Available: https://arxiv.org/abs/2004.09246

[4] E. T. Hall, *The Hidden Dimension*, vol. 609. Garden City, NY, USA: Doubleday, 1966.

[5] S. Weitz, *Nonverbal Communication: Readings With Commentary*. New York, NY, USA: Oxford Univ. Press, 1979.

[6] P. Mishra, "Proxemics: Theory and research," *Perspect. Psychol. Res.*, vol. 6, pp. 10–15, 1983.

[7] P. Veerajagadheswar, K. Ping-Cheng, M. R. Elara, A. V. Le, and M. Iwase, "Motion planner for a tetris-inspired reconfigurable floor cleaning robot," *Int. J. Adv. Robot. Syst.*, vol. 17, no. 2, 2020, Art. no. 1729881420914441.

[8] A. V. Le, P. T. Kyaw, R. E. Mohan, S. H. M. Swe, A. Rajendran, K. Boopathi, and N. H. K. Nhan, "Autonomous floor and staircase cleaning framework by reconfigurable sTetro robot with perception sensors," *J. Intell. Robot. Syst.*, vol. 101, no. 1, pp. 1–19, Jan. 2021.

[9] J. K. Burgoon, D. B. Buller, and W. G. Woodall, *Nonverbal Communication: The Unspoken Dialogue*. New York, NY, USA: Harpercollins College Division, 1989.

[10] J. R. Aiello and D. E. Thompson, "Personal space, crowding, and spatial behavior in a cultural context," in *Environment and Culture*. Boston, MA, USA: Springer, 1980, pp. 107–178.

[11] Y. Nakauchi and R. Simmons, "A social robot that stands in line," *Auton. Robots*, vol. 12, no. 3, pp. 313–324, 2002.

[12] J. W. Burgess, "Interpersonal spacing behavior between surrounding nearest neighbors reflects both familiarity and environmental density," *Ethol. Sociobiol.*, vol. 4, no. 1, pp. 11–17, Jan. 1983.

[13] M. M. Islam, S. Mahmud, L. J. Muhammad, M. R. Islam, S. Nooruddin, and S. I. Ayon, "Wearable technology to assist the patients infected with novel coronavirus (COVID-19)," *Social Netw. Comput. Sci.*, vol. 1, no. 6, p. 320, Nov. 2020.

[14] A. A. R. Alsaeedy and E. K. P. Chong, "Detecting regions at risk for spreading COVID-19 using existing cellular wireless network functionalities," *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 187–189, 2020.

[15] U. Stenz, J. Hartmann, J.-A. Paffenholz, and I. Neumann, "High-precision 3D object capturing with static and kinematic terrestrial laser scanning in industrial applications—Approaches of quality assessment," *Remote Sens.*, vol. 12, no. 2, p. 290, Jan. 2020.

[16] A. V. Le, R. Parween, P. T. Kyaw, R. E. Mohan, T. H. Q. Minh, and C. S. C. S. Borusu, "Reinforcement learning-based energy-aware area coverage for reconfigurable hRombo tiling robot," *IEEE Access*, vol. 8, pp. 209750–209761, 2020.

[17] A. V. Le, P. Veerajagadheswar, P. T. Kyaw, M. A. V. J. Muthugala, M. R. Elara, M. Kuma, and N. H. K. Nhan, "Towards optimal hydro-blasting in reconfigurable climbing system for corroded ship hull cleaning and maintenance," *Expert Syst. Appl.*, vol. 170, May 2021, Art. no. 114519.

[18] S. K. Udgata and N. K. Suryadevara, *COVID-19, Sensors, and Internet of Medical Things (IoMT)*. Singapore: Springer, 2021, pp. 39–53.

[19] C. A. S. Pouw, F. Toschi, F. van Schadewijk, and A. Corbetta, "Monitoring physical distancing for crowd management: Real-time trajectory and group analysis," *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0240963.

[20] D. Yang, E. Yurtsever, V. Renganathan, K. A. Redmill, and Ü. Özgüner, "A vision-based social distancing and critical density detection system for COVID-19," 2020, *arXiv:2007.03578*. [Online]. Available: http://arxiv.org/abs/2007.03578

[21] I. Ahmed, M. Ahmad, J. J. P. C. Rodrigues, G. Jeon, and S. Din, "A deep learning-based social distance monitoring framework for COVID-19," *Sustain. Cities Soc.*, vol. 65, Feb. 2021, Art. no. 102571.

[22] M. Antony, M. Parameswaran, N. Mathew, S. V. S. J. Joseph, and C. M. Jacob, "Design and implementation of automatic guided vehicle for hospital application," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2020, pp. 1031–1036.

[23] Z. H. Khan, A. Siddique, and C. W. Lee, "Robotics utilization for healthcare digitization in global COVID-19 management," *Int. J. Environ. Res. Public Health*, vol. 17, no. 11, p. 3819, May 2020. [Online]. Available: https://www.mdpi.com/1660-4601/17/11/3819

[24] T. W. Teng, P. Veerajagadheswar, B. Ramalingam, J. Yin, R. E. Mohan, and B. F. Gómez, "Vision based wall following framework: A case study with HSR robot for cleaning application," *Sensors*, vol. 20, no. 11, p. 3298, Jun. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/11/3298

[25] K. P. Cheng, R. E. Mohan, N. H. K. Nhan, and A. V. Le, "Multi-objective genetic algorithm-based autonomous path planning for Hinged-Tetro reconfigurable tiling robot," *IEEE Access*, vol. 8, pp. 121267–121284, 2020.

[26] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw.*, Kobe, Japan, 2009, vol. 3, nos. 3–2, p. 5.

[27] U. Yamaguchi, F. Saito, K. Ikeda, and T. Yamamoto, "HSR, human support robot as research and development platform," in *Proc. Abstr. Int. Conf. Adv. Mechatronics, Toward Evol. Fusion IT Mechatronics (ICAM)*, 2015, pp. 39–40.

[28] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," 2018, *arXiv:1812.08008*. [Online]. Available: http://arxiv.org/abs/1812.08008

[29] J. Yin, K. G. S. Apuroop, Y. K. Tamilselvam, R. E. Mohan, B. Ramalingam, and A. V. Le, "Table cleaning task by human support robot using deep learning technique," *Sensors*, vol. 20, no. 6, p. 1698, Mar. 2020.

[30] A. Le, P.-C. Ku, T. T. Tun, N. H. K. Nhan, Y. Shi, and R. Mohan, "Realization energy optimization of complete path planning in differential drive based self-reconfigurable floor cleaning robot," *Energies*, vol. 12, no. 6, p. 1136, Mar. 2019.

[31] A. V. Le, N. H. K. Nhan, and R. E. Mohan, "Evolutionary algorithm-based complete coverage path planning for tetriamond tiling robots," *Sensors*, vol. 20, no. 2, p. 445, Jan. 2020.
[32] A. Le, V. Prabakaran, V. Sivanantham, and R. Mohan, "Modified a-star algorithm for efficient coverage path planning in Tetris inspired self-reconfigurable robot with integrated laser sensor," *Sensors*, vol. 18, no. 8, p. 2585, Aug. 2018.
[33] A. V. Le, A. A. Hayat, M. R. Elara, N. H. K. Nhan, and K. Prathap, "Reconfigurable pavement sweeping robot and pedestrian cohabitant framework by vision techniques," *IEEE Access*, vol. 7, pp. 159402–159414, 2019.
[34] L. Yi, A. V. Le, A. A. Hayat, C. S. C. S. Borusu, R. E. Mohan, N. H. K. Nhan, and P. Kandasamy, "Reconfiguration during locomotion by pavement sweeping robot with feedback control from vision system," *IEEE Access*, vol. 8, pp. 113355–113370, 2020.
[35] Y. Shi, M. R. Elara, A. V. Le, V. Prabakaran, and K. L. Wood, "Path tracking control of self-reconfigurable robot hTetro with four differential drive units," *IEEE Robot. Autom. Lett.*, vol. 5, no. 3, pp. 3998–4005, Jul. 2020.
[36] M. A. V. J. Muthugala, A. V. Le, E. S. Cruz, M. R. Elara, P. Veerajagadheswar, and M. Kumar, "A self-organizing fuzzy logic classifier for benchmarking robot-aided blasting of ship hulls," *Sensors*, vol. 20, no. 11, p. 3215, Jun. 2020.
[37] A. K. Lakshmanan, R. E. Mohan, B. Ramalingam, A. V. Le, P. Veerajagadeshwar, K. Tiwari, and M. Ilyas, "Complete coverage path planning using reinforcement learning for tetromino based cleaning and maintenance robot," *Autom. Construct.*, vol. 112, Apr. 2020, Art. no. 103078.
[38] A. V. Le, R. Parween, R. E. Mohan, N. H. K. Nhan, and R. E. Abdulkader, "Optimization complete area coverage by reconfigurable hTrihex tiling robot," *Sensors*, vol. 20, no. 11, p. 3170, Jun. 2020.
[39] S. Kohlbrecher, O. von Stryk, J. Meyer, and U. Klingauf, "A flexible and scalable SLAM system with full 3D motion estimation," in *Proc. IEEE Int. Symp. Saf., Secur., Rescue Robot.*, Nov. 2011, pp. 155–160.
[40] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3645–3649.
[41] N. Wojke and A. Bewley, "Deep cosine metric learning for person re-identification," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2018, pp. 748–756.
[42] R. Kirby, "Social robot navigation," Ph.D. dissertation, Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, May 2010.

ANH VU LE received the B.S. degree in electronics and telecommunications from the Hanoi University of Technology, Vietnam, in 2007, and the Ph.D. degree in electronics and electrical from Dongguk University, South Korea, in 2015. He is currently with the Optoelectronics Research Group, Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam. He is also a Postdoctoral Research Fellow with ROAR Laboratory, Singapore University of Technology and Design. His current research interests include robotics vision, robot navigation, human detection, action recognition, feature matching, and 3D video processing.

BALAKRISHNAN RAMALINGAM received the B.Sc. and M.Sc. degrees in electronics from Bharathidasan University, Tiruchirapalli, India, in 2006 and 2009, respectively, and the Ph.D. degree in VLSI design from SASTRA University, Thanjavur, India. He is a Postdoctoral Research Fellow with ROAR Laboratory, Singapore University of Technology and Design, Singapore. His research interests include artificial intelligence, computer vision and embedded system design, reconfigurable hardware, information security, and multiprocessor system on chips.

BRAULIO FÉLIX GÓMEZ received the bachelor's degree in mechatronic and bachelor's degree in informatics engineering from the Technological Institute of Los Mochis, Mexico, in 2018 and 2019, respectively. He is a Research Assistant with ROAR Laboratory, Singapore University of Technology and Design (SUTD). His research interests include robotics vision, robotics control, and AI based robotics.

RAJESH ELARA MOHAN received the B.E. degree from Bharathiar University, India, in 2003, and the M.Sc. and Ph.D. degrees from Nanyang Technological University, in 2005 and 2012, respectively. He is currently an Assistant Professor with the Engineering Product Development Pillar, Singapore University of Technology and Design. He is also a Visiting Faculty Member with the International Design Institute, Zhejiang University, China. He has published over 80 papers in leading journals, books, and conferences. His research interests include robotics with an emphasis on self-reconfigurable platforms and research problems related to robot ergonomics and autonomous systems. He was a recipient of the SG Mark Design Award in 2016 and 2017, the ASEE Best of Design in Engineering Award in 2012, and the Tan Kah Kee Young Inventors' Award in 2010.

TRAN HOANG QUANG MINH received the B.S. degree in electronics and telecommunications from the Hanoi University of Technology, Vietnam, in 2007, and the Ph.D. degree in electronics and electrical from Dongguk University, South Korea, in 2015. He is currently with the Optoelectronics Research Group, Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam. He is also a Postdoctoral Research Fellow with ROAR Laboratory, Singapore University of Technology and Design. His current research interests include robotics vision, robot navigation, human detection, action recognition, feature matching, and 3D video processing.

VINU SIVANANTHAM received the B.Tech. degree in electronics and communication engineering from Amrita University, in 2017. From 2015 to 2017, he was a Junior Research Assistant with HuT Laboratory, where he was working on mechatronics and software development. He is currently a Visiting Fellow with the Singapore University of Technology and Design, where he is working on the mechatronics and software development of reconfigurable robots.