Fusion of Millimeter wave Radar and RGB-Depth sensors for assisted navigation of the visually impaired

Ningbo Long, Kaiwei Wang*, Ruiqi Cheng, Kailun Yang, Jian Bai
State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China

ABSTRACT

According to the data from the World Health Organization, 253 million people are estimated to be visually impaired worldwide, and 36 million are blind. It is very difficult for visually impaired people to perceive and avoid obstacles at a distance during their travelling. To address this problem, we propose a sensor fusion system, which combines the RGB-Depth sensor and millimeter wave radar sensor, to detect the surrounding obstacles. The range and velocity of multiple obstacles are acquired by the millimeter wave radar based on the principle of frequency modulated continuous wave. The positions of the obstacles are verified by the RGB-Depth sensor based on the contour extraction and MeanShift algorithm. The data fusion algorithm based on particle filters obtains accurate state estimation by fusing RGB-Depth data with millimeter wave radar data. The experiment results show that multiple obstacles with different ranges and angles are successfully detected by the proposed system. The measurement uncertainties are reduced by the data fusion system, meanwhile the effective detectable range is expanded compared to the detection with only RGB-Depth sensor. Moreover, the measurement results are stable when the illumination varies. As a wearable prototype, the sensor fusion system has the characteristics of versatility, portability and cost-effectiveness, which is very suitable for blind navigation application.

Keywords: visually impaired people, RGB-Depth sensor, millimeter wave radar, sensor fusion

1. INTRODUCTION

According to the data from the World Health Organization (WHO), 253 million people were estimated to be visually impaired worldwide, and 36 million were blind[1]. It is very difficult for visually impaired people (VIP) to perceive and avoid obstacles at a distance. To address this problem, we propose a sensor fusion system, which combines the RGB-Depth (RGB-D) sensor and millimeter wave (MMW) radar sensor, to perceive the surrounding obstacles and their positions.

The RGB-D sensors acquire both color and depth maps simultaneously at the rate of video frame, which revolutionizes the field of navigation assistance for VIP. The personal guidance devices based on RGB-D sensors provide much more information compared the traditional assistive tools. Assistive approaches based on RGB-D sensors have been widely investigated to help VIP to detect and avoid obstacles[2-8].

However, the RGB-D sensors, which includes light-coding sensors, time-of-flight sensors (TOF camera), and stereo cameras, could not solve the problems of remote obstacle detection and velocity detection perfectly. The detection range of low power light-coding sensors is too small, especially in sunny environments[2, 4, 9, 10]. The ranging precision of TOF camera is sensitive to ambient light and shows poor performance in outdoor environments[11-13]. The ranging results of remote objects derived from stereo cameras are not accurate, and the remote objects without texture are not robustly detected[14,15]. Furthermore, it is difficult to measure the velocity of object using all these kinds of RGB-D sensors.

The radar's performance is not influenced by illumination across indoor and outdoor environments[16, 17]. The range and velocity information of the obstacles are calculated at the same time, and the accuracy of the range is very high, e.g., several centimeters. Thanks to the technological development, the MMW radar sensors have become small, low-cost and accurate, which makes them especially suitable for portable low-power applications. Moreover, the single chip radar sensor has already appeared. However, the azimuth beam with of the MMW radar always covers more than several degrees due to the limited antenna distributions, which results in a low directional resolution compared with the camera. On the contrary, the directional resolution of the RGB-D sensor is very high, which makes it possible to be complemented through sensor fusion.

*wangkaiwei@zju.edu.cn; http://www.wangkaiwei.org/

In this paper, we propose a sensor fusion system based on the MMW radar and the RGB-D sensor to help VIP perceive and avoid the obstacles. As a wearable prototype, the sensor fusion system has the characteristics of versatility, portability and cost-effectiveness, which is very suitable for blind navigation application.

There are various advantages of fusing RGB-D sensor and radar sensor. The sensor fusion system is able to utilize the advantages of these two sensors and avoid their disadvantages. Firstly, the range information becomes more accurate compared to the detection with only RGB-D sensor, and the detectable distance is expanded with the help of the MMW radar. Secondly, the angle resolution is improved compared to the detection with only MMW radar. Thirdly, the velocity information could be easily measured compared the detection with RGB-D sensor. At last, the sensor fusion system increases the robustness of the overall system across to diverse lighting conditions.

The paper is organized as follows. In the next section, the hardware configuration of the fusion system is presented. In the section 3, the detection principle of the MMW radar and RGB-D sensor is described. The sensor fusion algorithm is described in detail in the section 4. After that, the experiments results are presented and discussed in section 5. Finally, conclusions are drawn in the section 6.

2. HARDWARE CONFIGURATION

In this paper, we propose the sensor fusion system to detect and perceive the surrounding obstacles based on the Intel RealSense D435[18] and the TI short range radar evaluation board[19], as shown in Figure 1. The size and power consumption of these two sensors is very small and suitable for being integrated in a wearable device. They are mounted on a bracket fabricated by 3D printing and their position is fixed. The sensors are closely spaced on the same plane, hence the sensor fusion is performed at the objects detection level. In view that the nature of radar measurements is coarse, the current fixed mode could achieve robust data fusion with sufficient precision.

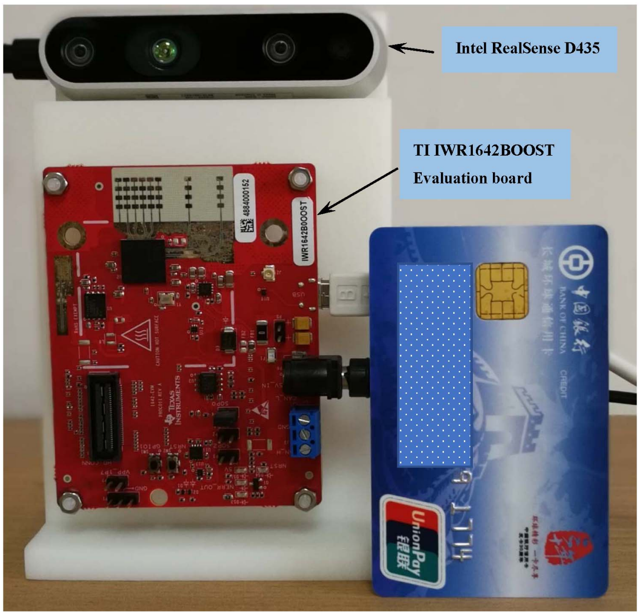

Figure 1. The hardware configuration. The fusion system is composed of an RGB-D sensor (Intel RealSense D435) and a short-range radar (TI IWR1642BOOST Evaluation board). As compared with a credit card, the size of the D435 and the MMW radar is so small that they are portable for visually impaired users.

The RealSense D435 developed by the Intel Company is the 2nd generation of the active stereo depth camera system. It offers complete depth cameras integrating vision processor, stereo depth module and RGB sensor with color image processing. The stereo vision module consists of a left imager, right imager and an infrared projector. The infrared projector projects invisible Infrared Radiation (IR) pattern to improve depth accuracy for scenes with low texture. With the principle of active projecting and passive stereo matching, the performance of RealSense D435 is excellent across indoor and outdoor circumstances. Therefore, the D435 is quite suitable for VIP navigation because of its environment adaptability, small size and cost-effectiveness.

However, the performance of the D435 is sensitive to lighting conditions. On the contrary, the range accurate of the MMW radar is high and the measurement results are stable. As shown in Figure 1, the TI IWR1642BOOST evaluation board, which is a short-range radar, is employed in our fusion system. The IWR1642 device is an integrated single-chip MMW sensor based on frequency-modulated continuous wave (FMCW) radar technology capable of operation in the 76 to 81 GHz band with up to 4 GHz continuous chirp. The IWR1642 includes a monolithic implementation of a 2TX, 4RX system with built-in PLL and A2D converters. The IWR1642 also integrates a DSP subsystem, which contains TI's high-performance C674x DSP for the Radio Frequency (RF) signal processing. The device includes an ARM R4F-based processor subsystem, which is responsible for front-end configuration, control, and calibration. The IWR1642 is an ideal solution for low-power, self-monitored, ultra-accurate radar systems in both industrial and consumer electronics applications.

3. OBSTACLE DETECTION

3.1 Image obstacle extraction

Unlike the ordinary digital image processing on the color image, this paper achieves the objects detection on the depth image produced by the RealSense D435. As the Figure 2 shown, the objects are detected in indoor and outside, the color images are shown in Figure 2(a), and the depth images are shown in Figure 2(b). The MeanShift algorithm[20] is applied to detect the objects in the depth images, the detection results are shown in Figure 2(c). The detection objects are indicated by the red bounding box. The distance of the detection object is decided by the average depth in the red bounding box. The world coordinate of the detection object is acquired with the help of the camera intrinsic and extrinsic parameters.

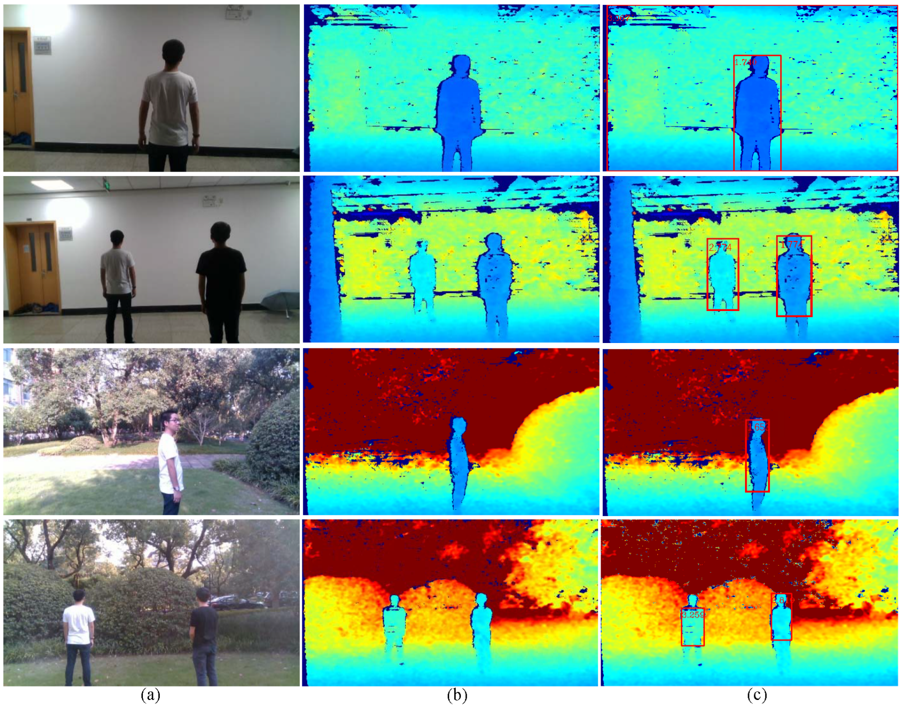

Figure 2. The color image, depth image and the feature extraction results of the objects. (a) The color images; the objects are detected in indoor and outside. (b) The depth images. (c) The detection results based on the MeanShift algorithm, where the detected objects are indicated using the red bounding box.

3.2 Frequency modulated continuous wave

FMCW radar is a technique that obtains range and velocity information from a radar by the way of frequency modulating a continuous signal[21]. The frequency modulation takes many forms, and the linear frequency modulation is the most commonly used approach. The basic principle of the sawtooth modulation FMCW radar and the objects detection algorithm have already been elaborated in many literatures[17,21-23]. Hence, we focus on the application of this low power

FMCW radar. The range, velocity and angle information of multiple target detection results are shown in Figure 3. The radar trihedral corners with different ranges and angles are placed in the courtyard as shown in Figure 3(a). The detection results with X-Z scatter mode and Doppler-Range mode are described in the (b) and (c) respectively.

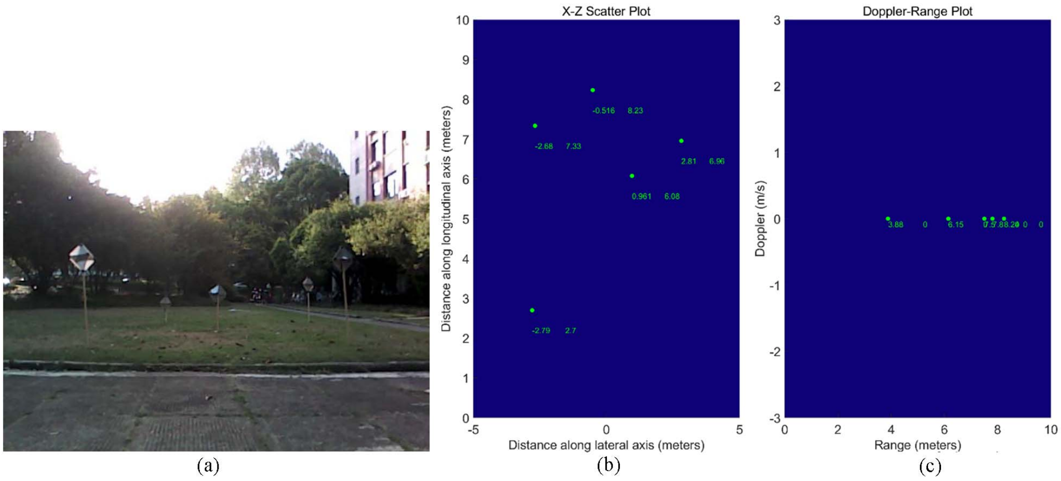

Figure 3. The multiple target detection result of the FMCW radar. (a) The radar trihedral corners placed in the courtyard with different ranges and angles. (b) The detection result shown with X-Z Scatter mode. (b) The detection result shown with Doppler-Range mode.

4. SENSOR FUSION

In multiple sensor system, each sensor performs measurements in its own coordinate system. Thus, one needs to transform these measurements into a global coordinate system.

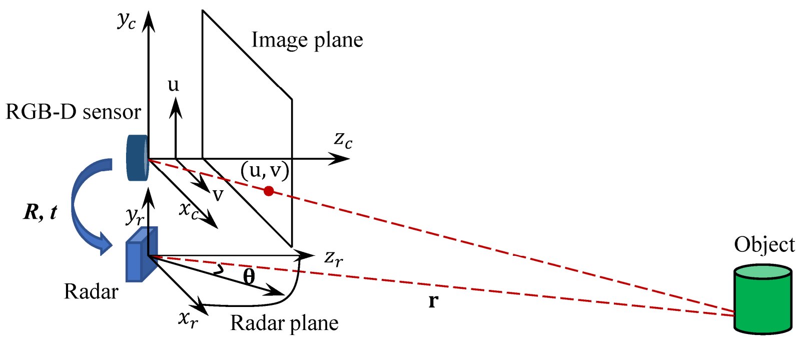

Figure 4. The RGB-D sensor coordinate and MMW radar coordinate

In this paper, the standard pinhole model is assumed for the camera. As shown in Figure 4, the (x_c, y_c, z_c) and (u, v) are the camera coordinate and the image plane coordinate respectively. The relationship between them is described as equation (1). The f_x, f_y, c_x, and c_y are the x, y direction focus length and principal point respectively, and the K is the intrinsic matrix.

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (1)$$

As described in Figure 4, the position of a target is represented as (x_r, y_r, z_r) in the radar coordinate system. The MMW radar only gives azimuth information in 2D without information on pitch angle in the 3D plane. In this sense, the

y_r coordinate has no meaningful value. The calibration matrix H^{RT} (including the rotation R and translation T) between the camera and the radar coordinate is obtained through

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = H^{RT} \begin{bmatrix} x_r \\ z_r \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_r \\ z_r \\ 1 \end{bmatrix}. \tag{2}$$

where the calibration matrix H^{RT} is made up of 9 elements.

As mentioned above, the object position in $x-z$ plane is measured by the two sensors. After calibration, we obtain two sets of measurements on the common plane. Considering the sensors resolution, noise and errors, we propose data fusion based on Particle Filter[24-25] to reduce uncertainties and obtain more accurate state estimation.

The target fusion task from different sensors is generally solved by the method of Kalman Filter. However, the Kalman Filter is unimodal, and the Joint Integrated Probabilistic Data Association (JIPDA) algorithm[26-27] is needed to track and label this multiple target before performing the Kalman Filter, which makes the full process very complex. Following this rationale, the Particle Filter based on the principle of Monte Carlo sampling is used in our application to accomplish the multiple object fusion. The Particle Filter is multimodal which is able to track more than one object simultaneously. It is also suitable when the multivariate, nonlinear behavior and non-Gaussian noise situation appeared.

5. EXPERIMENT

Before the experiment, the sensor features need to be considered. The RealSense D435 uses a combination of active projecting and passive stereo matching. The color camera image is up to 1280×1080 RGB resolution, and the depth image is up to 1280×720 stereo depth resolution. The experiments are carried out at the resolution of 640×360. The horizontal and the vertical field of view (FOV) of the infrared imager are 91.2° and 65.5°, respectively. The effective depth range of the RGB-D sensor is 0.2m to about 10m, which varies with lighting conditions. The horizontal and the vertical FOV of the color camera are 69.4° and 42.5°. The stereo image synchronization is executed internally, while the results are recorded in the auto-exposure mode. Comparatively, the antenna FOV of the MMW radar (IWR1642EVM) is $\pm 60^{\circ}$ with angular resolution of approximately 15°.

We use the linear least square method with the help of 57 data points to estimate the calibration matrix H^{RT}. The corresponding pixels on the depth image and radar detection points are matched manually, then the matched data point is acquired, as the Figure 5 shown. The 5 targets are placed at different heights and depths, as the color image shown in the Figure 5(a). They are detected by the RGB-D sensor and MMW radar simultaneously, and the detection results are shown in Figure 5(b) and (c) respectively. By changing the relative position between the sensors and radar trihedral corners, we get sufficient corresponding data points. In our paper, the calibration matrix H^{RT} between the RGB-D sensor coordinate and the MMW radar coordinate is

$$H^{RT} = \begin{bmatrix} 0.8650 & -0.0255 & 0.0605 \\ 0.0116 & 0.0247 & -0.1567 \\ 0.0296 & 0.7859 & 0.2759 \end{bmatrix}. \tag{3}$$

In order to test and verify the performance of our sensor fusion system, the field test is designed and performed. As shown in Figure 6, the color images are shown in Figure 6(a1-a3), and the depth images are shown in the Figure 6(b1-b3). The detection and fusion results are shown in Figure 6(c1-c3). The detection obstacles of the RGB-D sensor are indicated by the red bounding box, and the corresponding detection results by the MMW radar are indicated by the green vertical line. The fusion results, x, y, z coordinate and velocity information, are labelled in the depth images. In the scenario 1, two cars appear in the scene, as shown in Figure 6(a1). However, they are not able to be detected in the depth image because of the limited perception range of the RGB-D sensor, as shown in Figure 6(b1, c1). Still, the obstacles are correctly detected with the help of the MMW radar, and the detection results except the y coordinate are mapped and labelled in the depth image. In this regard, the effective detection range of the fusion system has been significantly expanded compared to the detection with only RGB-D sensor.

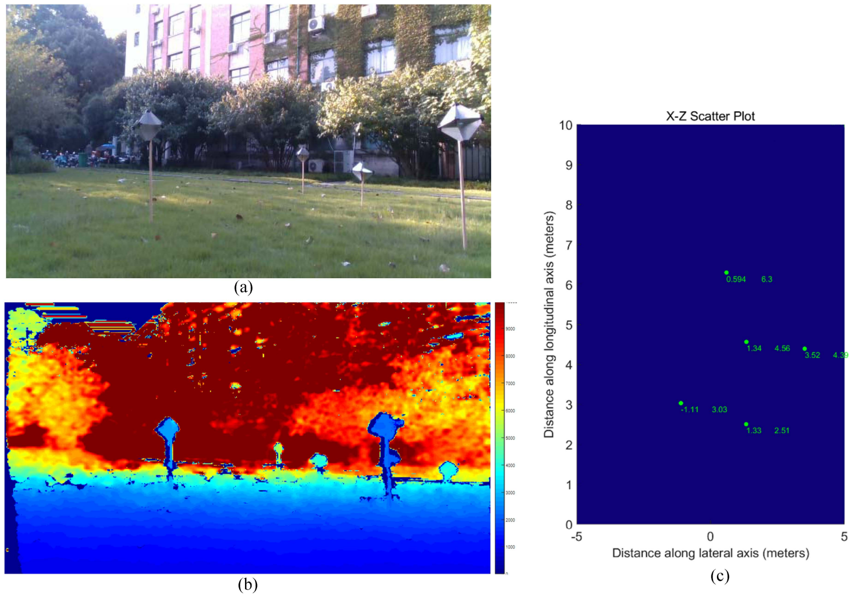

Figure 5. The processing procedure of the matched data point. (a)The color image including 5 targets which are placed at different height and depths. (b)The RGB-D sensor detection results. (c) The MMW radar detection results. The corresponding pixels on the depth image and radar points are matched manually, then the matched data points are acquired.

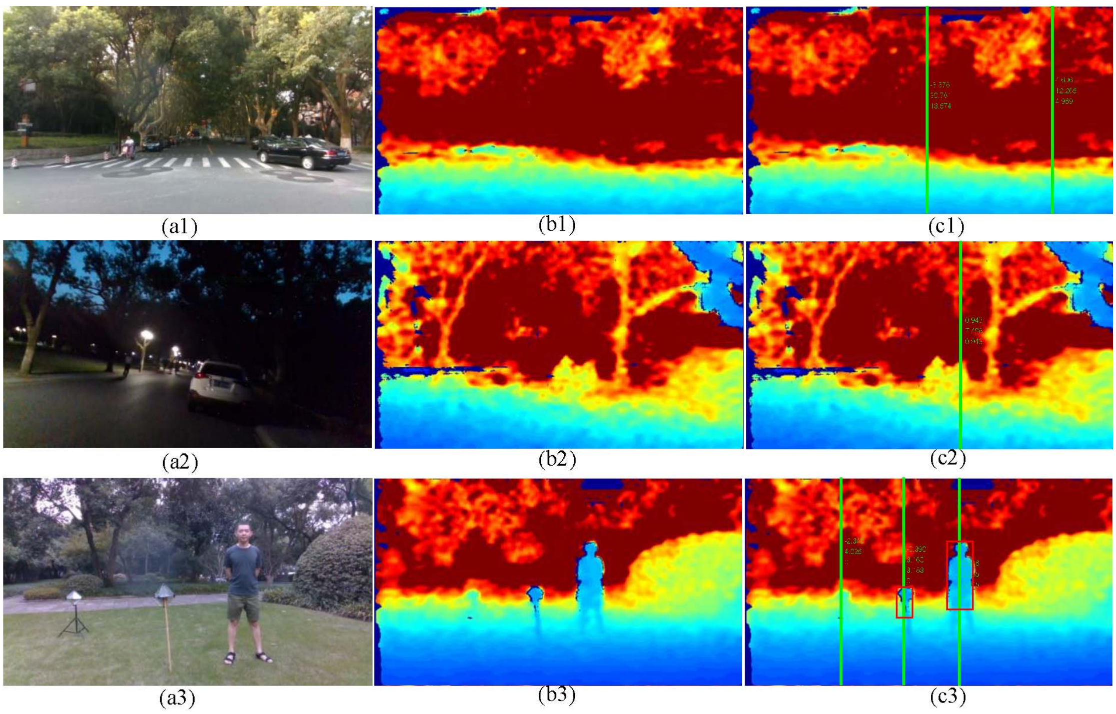

Figure 6. The field test. (a1-a3) The color images. (b1-b3) The depth images. (c1-c3) The detection and fusion results. The detection obstacles of the RGB-D sensor are indicated by the red bounding box, and the corresponding detection results by the MMW radar are indicated by the green vertical line. The fusion results are labelled in the depth images.

In the second scenario, a car is parked on the roadside at nightfall, as presented in Figure 6(a2). It is not found in the depth image because the quality of RGB-D depth image declines when the lights dim and the measurement distance becomes far, as shown in Figure 6(b2). However, the MMW radar detection result is stable when the illumination changes. It is mapped and labelled in the depth image, as shown in Figure 6(c2). In this sense, our fusion system enhances the robustness of obstacle detection across different illumination conditions.

In the third scenario, two radar trihedral corners and a person stand in the courtyard, as shown in Figure 6(a3). One of the radar trihedral corners and the person are found out in the depth image, as presented in Figure 6(b3). Their position, x, y, z coordinate, and velocity information are decided by the sensor fusion, then they are mapped and labelled in the depth image, as shown in Figure 6(c3). The y coordinate is added compared to the detection with only radar. The velocity information is added and the z coordinate is more accurate compared to the detection with only RGB-D sensor. The other radar trihedral corner is submerged in the background because it exceeds the effective distance of the RGB-D sensor. However, we could also find it, the x, z coordinate and velocity information is got with the help of the MMW radar. Compared with the single RGB-D sensor or the single MMW radar, the data fusion enriches the detection results, increases the accuracy and improves the robustness of the prototype.

6. CONCLUSION

In this paper, we proposed a sensor fusion system based on the RGB-D sensor and the low power MMW radar sensor to help visually impaired people to perceive and avoid obstacles at a distance. The experiment results show the multiple target with different ranges and angles are detected by the MMW radar and the RGB-D sensor. The data fusion based on particle filters reduce uncertainties and obtain more accurate state estimation. Meanwhile, the effective detectable range of the fusion system is expanded compared to the detection with only RGB-D sensor. Moreover, the measurement results maintain stable when the illumination is changed.

As a wearable system, the sensor fusion system has the characteristics of versatility, portability and cost-effectiveness, which is very suitable for blind navigation application. Moreover, this system could be flexibly applied in the fields of self-driving, unmanned aerial vehicle (UAV), robotics, surveillance and defence.

REFERENCES

[1] Bourne, Rupert RA, Seth R. Flaxman, Tasanee Braithwaite, Maria V. Cicinelli, Aditi Das, Jost B. Jonas, Jill Keeffe et al. "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis." The Lancet Global Health 5, 9: e888-e897 (2017).

[2] R. Cheng, K. Wang, K. Yang et al., "A ground and obstacle detection algorithm for the visually impaired," 2015 IET International Conference on Biomedical Image and Signal Processing (ICBISP 2015).

[3] R. Cheng, K. Wang, K. Yang et al., "Real-time pedestrian crossing lights detection algorithm for the visually impaired," Multimedia Tools and Applications. Papers 1-21, (2017).

[4] Yang, Kailun, Kaiwei Wang, Luis M. Bergasa, Eduardo Romera, Weijian Hu, Dongming Sun, Junwei Sun, Ruiqi Cheng, Tianxue Chen, and Elena López. "Unifying Terrain Awareness for the Visually Impaired through Real-Time Semantic Segmentation." Sensors 18, 5, 1506 (2018).

[5] X. Zhao, K. Wang, K. Yang et al., "Unconstrained face detection and recognition based on RGB-D camera for the visually impaired," Eighth International Conference on Graphic and Image Processing (ICGIP 2016). Vol. 10225. International Society for Optics and Photonics, 2017.

[6] K. Yang, K. Wang, R. Cheng et al., "Detecting Traversable Area and Water Hazards for the Visually Impaired with a pRGB-D Sensor," Sensors, 17(8), 1890 (2017).

[7] K. Yang, K. Wang, X. Zhao et al., "IR stereo RealSense: Decreasing minimum range of navigational assistance for visually impaired individuals," Journal of Ambient Intelligence and Smart Environments, 9(6), 743-755 (2017).

[8] R. Cheng, K. Wang, K. Yang et al., "Crosswalk navigation for people with visual impairments on a wearable device," Journal of Electronic Imaging, 26(5), 1-12 (2017).

[9] Yang, Kailun, Kaiwei Wang, Weijian Hu, and Jian Bai. "Expanding the detection of traversable area with RealSense for the visually impaired." Sensors 16(11), 1954 (2016).

[10] Zöllner, Michael, Stephan Huber, Hans-Christian Jetter, and Harald Reiterer. "NAVI–a proof-of-concept of a mobile navigational aid for visually impaired based on the Microsoft Kinect." In IFIP Conference on Human-Computer Interaction, 584-587 (2011).

[11] Zeng, Limin, Denise Prescher, and Gerhard Weber. "Exploration and avoidance of surrounding obstacles for the visually impaired." Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility. ACM, 111-118 (2012).
[12] Tamjidi, Amirhossein, Cang Ye, and Soonhac Hong. "6-DOF pose estimation of a portable navigation aid for the visually impaired." Robotic and Sensors Environments (ROSE), 2013 IEEE International Symposium on. IEEE, 178-183 (2013).
[13] Lee, Chia-Hsiang, Yu-Chi Su, and Liang-Gee Chen. "An intelligent depth-based obstacle detection system for visually-impaired aid applications." Image Analysis for Multimedia Interactive Services (WIAMIS), 2012 13th International Workshop on. IEEE, 1-4 (2012).
[14] Martinez, Juan Manuel Saez, and Francisco Escolano Ruiz. "Stereo-based aerial obstacle detection for the visually impaired." Workshop on Computer Vision Applications for the Visually Impaired. (2008).
[15] Lin, Kai Wun, et al. "A wearable stereo vision system for visually impaired." Mechatronics and Automation (ICMA), 2012 International Conference on. IEEE, 1423-1428 (2012).
[16] G. L. Charvat, [Small and Short-Range Radar Systems], CRC Press, New York, 40-43 (2014).
[17] Kim, Jong-Hwan, Joseph W. Starr, and Brian Y. Lattimer. "Firefighting robot stereo infrared vision and radar sensor fusion for imaging through smoke." Fire Technology. 51(4) 823-845 (2015).
[18] Keselman, Leonid, John Iselin Woodfill, Anders Grunnet-Jepsen, and Achintya Bhowmik. "Intel realsense stereoscopic depth cameras." arXiv preprint arXiv.1705.05548 (2017).
[19] TI, "IWR1642 Single-Chip 76-to-81GHz mmWave Sensor Integrating DSP and MCU Evaluation Module," 20 July 2018, http://www.ti.com/tool/iwr1642boost (20 July 2018).
[20] Yang, Kailun, Kaiwei Wang, Hao Chen, and Jian Bai. "Reducing the minimum range of an RGB-depth sensor to aid navigation in visually impaired individuals." Applied optics 57(11), 2809-2819 (2018).
[21] Stove, Andrew G. "Linear FMCW radar techniques." IEEE Proceedings F (Radar and Signal Processing). 139(5), (1992).
[22] Molchanov, Pavlo, et al. "Short-range FMCW monopulse radar for hand-gesture sensing." Radar Conference, IEEE, 1491-1496 (2015).
[23] Rohling, Hermann, and Marc-Michael Meinecke. "Waveform design principles for automotive radar systems." Radar, 4(1) (2001).
[24] Arulampalam, M. Sanjeev, Simon Maskell, Neil Gordon, and Tim Clapp. "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking." IEEE Transactions on signal processing. 50(2), 174-188 (2002).
[25] Zhou, Shaohua Kevin, Rama Chellappa, and Baback Moghaddam. "Visual tracking and recognition using appearance-adaptive models in particle filters." IEEE Transactions on Image Processing. 13(11), 1491-1506 (2004).
[26] Kim, Du Yong, and Moongu Jeon. "Data fusion of radar and image measurements for multi-object tracking via Kalman filtering." Information Sciences. 641-652 (2014).
[27] Ćesić, Josip, Ivan Marković, Igor Cvišić, and Ivan Petrović. "Radar and stereo vision fusion for multitarget tracking on the special Euclidean group." Robotics and Autonomous Systems. 83, 338-348 (2016).