2021 IEEE International Conference on Progress in Informatics and Computing (PIC) | 978-1-6654-2655-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/PIC53636.2021.9687018

A Multi-Sensory Blind Guidance System Based on YOLO and ORB-SLAM

Chufan Rui, Yichen Liu, Junru Shen, Zhaobin Li, and Zaipeng Xie*
Key Laboratory of Water Big Data Technology of Ministry of Water Resources
Hohai University, Nanjing, China
zaipengxie@hhu.edu.cn
**corresponding author*

***Abstract*—Blind guidance system has always been a research hotspot for years. Although there are many kinds of blind guidance systems on the market, most of them prompt from the perspective of a single sense of tactile or auditory. The blind guidance method of single sense can be unstable and it does not fully mobilize other general senses of the with vision impairment. This paper designs and implements a multi-sensory blind guidance system that provides tactile and auditory sensations by using the ORB-SLAM and YOLO techniques. Based on the RGB-D camera, the local obstacle avoidance system is realized at the tactile level through the point cloud filtering that feedback the results to the user through vibrating motors. The improved ORB-SLAM can generate a dense navigation map to implement a global obstacle avoidance system through the coordinate transformation. Real-time target detection and the YOLO-based prompt voice system is implemented at the auditory level. The system can detect the specific category and give the location of obstacles as real-time voice messages. The functions mentioned above are integrated and verified as a smart cane. Experimental results show that the position and category of the obstacles in the surrounding environment can be detected accurately in real-time through our system. By combining YOLO and ORB- SLAM, we can provide a piece of useful auxiliary equipment to the community of vision impairment and enable users to move about safely.**

***Index Terms*—multi-sensory guidance system, ORB-SLAM, YOLO, obstacle detection and avoidance**

I. Introduction

In recent years, research on blind guidance systems has been a hot field [1] [2] [3]. Vision loss can confine a person with vision impairment, restricting traveling and the ability to detect obstacles [4]. Therefore, a blind person typically requires the support of a wide range of tools and techniques [5]. Conventional methods to facilitate a visually impaired person are to wear a cane or be accompanied by a guide dog [6]. However, the conventional cane is challenging in providing adequate information about the environment [7]. Furthermore, the techniques for cane travel can vary depending on the user or the situation. The helpfulness of guide dogs is limited by the inability of dogs to understand complex directions. Dogs may require extensive training, and their service life cannot exceed several years, which may not be an optimized solution to the blind community [8]. Some blind guidance system utilizes GPS positioning [9] [10], UWB indoor positioning [11], and ultrasonic positioning [12], but these methods depend on the environment and have several limitations. Traditional GPS positioning cannot provide the indoor navigation of the blind, especially when there are poor GPS satellite signals [13]. UWB indoor positioning needs to deploy positioning anchor points, which may not be feasible for the outdoor navigation requirement in the city [14]. The cost of ultrasonic hardware can be high and may also be limited by acoustic effects [15].

The guidance system based on SLAM can be a feasible solution. Simultaneous localization and mapping technology (SLAM) [16] can simultaneously determine the position and orientation of the sensor relative to the surrounding environment and map the surrounding environment. Visual SLAM is a particular SLAM system, and it can implement the function of SLAM positioning and mapping when the environment and sensors are unknown by using 3D vision. For instance, ORB-SLAM [17] is one of the visual SLAM algorithms based on image ORB features. The ORB method based on the FAST and BRIEF algorithm can reasonably deal with the camera shaking during the walking and has little effect on the change of each pixel position feature vector.

The existing research on blind navigation technology generally combines visual SLAM with other optimization methods. For example, Nadia Kanwal et al. [18] used Microsoft Kinect somatosensory peripherals to combine the RGB-D camera of Kinect with the image information obtained by infrared sensors. Jinqiang Bai et al. [19] integrated a virtual blind road following method based on ORB-SLAM to provide automatic navigation for visually impaired people. The technique relies on optical perspective glasses to feedback the navigation results through vision and hearing. Zhuo Chen et al. [20] combined the semantic segmentation method to obtain semantic information in the environment and provide voice feedback to the blind. However, these three methods mainly utilize the auditory sense for blind navigation, and the effectiveness can be limited in a noisy environment. If such a system can use tactile and auditory senses simultaneously, it may be more intuitive to facilitate the mobility of blind users.

This paper proposes a multi-sensory blind guidance system by combining ORB-SLAM [17], YOLO target detection technology [21], RGB-D camera, and computer vision technology. Based on the research of the existing visual SLAM navigation system, the system combines the local navigation results obtained by detecting the surrounding obstacles through the RGB-D camera with the global navigation results obtained by synchronously positioning the ORB-SLAM on the map and transmits the obstacle information to the blind through a vibration prompt method. In addition, the system can recognize the obstacle type information combined with YOLO target detection and transmit the obstacle information in the global navigation direction through voice. The combination of the two produces a multi-sensory blind guidance system. After implementation, the accuracy of the system was evaluated through experiments.

The main contributions of this paper are as follows:

- This paper evaluates the great potential of combining YOLO and ORB-SLAM by analyzing the actual needs of the blind guidance system. Therefore, a multi-sensory blind guidance system combining the touch and the hearing was proposed and implemented. The experiment shows that the proposed system can achieve a reliable guide effect.
- This paper proposes improving the algorithm to improve the accuracy by generating a dense navigation map. The system utilizes the position information from the ORB-SLAM and combines it with the Bresenham algorithm [22], the ideal dense navigation map can be obtained through transforming coordinates, "emitting light" to feature points and other steps.
- This paper also proposes a local obstacle avoidance algorithm based on an RGB-D camera, which can identify short-range obstacles through point cloud filtering and angle sampling. The integration of the techniques has provided a piece of useful auxiliary equipment to the community of vision impairment and enabled a blind person to move about safely.

The rest parts of the paper are arranged as follows: Section II introduces the specific implementation of the multi-sensory blind guidance system, and Section III introduces our experiment and test results of the system. Section IV is our conclusion.

II. Architecture and Methods

A. System Architecture

The proposed blind guidance system comprises three layers, including device, edge, and cloud layers, as shown in Fig. 1.

The device layer is a guide crutch. Sensors, interactive devices, and the controller are embedded in it. The sensor part of the device is an RGB-D camera. The camera is used to capture the image and distance information of objects in front of the user. The controller transmits data from the RGB-D camera and audio input device to the edge layer for further processing. After the return of data, the controller controls interactive devices to output obstacle prompts. The interactive devices contain an audio input/output device and a vibrating motor array. The vibrating motor array prompts obstacles in the user's front, left, right, upper and lower sides. The audio output device outputs a 3D sound field. Therefore, the user can understand the environment and obstacles by sensing the direction of sounds. The audio input device receives voice commands from the user.

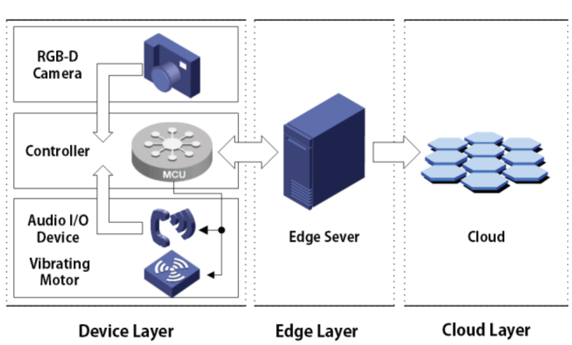

Fig. 1: The architecture of the blind guidance system.

The edge layer is a base station. The base station receives images and audio data from the controller. When realizing real-time navigation and obstacle avoidance, the base station generates maps and obstacle information by processing image data. Secondly, the base station processes audio data using the speech recognition algorithm and feeds the result back to the device layer.

In this system, the cloud layer regularly receives maps uploaded by the base station and saves backup copies of the maps in case of data loss. This system is divided into three subsystems: local obstacle avoidance system, global obstacle avoidance system, and object detection and voice prompt system.

The local obstacle avoidance system is used to detect surrounding obstacles. The global obstacle avoidance system is based on local obstacle avoidance and provides real-time navigation for the blind. The object detection and voice prompt system can identify obstacles and generate feedback via voice message. In actual application scenarios, there're two modes. The first mode is the free-walking mode which mainly uses a local obstacle avoidance system. The second mode is real-time navigation mode which uses the global obstacle avoidance system. Both modes use object detection and voice prompt systems.

In this paper, algorithm modules applied in the system are local obstacle avoidance system based on RGB-D, global obstacle avoidance system based on RGB-D, global obstacle avoidance systems based on improved ORB-SLAM, object detection, and voice prompt based on YOLO. The local obstacle avoidance system uses depth images from RGB-D cameras to calculate the distance between obstacles and the device. Based on distance data, the system can notify the user whether there is any obstacle ahead. The global obstacle avoidance system uses ORB-SLAM to plan a path without obstacles for the blind. The object detection and voice prompt system uses the YOLO algorithm to process RGB images to identify objects and prompt the user via voice. The specific algorithm modules are introduced below.

B. Local Obstacle Avoidance System Based on RGB-D

The local obstacle avoidance system can detect and prompt the obstacles around the blind who walk freely. It is the primary function of the blind guidance system for the blind

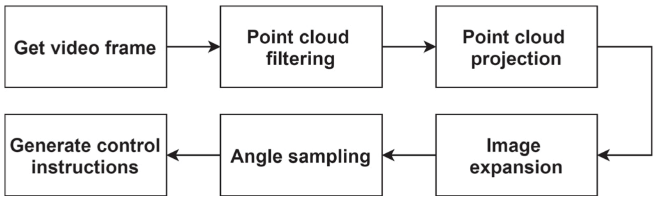

Fig. 2: Core process of obstacle detection in front and below.

when walking without a destination. The camera is the eye for the blind to achieve local obstacle avoidance. In order to more accurately perceive the distance between obstacles and meet the needs of blind people traveling at night, the RGB-D camera [23] was chosen, rather than a monocular or binocular. This paper uses two RGB-D cameras to detect obstacles in front of and below the blind, respectively. The detection and prompt are carried out according to the process in Fig. 2.

Firstly, the depth image video frame is obtained by RGB-D cameras. A depth image is a single channel image. Each pixel position stores a 16 bit unsigned integer, which is the scaled distance value, indicating the distance from the camera to the point. If there is no object at a pixel position, the distance value at that position is 0. Then, filtering technology is used to process the obtained point cloud data further. This paper uses the following methods for filtering.

1. Direct filtering. The points whose depth exceeds the obstacle avoidance detection distance threshold (the RGB-D camera in front is set to 3m) are directly removed. The points whose height exceeds a certain threshold or whose height is lower than a certain threshold are also directly removed.

2. Statistical filter. When a point cloud somewhere is less than a specific density, the point cloud is invalid. When implemented, the average distance from each point to the nearest k points around it is calculated. Remove points where the average value is outside the specified standard range.

Therefore, the threshold of the lower RGB-D camera depth is calculated as:

$$\frac{H-h}{\sin\theta} \tag{1}$$

where h is the height of the obstacle, θ is the deflection angle of the camera, and H represents the distance from the crutch to the ground.

After filtering, the point cloud projection can be obtained. For each point (x, depth), map it to (x, [depth × Scale]), where scale can control the resolution, and [·] represents rounding down. This paper uses the matrix mat of OpenCV to store the filtered point cloud projection as a plane map for local obstacle avoidance. The point mapped to is the point on the obstacle.

In order to avoid obstacles better, the image needs to be expanded. The effect of expansion is to make the obstacle appear to expand a circle and fully surround the boundary to prevent the user from hitting the boundary of the obstacle as much as possible during planning.

After that, angle sampling is performed. This paper assumes that the motion model of the human body is linear. Starting from the camera position, n rays in different directions are emitted as sampling rays. The system divides each sampling ray into feasible and infeasible rays according to the ratio of obstacle points to total points for each sampling ray.

Finally, control instructions are generated according to the sampling results. The algorithm divides the 180-degree area in front of the blind into several regions. The ratio of the number of feasible sampling rays to the total number of sampled rays in the region is calculated. Once the proportion is lower than the threshold, it is determined that there are obstacles in the area. Then the vibrating motor in the corresponding area prompts the user.

C. Global Obstacle Avoidance System Based on Improved ORB-SLAM

Global obstacle avoidance means that the system carries out path planning for the blind, avoids obstacles in the current area, and guides them to the end. The global obstacle avoidance system adds the global path planning results based on local obstacle avoidance. In this situation, the function of the vibrating motor is to indicate the direction that should go forward. In order to realize the global obstacle avoidance of blind guidance system, this paper improves the existing ORB-SLAM.

SLAM is a mapping technology that can gradually depict a complete environment map with movement in an unknown environment. The system mainly uses the ”real-time positioning” feature of SLAM to determine the current position and orientation of the blind on the map. The system makes further decisions by comparing the image in the camera's field of view with the key frame of the map. Finally, a vibration motor is used to indicate the direction the blind person should go. On the other hand, when a blind person walks by if the map of the scene itself does not exist or is incomplete, the mapping feature of SLAM enables the map to be continuously updated incrementally [24]. In this way, the next blind person can get a complete map after walking through a similar area. In addition, based on the map method, the map key frames of similar regions can be merged using the algorithm.

This paper improves ORB-SLAM because of its lack of dense navigation maps. Obstacle points and unexplored points are classified as non-passable points while other points as passable points. Then a point cloud set is used to save a dense map.

ORB-SLAM will provide us with the camera's position information, and the actual position in the plane can be obtained through coordinate transformation. At the same time, ORB-SLAM will also provide us with the orientation of feature points in the current key frame. For all feature points on the map, the readers can suppose that light is emitted to these feature points from the current camera position. The end of each ray is a feature point, which represents the boundary between two objects or parts. Therefore, in the current key frame, it can be considered that all points between the light source and the endpoint are passable points in this key frame. Therefore, add one to the ”passability” of the point on each light (except the endpoint). Since the feature point itself must be the point on the obstacle, add one to the 'impassability' at the feature point. In the system, the points are actually grid points. Since the grid points are all integer points, the Bresenham algorithm is used to obtain the integer points passed by the light between the current camera position and the feature points. After each calculation, count the 'passability' and 'impassability' of each grid point. If the ratio of the two exceeds a certain threshold, it indicates that it is passable, otherwise, it is not passable. If the passability is 0, it also indicates that it is not passable (unexplored point).

The dense map is constructed by this method. At first, the map is only composed of some radial rays. However, with the progress of users, all parts of the map will be gradually supplemented. Finally, a Gaussian kernel is used to filter the map to reduce the noise. This approach is an improvement on the lack of dense maps in ORB-SLAM itself. Finally, according to the generated dense navigation map, the blind can be prompted with direction through the vibrating motor, and global obstacle avoidance is realized.

D. Object Detection and Voice Prompt Based on YOLO

The target of object detection is to find every object and determine the types and positions of the things in the image. In this system, the object detection algorithm processes RGB images from RGB-D cameras and identifies their types. When manipulating the device, the system will process images and prompt the user whether there is any obstacle by voice as a part of the obstacle avoidance function.

Object detection algorithms can be divided into two categories, namely one-stage and two-stage algorithms. Because the object detection speed of the two-stage algorithm is slow, it cannot meet the real-time requirements of the system. The one-stage algorithm has a higher speed and lower background error detection rate than the two-stage algorithm. Thus one-stage algorithm has higher practical value for the guide equipment which requires strong real-time performance. Therefore, the object detection algorithm deployed in this system uses the YOLO algorithm of the one-stage algorithm [21], You Look Only Once.

YOLO algorithm uses a convolutional neural network to extract features and get predicted values. The structure of the convolutional neural network consists of 24 convolutional layers and two fully connected layers. The convolution layer is used to extract the features of objects and the full connection layer is used to predict the types and positions of objects. In order to simplify the model and enhance the real-time performance, a lightweight YOLO is adopted in this system. This algorithm uses only nine convolution layers and two fully connected layers, and fewer convolution kernels are used in the convolution layer. [25]

YOLO algorithm firstly slices the image into the S*S grid. Then, the convolutional neural network is used for each grid to output B vectors, namely (x,y,h,w,c). In addition, the vector also contains a One-hot coded classification, which represents the types of objects in the grid. In the vector above, c represents confidence, and (x,y,h,w) represents the position and size of the five bounding boxes in this grid. Therefore, the output vector dimension of the full connection layer of the convolutional neural network is, and N indicates the number of category labels. The calculation method of confidence c is as follows:

$$c = P_r(o) \cdot IOU_{\text{truth}}^{\text{pred}} \tag{2}$$

where P_r indicates whether an object is detected in the grid. P_r is 1 if an object is detected, and 0 if not. And IOU(Intersection of Union) represents the cross-ratio of the actual bounding boxes to the predicted bounding boxes. The intersection of union refers to the ratio of area intersection and area union of two rectangular areas, which can well measure the similarity of two rectangular areas.

There may be overlapping bounding boxes of multiple prediction results in the prediction results output by a convolutional neural network. Therefore, non-maximum suppression (NMS) algorithm was adopted to remove redundant bounding boxes. The core idea of NMS is that for two bounding boxes A and B if categories are different or A and B do not intersect, they will be reserved. And if there is an intersection part, the bounding box with high confidence in A and B will be reserved.

In this system, the YOLO algorithm is deployed in the edge layer. After the base station finishes image processing, object type and location information are passed to the coordinator module on the device side. This module calls the name text of the object and 3D sound [26] [27] generation module and gives feedback to the user through 3D sound to inform the position and types of obstacles.

III. Experiments and Results

A. Platform and Environment

The test platform is divided into edge end and device end. Uzel US-M5422 edge server is used for edge end, and parameters are shown in Table 1. Raspberry Pi 4B is used as the test platform on the device side. The overall device end is shown in Fig. 3.

B. Performance of the System

Firstly, the local obstacle avoidance system was tested. As is shown in Fig. 4, the tester held the guide device. There was a chair in front of the subject as an obstacle.

Fig. 5 is the image of the depth camera showing obstacles detected by the RGB-D camera. In this image, the white part

TABLE I: **Parameters of the test platform**

Name	Parameter
CPU	Intel WL-U I3 8145UE
Memory	4GB DDR4L
AI Card	HDDL-R8
SSD	500GB M.2 2280
IDE	CLion

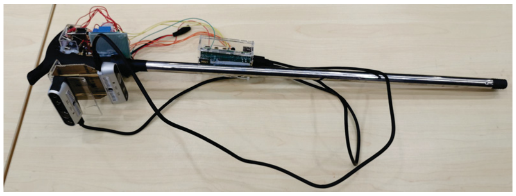

Fig. 3: The overall device.

Fig. 4: The tester holds the blind guide device.

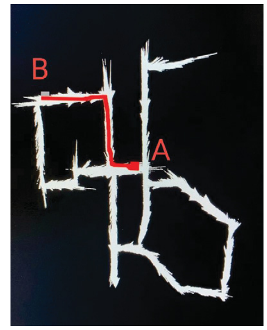

Fig. 7: Path planning result.

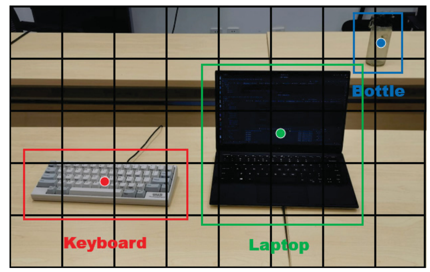

Fig. 8: Object detection result.

represents the obstacle. Blue and green lines are the sampling track, and green means the track is feasible. Fig. 6 shows the output of the vibrating motor array. In this figure, white means no vibration and red means vibration. The vibration position is the direction of the obstacle. Through the visual image of the vibrating motor, it can be seen that the system successfully detects the obstacle.

Secondly, the global obstacle avoidance system was tested. This system realizes navigation based on maps. When the user arrives at a strange area, the system will automatically establish a map of the area as the user explores and walks, which is convenient for the subsequent users.

The experiment used the KITTI-02 dataset as the real-time mapping test dataset. The initial point cloud image generated by ORB-SLAM is very sparse and difficult to navigate. With our improved ORB-SLAM algorithm, the point cloud is converted into a dense map. The researcher set the blind location as A and destination as B and obtained the global navigation diagram planned by the system, as shown in Fig. 7. The red square position is the current camera position. It can be found that the dense map can be constructed under the transformation of key frame points and the Bresenham algorithm, which significantly improves the accuracy of path planning.

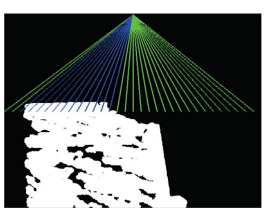

Fig. 5: The depth image of the RGB-D camera.

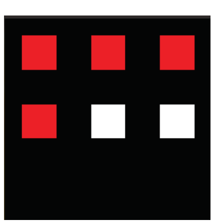

Fig. 6: The output of the vibrating motor array.

Finally, the object detection and voice prompt system was tested. The system detects the result as shown in Fig. 8. It can be seen that the system has obtained accurate detection results and made a correct judgment of the item category. Finally, the detection result will be voice feedback to the user.

As seen from the above testing process, our system can provide a primary obstacle avoidance guide for blind people with tactile signals based on vibration and feedback to blind people about the objects around and even the location of each object through the audio auditory signal. The whole system allows the blind to perceive the surrounding environment from both tactile and auditory aspects, with high detection accuracy and good target recognition ability, which can bring convenience for the blind to travel.

IV. Conclusion

This paper proposed a blind guidance system using both the tactile and auditory senses by combining ORB-SLAM and YOLO. An obstacle avoidance algorithm based on the depth camera was implemented to realize local obstacle avoidance. At the same time, an algorithm is developed to convert the sparse point cloud generated by ORB-SLAM into a dense navigation map. Additionally, image recognition based on the YOLO algorithm is utilized in the proposed system to detect the obstacle target in real-time, which can assist the travel of a visually impaired person through a voice prompt. This paper evaluates the robustness and feasibility of our system. Experimental results show that the combination of ORB-SLAM and YOLO can greatly improve the blind guidance system's reliability and bring much convenience to the visually impaired community.

ACKNOWLEDGMENT

This work was supported by the National College Student Innovation and Entrepreneurship Training Program of China.

REFERENCES

[1] E. Wise et al., "Indoor navigation for the blind and vision impaired: Where are we and where are we going?," 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2012, pp. 1-7.

[2] M. D. Mashat and A. A. Albani, "Intelligent Blind Cane System," 2017 UKSim-AMSS 19th International Conference on Computer Modelling & Simulation (UKSim), 2017, pp. 34-38.

[3] S. Khan, S. Nazir and H. U. Khan, "Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review," in IEEE Access, vol. 9, pp. 26712-26734, 2021.

[4] A. Krishnan, G. Deepakraj, N. Nishanth and K. M. Anandkumar, "Autonomous walking stick for the blind using echolocation and image processing," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016, pp. 13-16.

[5] S. R. A. W. Alwi and M. N. Ahmad, "Survey on outdoor navigation system needs for blind people," 2013 IEEE Student Conference on Research and Developement, 2013, pp. 144-148.

[6] K. Kumar, B. Champaty, K. Uvanesh, R. Chachan, K. Pal and A. Anis, "Development of an ultrasonic cane as a navigation aid for the blind people," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014, pp. 475-479.

[7] S. S. Bhat lawande, et al., "Ultrasonic spectacles and waist-belt for visually impaired and blind person," in Communications (NCC), 2012 National Conference on, 2012, pp. 1-4

[8] P. Valsecchi, et al., "Development of the at tachment bond in guide dogs," Applied animal behaviour science, vol. 123, pp. 43-50, 2010.

[9] G. I. Hapsari, G. A. Mutiara and D. T. Kusumah, "Smart cane location guide for blind using GPS," 2017 5th International Conference on Information and Communication Technology (ICoIC7), 2017, pp. 1-6.

[10] S. Liu et al., "Mind Controlled Vehicle Based on Lidar SLAM Navigation and SSVEP Technology," 2021 9th International Winter Conference on Brain-Computer Interface (BCI), 2021, pp. 1-4.

[11] Y. Xu, X. Chen, J. Cheng, Q. Zhao and Y. Wang, "Improving tightly-coupled model for indoor pedestrian navigation using foot-mounted IMU and UWB measurements," 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, 2016, pp. 1-5.

[12] N. H. B. A. Khafar, D. M. Soomro, R. A. Rahman, M. N. Abdullah, I. A. Soho and A. A. Rahimoon, "Intelligent Navigation System using Ultrasonic Sensor for Visually Impaired Person (VIP)," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), 2019, pp. 1-4.

[13] M. Ochiai, M. Fujii, A. Ito, Y. Watanabe and H. Hatano, "A study on indoor position estimation based on fingerprinting using GPS signals," 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2014, pp. 727-728.

[14] C. Fowler, J. Entzminger and J. Corum, "Assessment of ultra-wideband (UWB) technology," in IEEE Aerospace and Electronic Systems Magazine, vol. 5, no. 11, pp. 45-49, Nov. 1990.

[15] D. Tudor, L. Dobrescu and D. Dobrescu, "Ultrasonic electronic system for blind people navigation," 2015 E-Health and Bioengineering Conference (EHB), 2015, pp. 1-4.

[16] Y. Jia, X. Yan and Y. Xu, "A Survey of simultaneous localization and mapping for robot," 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2019, pp. 857-861.

[17] R. Mur-Artal, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," in IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, Oct. 2015.

[18] Nadia Kanwal, Erkan Bostanci, Keith Currie, Adrian F. Clark, "A Navigation System for the Visually Impaired: A Fusion of Vision and Depth Sensor", *Applied Bionics and Biomechanics*, vol. 2015, Article ID 479857, 16 pages, 2015.

[19] J. Bai, S. Lian, Z. Liu, K. Wang and D. Liu, "Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People," in IEEE Transactions on Consumer Electronics, vol. 64, no. 1, pp. 136-143, Feb. 2018.

[20] Z. Chen, X. Liu, M. Kojima, Q. Huang, and T. Arai, "A Wearable Navigation Device for Visually Impaired People Based on the Real-Time Semantic Visual SLAM System," Sensors, vol. 21, no. 4, p. 1536, Feb. 2021.

[21] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[22] J. E. Bresenham, "Algorithm for computer control of a digital plotter," in IBM Systems Journal, vol. 4, no. 1, pp. 25-30, 1965.

[23] K. Melbouci, S. N. Collette, V. Gay-Bellile, O. Ait-Aider and M. Dhome, "Model based RGBD SLAM," 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 2618-2622.

[24] A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 1052-1067, June 2007.

[25] R. Huang, J. Pedoeem, and Cuixian Chen. Yolo-lite: A real-time object detection algorithm optimized for non-gpu computers. In 2018 IEEE International Conference on Big Data (Big Data), 2018.

[26] R. Nicol and M. Emerit. 3d-sound reproduction over an extensive listening area: A hybrid method derived from holophony and ambisonic. In Proc Aes International Conference, 1999.

[27] V. Pulkki. Multichannel Sound Reproduction. Handbook of Signal Processing in Acoustics, 2008