OBSTACLE SEGMENTATION AND WARNING SYSTEM FOR VISUALLY IMPAIRED PEOPLE

Dr. Md. Mojahidul Islam[1], Dr. Md. Robiul Hoque[2], Md. Atiqur Rahman[3], Dr. Md. Imran Hossain[4]

[1]*Associate Professor, Dept. of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh*
[2]*Professor, Dept. of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh*
[3]*Associate Professor, Dept. of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh*
[4]*Associate Professor, Dept. of Information and Communication Engineering, Pabna University of Science and Technology, Pabna, Bangladesh*

Abstract

This work describes a wearable obstacle object segmentation and warning system that can improve the dynamism and safety of blind and visually impaired people, particularly in unknown environments. The system uses Simultaneous Localization and Mapping and semantic path planning to complete localization and navigation tasks while collaborating with the visually impaired user. The RGB-D sensor camera detects and identify the location obstacles, and then sends information about the obstacles to visually impaired people via various modalities such as voice, tactile, and vibration. The path and motion guidance are generated to direct the user to their desired destination. The hardware of the system is consists of RGB-D camera, a Raspberry Pi module, and an eyeglasses. The software modules are implemented in MATLAB and OpenCV, and the navigation system is tested with blindfolded observers. The system employs computer image processing and motion planning strategies to (1) determine the accessible space, (2) construct an easy and safe movement path in space, and (3) identify and recognized particular kinds of objects, such as an empty chair. All of this information is transmitted to those who are wearing the device via vibration. We present findings from user studies on lower and high-level activities such as avoiding collisions in a maze, locating a chair, and navigating an overcrowded space while avoiding people.

Keywords: *visually impaired people, RGB-D camera, obstacle detection, computer vision, segmentation*

---

1. INTRODUCTION

When we are traveling, such as when we go to the workplace from our residence every day, we can freely perceive the three-dimensional information of the environment and recognize the scenes and objects we see. This ability allows us to easily navigate complex places and avoid various obstacles in our daily lives. Even in unfamiliar scenes, we can complete precise navigation and walk freely. We usually take this ability for granted. However, for the visually impaired, they cannot perceive the world like we do, and move freely in various indoor and outdoor scenes.

Identifying obstacles is a significant aspect to take into account when developing mobility assistive devices for those who are blind. Current mobility assistive technology include this function as an obstacle detector while walking, whether indoors or outdoors. Detecting obstacles on the pathway walking area is critical to avoiding a collision with obstacles that might cause a person to slip and fall. Accidents are a major concern in an ageing population.

According to WHO global statistics, 188.5 million individuals have minor vision impairment, 217 million have medium to serious eyesight disability, and 36 million individuals are completely blind [1]. People with normal eyesight can orient themselves in physical space and move around with ease. However, even with the use of electronic travel aids and vision techniques, it is difficult for people who are blind or have significant visual impairment to navigate an unfamiliar environment. Vision impairment has a significant impact on people's lives, particularly those who can travel and recognize their surroundings individually. Vision impairment has an important effect on people's lives, particularly their ability to navigate and recognize their surroundings individually. Visually impaired people cannot obtain visual information and perceive the environment normally like people with eyesight. However, in modern society, vision is a vital channel for people to obtain information. Therefore, blind people face great difficulties in many aspects of life, including daily life, social interaction, work, and travel. Although my country's economy is developing rapidly, the protection of minority groups is not perfect. The infrastructure construction for the visually impaired is also imperfect. For example, there are many irrational designs for blind roads in my country, which also cause external difficulties for the blind. Apart from the success of healthcare, the field of neuroscience, and biological technologies in finding a permanent solution to vision impairment problems [2], electronic and technological advances in computers can offer assistive tools to enhance their standard of life and allow for greater integration into community.

The most commonly used travel aids for blind people including guide sticks and dogs. From both of these, most the blind people used guide sticks, but the guide sticks can only detect obstacles and road conditions at a close distance of 1-2m, and the information acquisition is very limited. At the same time, the guide stick occupies one hand of the blind, which also causes inconvenience to use. Furthermore, there are many blind individuals who pay great attention to protect their self-esteem and refuse to travel with a guide cane, which gives them a kind of blind label. Guide dogs have also brought many pain points. For example, the total number of assistance dogs in the country is extremely limited, and the training and everyday maintenance costs are high.

To address the previously mentioned requirements, the current effort presented an improved layout that combines guidance and identification capability into one working prototype. The gadget is based on computer-vision techniques because the vision sensor provides sufficient data while being lightweight and inexpensive in comparison with different sensors. The system can operate in outdoor as well as indoor environments. As soon as the device is turned on, the navigation component will begin to instruct those with visual impairments whenever they want to go somewhere. To prevent overload of data, the recognition engine turns on based on the user's preferences through an easy procedure of double-clicking on the cell phone screen. The prototype will be constructed and evaluated in environments that are both indoor and outdoor, and outcomes are expected to show satisfactory results in terms of both guidance and identification efficiency, as well as providing a good travelling experience for blind and visually impaired people.

The major implications of this study are outlined below:

• Create an inexpensive object identification system that utilizes depth based on pictures recognition of objects to provide obstacles class, the location, and direction information, thereby improving environment perception and promoting navigation.

• Develop an adaptable surface segmentation method that employs an adaptive threshold calculation technique and ground height consistency across adjacent frames.

• Provide an accessible direction searching technique to help blind people get to their destination. Government or private agencies that works for blind people to improve their lifestyle can also profit this study.

• Future research group can be used this work as a reference.

2. LITERATURE REVIEW

Compared with limited assistance, the need for blind people is increasing day by day. On the one hand, economic development has not only increased the income of everyone, but also the income of the blind, but the blind has not enjoyed the benefits brought about by economic growth like the discerning people. For example, most of the new high-tech products are designed for discerning people. However, the increased purchasing power of blind people does not give them full opportunities to consume. On the other hand, the development of smart phones and the Internet has enabled blind people to learn about this rich and colorful world through mobile phones and the Internet, and have more needs for travel and awareness of the world. They all want to be able to fully perceive the world, freely travel and explore their homes, just like the discerning people.

The exponential development of computer vision and deep learning makes efficient and accurate image detection and scene analysis possible. In recent years, with the advent of Convolutional Neural Networks (CNNs), various visual tasks, such as image recognition [3], object detection [4], semantic segmentation [5], instance segmentation [6] and the whole picture segmentation [7] have made breakthrough progress. This also brings new ideas to blind people's assistance, for example, image segmentation can help blind people determine the paths that they can walk, and can predict various terrains and obstacles. Table 2.1 describes the main user of the portable mobility assistive devices and its functionality.

The color-depth (RGB-Depth) sensor can be referred to as the RGB-D sensor for short. Its invention and application have also revolutionized visual aids for the blind. The current conventional RGB-D sensors include Microsoft Kinect [8], Intel RealSense [9] and Stereo Lab ZED [10]. The RGB-D sensor can simultaneously collect the color, distance, and even more information of the scene at a real-time or even super-real-time speed. In addition, the miniaturization, low power consumption, and low cost of RGB-D sensors are very suitable for integration into wearable devices, such as visual aid glasses for the blind as shown in Fig-1.

Fig -1: Visual aid glasses for the blind

The development of computer vision and RGB-D sensors has made new travel assistance for the blind attractive. On the one hand, integrating RGB-D sensors into wearable devices and equipped with computer vision environmental perception methods enables blind people to obtain information about the surrounding scenes and travel safely and independently, which has application and social value. On the other hand, the existing RGB D sensors have problems such as single mode and limited detection range, while the perception methods of computer vision transform from simple image data domains to complex real scenes. The real-time, robustness, and sufficiency have also been tested, and these problems have brought abundant scientific research space to the research of this subject. The research on this subject is based on this new paradigm: computer vision and RGB-D sensor-based assistance for the blind. Among them, this topic is mainly aimed at the perception of the environment of the blind when traveling, and aims to provide the blind assist system with fast, reliable and comprehensive scene analysis and scene understanding.

A study [11] of 57 blind or visually impaired people, their carers, and professionals in rehabilitation discovered that those with visual impairments need and anticipate an electronic assistance device with sufficient knowledge about their surroundings, a straightforward interface for users, small size, protection, and cost effectiveness. To meet the aforementioned requirements, numerous electronic aids [12-15] have been suggested in the last few years. In terms of overall classification, these types of designs fall into two different groups. One is for guidance/navigation, while the other is for identifying nearby barriers or objects. However, only a few of these devices combine both capabilities (navigation and recognition) and incorporate recognition algorithms into navigation systems [16]. Adding recognition and navigation functions to a single system may significantly enhance everyday travel. For instance, a system for navigation without a recognition function may interpret a chair as an obstacle when assisting a blind person in looking for a seat or determining no route for entering a room with doors that are closed. However, if the identification ability exists, it will assist people in finding a place to sit or opening the door to get into the room. Planning a path for the blind or visually impaired differs from that for robots. There are numerical optimal/suboptimal solutions for dealing with unpredictability and insufficient data in robots. However, as a human-centered system, the navigation system must strive to offer a pleasant and uniform user experience.

The motivation of this study is explained particularly in terms of hands-free gadgets for identifying obstacles while walking, because the current equipment are not fully optimized in numerous ways, such as::

- Not appropriate for practical or outdoor applications.
- Not inexpensive
- Not allowing effective signal analysis.
- Not fully integrated for greater dependability and long-term use.
- Not taking into account the necessary gait requirements for a multiple user's application.
- Not supporting a successful alarming system.

3. PROPOSED METHODOLOGY

This study develops and implements an assistance system that allows people with impaired vision to safely move about in indoors as well as outdoors. In addition, this assistance system may help those who are blind in recognizing their surroundings and obstacles, as well as guiding them in a safe and appropriate route to walk.

3.1 System Architecture

As shown in Fig- 3.1, the architecture of the system in this section includes a pair of wearable smart glasses, a pathfinder worn on the waist, a portable processor and a pair of bone conduction headphones. The framework has two processing threads, which continuously acquire images from the sensor and complete detection at different frame rates. The first processing thread is to obtain images from real indoor and outdoor environment using RGB-D camera.

3.2 Plane Partition

The main goal of scene segmentation in this work is to divide indoor or outdoor scenes into several planes, which are used to guide visually impaired users. Visually impaired people usually expect assistive devices to be very smart, telling them where the walls are or whether there are stairs in the scene. In this work, the Z direction is defined as being parallel to the horizontal plane and pointing forward. The Y direction is perpendicular to the horizontal plane and points upward. The three directions of X, Y, and Z are in a right-handed coordinate system. Points on the same plane have similar normal vectors. First, we have used the normal vector in the Z direction to distinguish the ground and ceiling points from other points. In this step, the Z-direction components of all normal vectors need to be set to positive. Similarly, when dealing with the segmentation of the point cloud in other directions, this step is also required to deal with the ambiguity of the direction. The adaptive threshold method is adopted to eliminate misjudgments caused by fixed and manual thresholds. When segmenting in the Z direction, the adaptive threshold is multiplied by a larger factor, because the ground and the horizontal plane are considered parallel. Next, the X and Y direction normal vectors are used to distinguish walls or slopes.

Multiplying the normal vectors in the (Z, X) or (Y, Z) directions can help eliminate the correlation between different directions, and can also distinguish the walls. Then, the remaining unmarked points are used in the detection field

Whether there are stairs in the scene. In this work, a quick way to determine whether there is a staircase is to look at the coordinates in the Y direction of the three-dimensional point cloud in the world coordinate system, that is, the histogram of the y coordinates. If there are stairs in the scene, there will be several peaks in the histogram, so the order of stairs can also be counted by counting the number of peaks to remind the user. Finally, the result of segmentation is transmitted to the user through sound or vibration belt.

3.3 Saliency Based Obstacle Detection

In this section, three cluster-based cues are introduced to measure the cluster-level saliency.

3.3.1 Contrast Cue

Contrast cue represents the visual feature uniqueness on the single or multiple images. Contrast is one of the most widely used cues for measuring saliency in single image saliency detection algorithms [17-19], since the contrast operator simulates the human visual receptive fields. This rule is also valid in the case of cluster-based method for the multiple images, while the difference is that contrast cue on the cluster-level better represents the global correspondence relationship than the pixel/patch level.

The contrast cue $w^c(k)$ of cluster C^k is defined using its feature contrast to all other clusters:

$$w^c(k) = \sum_{i=1, i\neq k}^{k} \left(\frac{n^i}{N} \mu^k - \mu^i\right) \qquad (1)$$

where a L_2 norm is used to compute the distance on the feature space, n^i represents the pixel number of clusters C^i, and N denotes the pixel number of all images.

3.3.2 Spatial Cue

In human visual system, the regions near the image center draw more attention than the other regions [20–22]. When the distance between the object and the image center increases, the attention gain is depreciating. We extend this concept to the cluster-based method, which measures a global spatial distribution of the cluster. The spatial cue $w^s(k)$ of cluster C^k is defined as:

$$w^s(k) = \frac{1}{n^k} \sum_{j=1}^{M} \sum_{i=1}^{N_j} \left[N\left(\| z_i^j - o^j \|^2 \middle| 0, \sigma^2 \right) . \delta\left[b\left(p_i^j \right) - C^k \right] \right] \qquad (2)$$

where $\delta(.)$ is the Kronecker delta function, o^j denotes the center of image I^j, and Gaussian kernel $N(.)$ computes the Euclidean distance between pixel z_i^j and the image center o^j, the variance σ^2 is the normalized radius of images.

3.3.3 Corresponding Cue

We employ the variances of clusters to roughly measure how widely is the cluster distributed among the multiple input images. Firstly, a M -bin histogram $\hat{q}^k = \{\hat{q}^k\}_{j=1}^{M}$ is adopted to describe the distribution of cluster C^k in M images:

$$\hat{q}_j^k = \frac{1}{n^k} \sum_{i=1}^{N_j} \delta[b(p_i^j) - C^k], \; j = 1....M \qquad (3)$$

where n^k is the pixel number of clusters C^k, which enforces the condition $\sum_{j=1}^{M} \hat{q}_j^k = 1$. Then, our corresponding cue $w^d(k)$ is defined as

$$w^d(k) = \frac{1}{\operatorname{var}(\hat{q}^k) + 1} \tag{4}$$

where $\operatorname{var}(\hat{q}^k)$ denotes the variance of histogram $\hat{q}^k$ of the cluster C^k . The cluster with the high corresponding cue represents that the pixels of this cluster evenly distribute in each image.

3.3.4 The Co-Saliency Maps

Each cue, if used independently, has its advantages and, of course, disadvantages. A common fusion is formulated as a linear summation [17], [23] or point-wise multiplication [24] of static salient features. For saliency detection, however, the precision is more important than recall [25]. In our work, our also prefer a precise, rather than a large, saliency map. Therefore, we employ the multiplication operation to integrate the saliency cues. Before combining saliency cues, we normalize each cue map to standard Gaussian using the distribution of scores across all clusters. Then the cluster-level co-saliency probability $p(k)$ of cluster k is defined as

$$p(C^k) = \prod w_i(k), \tag{5}$$

where $w_i(k)$ denotes saliency cues.

Now that the cluster-level co-saliency value is computed, which provides the discrete assignment. Then we smooth the co-saliency value for each pixel. The saliency likelihood of the pixel x belonging to the cluster C^k satisfies a Gaussian distribution N as:

$$p(x \mid C^k) = N\left(\left\| v_x, \mu^k \right\|_2 \mid 0, \sigma_k^2\right), \tag{6}$$

where v_x denotes the feature vector of pixel x, and the variance σ_k of Gaussian uses the variance of cluster C^k . Hence, the marginal saliency probability $p(x)$ is obtained by summing the joint saliency $p(C^k)p(x \mid C^k)$ over all clusters:

$$p(x) = \sum_{k=1}^{K} p(x, C^k) = \sum_{k=1}^{K} p(x \mid C^k) p(C^k) \tag{7}$$

The designs of the software prototype of an obstacle segmentation and detection system for the visually impairment locomotion is described. The segmentation and detection parts of the project are clearly explained which include the plane partition, ground flatness detection and saliency-based obstacle detection. All the steps are performed successfully.

4. RESULTS AND DISCUSSIONS

The proposed system is evaluated in both indoor and outdoor environments. The image resolution is 640 × 480 and the depth map capture rate is 30 frames per second. The sensing range is 0.8 to 4.0 m. A Kinect sensor uses structured light methods to give an accurate depth map of a scene. Both the video and depth sensor cameras in the Kinect sensor have a 640 × 480-pixel resolution and run at 30 FPS (frames per second). Our camera device is totally fixed on a helmet or chest and waist. All of the experiment images are random images taken from the environment. The experiments are divided into two different environments: simple and complicated. A simple environment does not include stairs and a complicated environment has stairs. Both environments are situated indoors and outdoors, with sufficient and insufficient light. When obstacles are in front of the user, the system vocally informs the user of the distance to the obstacle.

4.1 Experimental Analysis

We evaluate the static obstacle detection method with 300 images captured at two different times with visually impaired people. We named them dataset-1 and dataset-2. Each dataset contains 200 frames including color image, depth image and accelerometer data. With dataset 1, the ground plane in depth image has a large area; whereas the dataset-2 ground only takes a small area.

For object level, we define manually obstacles of the scene. Each obstacle is determined by a rectangle. A detection result is a true detection if the ratio between the intersection of the detected and the ground-truth rectangles and the union of these rectangles is larger than 0.5.

We employ three evaluation measures that are precision, recall and F-measure to evaluate the performance of obstacle object detection results. These measures are defined as follows:

$$\mathrm{Pr}\,ecision(p) = \frac{TP}{TP + FP} \tag{8}$$

$$\mathrm{Re}\,call(R) = \frac{TP}{TP + FN} \tag{9}$$

$$F = 2\frac{\mathrm{Pr}\,ecision * \mathrm{Re}\,call}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call} \tag{10}$$

4.2 Plane Segmentation Results

A large number of tests have been carried out in this work to test the proposed algorithm. First, this work considers a typical static in the state of the scene, there are no stairs in this scene, and the images are collected in the laboratory corridor. After that, we analyzed a scene with stairs and moving pedestrians.

Fig-2 shows a typical indoor scene, including the ground and two walls. The image in Fig-2 is the color image of indoor environment. In this experiment, in order to calculate the normal vector, the search radius of the nearby point search is 0.03m, taking into account the effective depth output range of Kinect (0.5-8m) and the general size of the wall corner. The ground and ceiling are marked as red, and the two walls are marked as blue and green respectively.

Fig -2: Color image of indoor environment

Fig-3 shows an indoor scene with stairs. In addition to dividing the ground and the wall, we also used the histogram of the y coordinate in the world coordinate system to detect stairs. In the histogram, the number of peaks and the interval between peaks can be used to predict the number and width of stairs. This method works well when going up or down stairs. The entire segmentation program runs at approximately 0.6 frames per second on a 2.6GHz processor. This speed is faster than the RANSAC algorithm for normal vector estimation, but there is still a lot of room for optimization.

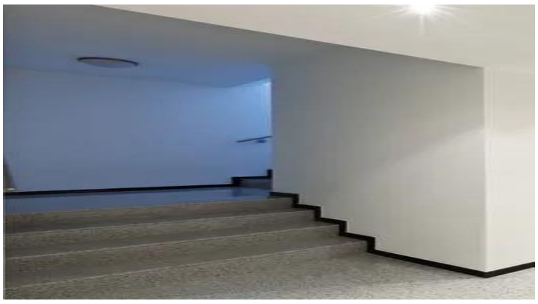
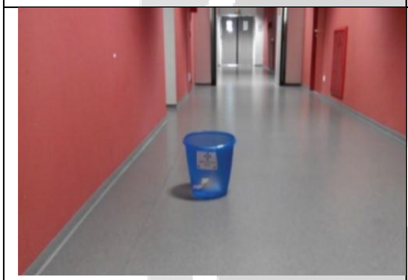
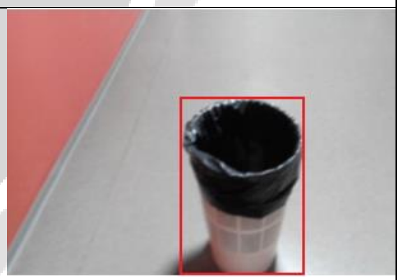
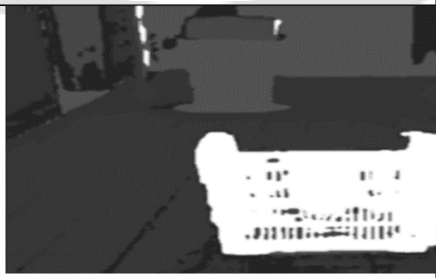


Fig -3: Color image of indoor scene with stairs

4.3 Experimental analysis of Obstacle Object Detection

This part describes the success rate for detecting obstacles in a simple environment without stairs. An obstacle is defined in this study as any object that interferes with a user's path. If an obstacle is labelled, then its identification is successful. If not, there has been a detection failure. The experimental results of the experiment are shown in Fig. 3. The first column displays the input picture, the second displays the saliency map detected by the detection technique, and the third displays the desired obstacle object. The saliency cue has the advantage of making rare clusters, such as baskets, appear more salient. However, the power of the saliency cue decreases in the presence of a complex background (e.g., the audience). Furthermore, it does not address the detection of common patterns across multiple images.

Input Image	Object Saliency Detection	Obstacle Detection
		
		
		

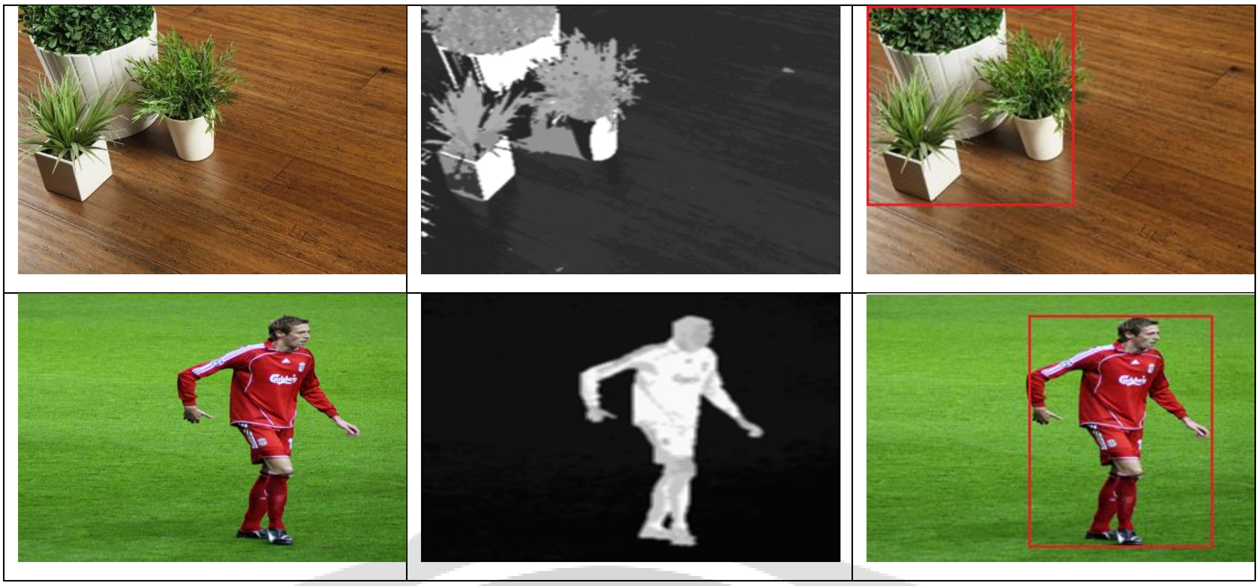


Fig -4: Obstacle detection with saliency map and the output marked by bounding box with red color.

Table-1: The success rate and the failure rate for obstacle detection of our proposed method

	Frame Amount (Total 300 frame)	Success Rate (%)	Failure Rate (%)
Overall	250	94.62	5.38%
Dataset-1	200	92.49%	7.51%
Dataset-2	100	96.75%	3.25%

Fig-4 illustrates some examples of detection while Table-1 and Table-2 shows the quantitative evaluation. Our algorithm provides better result in dataset-2, where the ground level takes small area. Overall, our method produces less false alarms with an acceptable rate of true detection. When the ground plane is wrongly detected or missed, it tends to consider the whole ground plane as an obstacle. That is why the overall precision score of dataset 1is significantly lower than dataset 1.

Table-2: Obstacle detection results of our proposed method

	Precision (P)	Recall (R)	F-Measure (F)
Overall	88.27	87.03	87.64
Dataset1	85.65	80.30	82.89
Dataset2	90.89	93.76	92.30

4. CONCLUSION

In this paper, we developed a system that combines RGB-D sensor cameras with machine learning to assist those with visual impairments in moving around obstacles avoiding colliding. Our method is intended to serve as a mobility aid while also detecting and warning of potential obstacles. The semantic segmentation method recognised the user's surroundings, while the co-saliency map detected the obstacles in front of the user. We designed a walking guide approach for visually impaired people by combining the two types of environmental information mentioned above. Considering the fact that those using it are visually impaired, the data presentation is straightforward, movable, hands and ears-free, and uses the human tongue as an interface. The system that was developed successfully detected both obstacles and humps in front of the user. It also identifies any objects that move in the foreground. The outcomes show that, under certain conditions, the method of imaging has been able to offer guidance cues, identify both static and moving obstacles, and calculate relatively accurate depth data in order to give warning signals at the appropriate time. The user receives immediate updates about their surroundings via a mobile application, and guardians are notified immediately if the user is in distress or falls down. Therefore, it ensures the user's safety. However, if an object appears suddenly in the camera's blind spot within 0.2-1 m of the user, they may not have enough time to avoid it. In this case, the white cane is still useful for the visually impaired individual. To ensure maximum safety when using the assistive system, we recommend that the visually impaired person continue to use the white cane. Hopefully, the system will assist visually impaired people with navigation, making their lives safer and easier. In addition, we will conduct additional experiments with visually impaired users to ensure the effectiveness of our proposed system. For the limited number of testers, we repeated various experiments several times, and the results are nearly consistent.

4. REFERENCES

[1] WHO. Available online on: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (accessed on 11 October 2018).

[2] Mekhalfi, M.L.; Melgani, F.; Zeggada, A.; De Natale, F.G.; Salem, M.A.M.; Khamis, A. Recovering the sight to blind people in indoor environments with smart technologies. Expert Syst. Appl. 2016, 46, 129–138.

[3] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks; proceedings of the Advances in neural information processing systems, F, 2012 [C].

[4] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks; proceedings of the Advances in neural information processing systems, F, 2015 [C].

[5] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation; proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), F, 2015 [C]. IEEE.

[6] He K, Gkioxari G, Dollár P, et al. Mask r-cnn; proceedings of the Computer Vision (ICCV), 2017 IEEE International Conference on, F, 2017 [C]. IEEE.

[7] Kirillov A, He K, Girshick R, et al. Panoptic Segmentation [J]. arXiv preprint, 2018, arXiv:.00868.

[8] Zhang Z. Microsoft kinect sensor and its effect [J]. IEEE multimedia, 2012, 19(2): 4-10.

[9] Keselman L, Iselin Woodfill J, Grunnet-Jepsen A, et al. Intel RealSense Stereoscopic Depth Cameras; proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, F, 2017 [C].

[10] Ortiz L E, Cabrera E V, Gonçalves L M. Depth Data Error Modeling of the ZED 3D Vision Sensor from Stereolabs [J]. ELCVIA: electronic letters on computer vision image analysis, 2018, 17(1): 0001-15.

[11] Bhatlawande, S.; Mahadevappa, M.; Mukherjee, J.; Biswas, M.; Das, D.; Gupta, S. Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation. IEEE Trans. Neural Syst. Rehabil. Eng. 2014, 22, 1148–1159.

[12] Islam, M.M.; Sadi, M.S.; Zamli, K.Z.; Ahmed, M.M. Developing walking assistants for visually impaired people: A review. IEEE Sens. J. 2019, 19, 2814–2828.

[13] Sivan, S.; Darsan, G. Computer vision based assistive technology for blind and visually impaired people. In Proceedings of the International Conference on Computing Communication and Networking Technologies, Dallas, TX, USA, 6–8 July 2016.

[14] Fei, Z.; Yang, E.; Hu, H.; Zhou, H. Review of machine vision-based electronic travel aids. In Proceedings of the 2017 23rd International Conference on Automation and Computing (ICAC), Huddersfield, UK, 7–8 September 2017; pp. 1–7.

[15] Jafri, R.; Ali, S.A.; Arabnia, H.R.; Fatima, S. Computer vision-based object recognition for the visually impaired in an indoors environment: A survey. Vis. Comput. 2014, 30, 1197–1222.

[16] Tian, Y. RGB-D sensor-based computer vision assistive technology for visually impaired persons. In Computer Vision and Machine Learning with RGB-D Sensors; Shao, L., Han, J., Kohli, P., Zhang, Z., Eds.; Springer: Cham, Switzerland, 2014; pp. 173–194.

[17] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[18] M. Cheng, G. Zhang, N. J. Mitra, X. Huang, and S. Hu, "Global contrast based salient region detection," in Proc. Comput. Vis. Pattern Recognit., 2011, pp. 409–416.

[19] R. Valenti, N. Sebe, and T. Gevers, "What are you looking AT?: Improving visual gaze estimation by saliency," Int. J. Comput. Vis., vol. 98, no. 3, pp. 324–334, 2012.

[20] L. Duan, C. Wu, J. Miao, L. Qing, and Y. Fu, "Visual saliency detection by spatially weighted dissimilarity," in Proc. Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 473–480.

[21] B. W. Tatler, "The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions," J. Vis., vol. 7, no. 14, pp. 1–17, 2007.

[22] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in Proc. Int. Conf. Comput. Vis., 2009, pp. 2106–2113.

[23] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum, "Learning to detect a salient object," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 2, pp. 353–367, Feb. 2011.

[24] C. Lang, T. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan, "Depth matters: Influence of depth cues on visual saliency," in Proc. 12th Eur. Conf. Comput. Vis., 2012, pp. 101–115. [36] L. Liu, R. Chen, L. Wolf, and D. CohenOr, "Optimis.

[25] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1597–1604.