Blind Navigation using Stereo Vision

Alka Leekha, Mukul Kumar, Shruti Bhardwaj#, Sarthak, Nitant Sood
Department of Information Technology,
Bharati Vidyapeeth's College of Engineering, New Delhi
#Corresponding Author, Email: *bhardwajshruti97@gmail.com*

Abstract— Blindness and visual impairment which are caused by contagious diseases has been extremely diminished, but rising number of people are in the danger of visual impairment which are age-related. Today 285 million people in the world are living with a vision impairment. A visual impairment refers to when one loses part or all of one's capability to see (vision). The impairment persists even with the use of eyeglasses, contact lenses, assistance, or surgery. Navigational tasks are based on visual information, individuals suffering with visual impairment face problems because appurtenant data of the surrounding environment of impaired person is not obtainable. Giving assistance to the visually impaired people on their navigational paths to their destination is an effortful task which has drawn the attention of various researchers. With the current progresses in the inclusive technology it has been made possible to augment the support that can be given to the people suffering with visual impairment at the time of their mobility. This literature review categories the current findings into distinct concept categories like stereo vision in blind navigation to enable advanced analysis into the area.

1. INTRODUCTION

According to WHO Fact Sheets, it has been estimated that approximately 1.4 billion people are living with some kind of distance or near visual impairment. According to distance vision, people with mild vision impairment 188.5 million in number, people with moderate to severe vision impairment are 217 million, and about 36 million people are blind [1]. Visual system is indeed an imperative and inescapable necessity to all the humans for exploring the ambient information for performing their day to day activities and therefore lead a comfortable life. Visual system is a wonderful characteristic gifted to organisms that allows us to appreciate the beautiful environment around us. Bereft of this feature, people will have to face a lot of complications in their daily life. So, there is a need to understand the problem of absence of human sight. The struggle of somehow, restoring the vision of a visually impaired or blind is the topic of an exhaustive research for both engineering and medical field. One major problem faced by impaired people is moving around in the environment without crashing into obstacles and reaching the destination successfully. From several years, long canes or guide dogs are used by visually impaired or blind people for their navigation. However, these guide dogs and long canes only provide the information of the obstacles which are at a short distance but failing in retrieving any information about the environment around. Obstacle detection is a process of detecting some barrier or objects near the autonomous system which can make the system change its movements. This method is being used to prevent any collisions in paths of self-driving vehicles or robots and hence ensuring safety of the vehicle, robot etc. Visually impaired and blind people face difficulties in navigation.

Obstacle detection plays an important role in many navigational systems. Generally, visually impaired have used the stick cane to navigate and to detect nearby objects, though this method does not provide much accuracy for navigational purposes at a long distance. Obstacle detecting system can help blind people in path navigations and in avoiding collisions. In the literature review we present the current findings in stereo vision in blind navigation to enable advanced analysis into the area. The paper has been organized as follows. Section I explains the algorithm and the techniques relevant to development of blind navigation systems. Section II describe recently developed navigation systems which are capable of assisting the blind or the visually impaired people.

2. RELEVANT TECHNIQUES

2.1 Computer Vision

Computer vision is a branch of computer science which helps computers to see, classify and process images in the similar way that human vision does, and then contribute to useful output. It is like conveying human judgement and instincts to a computer machine. Computer vision assignments include approach for collecting, processing, evaluating and also understanding the digital images, and deriving high-dimensional data from our real world for providing numerical or symbolic information, like in the form of decisions.

2.2 Stereo vision

Computer stereo vision refers to the extraction of three-dimensional information using digital images, like those obtained by a Charged Coupled Device camera. Any Information about the scene is compared using two vantage points. By examining relative positions of the objects in two panels, 3D information can be derived. This process is quite like our own biological process stereopsis. Stereoscopic images are usually stored as the MPO (multi picture object) files. Recently, for maintaining the high quality of stereo images, researchers have pushed to develop methods which can reduce the storage needed for MPO files. By using stereo cameras, objects that lie in the view of camera's field, appear at a little different location in the two images because of the different perspectives of cameras on the scene.

2.3 Estimation of depth map using single image

Single image is used along with, canny edge detection along with morphological operations is used to find the obstacles. Local depth hypothesis is used to find depth information. Canny edge detection is one of the optimal edge detectors. The noise is first removed by blurring the image with the help of Gaussian filter. Edge detection is performed on the obtained blurred image by studying pixel's gradient magnitude. The sharp edges are then detected by removing everything around the local maxima. Edge linking is then performed where broken edges are joined to retrieve a meaningful image. Edge linking is done by using the technique of morphological operation also known as dilation. The obstacles are then filled with white i.e. all the pixels in closed boundaries are colored white sing flood fill operation.

The point farthest from the user is called Vanishing Point. The brightness of the pixels increases from the vanishing point pixels to the pixels near the user. There are four depth hypotheses: top to bottom, left to right, bottom to top and right to left. According to the vanishing points one of the hypothesis is selected. When there are obstacles with no vanishing points then the default hypothesis is selected which is bottom to top. Final Depth for obstacle is estimated by comparing the depth map obtained on the bases of hypothesis using vanish points to the depth map for a given flat surface without any obstacles as default. This information obtained about the spatial data of the detected objects is then given to the visually impaired person.

2.4 Depth Map generation

In 3D computer graphics, depth map refers about the distance of surfaces of the objects from a given viewpoint. A pair of stereo images are rectified and then used to compute it. The intuition is proved by the image given below and mathematical formulas.

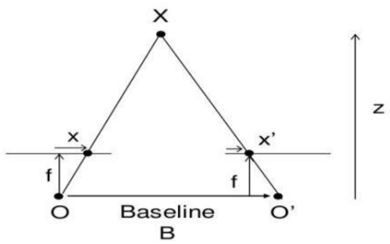

Fig.1: Depth map generated using stereo images.

This diagram contains equivalent triangles. Using the equivalent equations leads to the following result:

$$\text{Disparity} = x - x' = (B * f) / Z \quad (1)$$

x - x' denotes the distance between two points in an image plane which is correspondent to their scene point 3D and center of the camera. f is focal length of the camera which is already known and the distance between two O and O' cameras is B. So according to the equation the depth of any point in the image is inversely proportional with difference in their camera centers and distance of corresponding image points. So, using the information, the depth of all the pixels in an image can be derived. A simple algorithm is used to find matching patches in images. It is a simple operation called the sum of absolute differences. In order to have one value from 0 to 255 for each pixel, the two images are converted to grayscale before computation of depth map.

For computing the sum of absolute differences between a block and the template, value of each template's pixel is subtracted from the value of the corresponding block's pixel and then we take the differences' absolute values. Thereafter sum of all of the differences is taken which gives out a single value that approximately measures the similarity in two patches of the image. A higher value denotes that the patches are less similar. After finding the matches, disparity is found which is then used for computing distance of object from camera. By adjusting the values of number of Disparities and Size of block while computing disparity, we can obtain better results.

2.5 Acoustic Feedback

According to [2], the acoustic feedback helps in notifying the person of the existence of different kinds of obstacles, either dynamic or static. The main aim is to send warning messages which are fast enough, so as to avoid dangerous situations and assist the person to walk normally. Bluetooth bone conduction headphones are used at the hardware level. In this way, the user can continue with the normal process of capturing the external sounds generated by the surrounding environment as the user's ears are not covered.

But acoustic alerts are easier and preferrable to use as compared to alerts through tactile stimulation. Overall dynamics of the surroundings is insufficiently captured by vibrations of tactile stimulation. As a physical connection to the person's skin is required by tactile stimulation, the systems adopting this technique are considered invasive.
The detection of obstacle distribution in the scene, followed by the identified objects' semantic interpretation as obstacles and carrying out the effective transmission of the generated information to the user are the primary key elements. However, the manner in which transmitting of the acoustic feedback is done, plays a key role in approval of the developed device by the visually impaired users.

3. Related Work

There are many approaches for obstacle detection. One of the approaches is using smartphone where the set of points of interest are identified in the image and are tracked by an algorithm of Lucas Kanade. Detected obstacles are categorised as normal based or urgent upon their distance of the detected subject and its associated vector orientation. The obstacle that has been detected is sent to an object classifier which determines the degree of danger. Another technique is using the stereo-based color homography and the color segmentation technique. Another technique to detect obstacle uses a camera along with ultrasonic sensors. The obstacle is detected with the use of ultrasonic sensor and from the camera image, approximate size of the object is calculated.

3.1 Blavigator

According to [8]**,** the Blavigator model is composed of three main components which are the orientation and position unit which has the responsibility for providing the navigational system using the user's location which are in coordinates form, the Geographic Information System (GIS) contains geo-referenced data, stored in a database and the user interface. It helps the visually impaired people in moving around because it acts as an alternative for sensing of natural vision and interacts via audio interfaces.

3.2 Voice

In [3] , Voice uses a single video camera captures the image which is scanned and converted into sound.The loudness is proportional to the brightness of the pixel in the image.High frequency is used for top portion and low frequency is used for lower portion.

3.3 Navi

In [4], NAVI uses image processing techniques, differentiation between the background and objects is made,using brighter pixels for closer objects.Image is then used for conversion into stereo sound with loudness proportional to the brightness of the pixel in the image.

3.4 Optophone

In [5], Optophone uses edge detection technique a depth map is created,which is converted into sound with the technique similar to one used in Voice system.

3.5 Yoshihiro Kawai and Fumiaki Tomit's prototype system

In [6] the prototype system, there is a computer, a headset with an attached microphone, a headphone, three small cameras, and also a sound processor. The three images are used to prepare a 3D layout and for object detection and recognition which is then converted into 3D virtual sound.

3.6 An idea introduced by Zelek et. al.

In [7] the model images captured by the two video cameras are used to get disparity information, which is used to provide the information of object through tactile feedback.

3.7 Ifukube et al Navigation System

In [9], Ifukube et al developed a navigation system using two ultrasonic sensors which interacts with the user using sound. The idea behind this model was the reflection of ultrasonic waves by the nearby objects.

3.8 Smartvision

In SmartVision [11], Geographic Information System data is used to make geographic decision, indoor or outdoor and stereo vision system for safe and accurate navigation and orientation.

It consists of wifi for indoor and Global Positioning Unit (GPS) for outdoor navigation and apart from this computer vision along with Radio Frequency Identification (RFID) is used for both indoor as well as outdoor landmark or obstacle detection.

The GIS and Location are used to provide the orientation and navigation instructions to the user about the nearby obstacle and the appropriate direction of movement. The disparity in the two captured images are used to generate a depth map which in turn is used to get the information about the distance between the landmark or the obstacle to the user. The disparity in the image has been kept inversely proportional to the distance. This means that the object closer to the user is represented with brighter pixels as compared to the objects far away from the user.

Feature detection is also performed and in order to improve the performance only a portion of interest in the image is selected. The portion selected is usually the one near to the user. The example here taken is of a circle for which the image has been transformed into binary image with the circle considered as landmark to be white and the background to be black.

The path is decided by nearest circle as the immediate point of destination and the position of user along with position of circle used to find the correction angle which will be used by the user to correct his trajectory.

The instruction provided to the user to change its trajectory is not in absolute values, but the range of angles is divided into zones. The first zone being of minimal correction and the third zone of maximum correction.

3.9 Navbelt

In [10] , Navbelt developed by Shoval et al helps the user to navigate using a beep sound and employs ultrasonic sensor for collecting data. The model was found to be bulky and difficult to carry.There are many rangefinder techniques like ultrasonic , Laser etc. but are usually used to find the distance at which obstacle is and are unable to detect the change in the surface.

Huiying Shen and James Coughlan proposed a method which can detect curbs based on stereo vision but was not successful as the alignment of the stereo camera was not stable. Stereo Camera and two calibrated laser sensors were used by Aniket Murarka et al. [12] for obstacle detection based on motion and color segmentation. It is expensive and complicated and still suffers from alignment problem.

4. Conclusion

Assisting the impaired for navigation has been a difficult task which has drawn the interest of various researchers. Also, a ton of methods depending on GPS, RFID as well as various computer vision techniques have been developed for assisting blind navigation. A depth estimation method without any user interference from images was used for design and development of many navigation systems and the application which can help the visually impaired was developed. The main purpose of proposed works was to design a system that finds out particular landmarks in the person's environment, give navigation and orientation instructions and also provide information of distance to the visually impaired or a blind user. The proposed models have been able to provide correct and simple instructions to the visually impaired or blind user, helping him navigate in a non-intrusive manner.

Till now the disparity information from the depth maps has not been employed for distance calculation along with the correct indications and this side is still unexplored. Since every pixel in the depth map of the captured image provides the user's distance to the surrounding objects in the corresponding pixels in the image captured, along with the information about the distance to the correction output is the next step.

References

[1] R. Gnana Praveen and Roy Paily,"Blind Navigation Assistance for Visually Impaired Based on Local Depth Hypothesis from a Single Image", International Conference on Design and manufacturing, IConDM 2013, Procedia Engineering, Elsevier, Volume 64, pages 351 – 360, 2013

[2] Bogdan Mocanu, Ruxandra Tapu, Titus Zaharia, "When Ultrasonic Sensors and Computer Vision Join Forces for Efficient Obstacle Detection and Recognition" , Sensing Technology for Healthcare System ,2016.

[3] P. Meijer, "An Experimental System for Auditory Image Representations", IEEE Transactions on Biomedical Engineering, 1991, pp.112-121.

[4] G. Sainarayanan, "On Intelligent Image Processing Methodologies Applied to Navigation Assistance for Visually Impaired", Ph.D. Thesis, University Malaysia Sabah, 2002.

[5] M. Capp and P. Picton, "The Optophone: an electronic blind aid", Engineering Science and Education Journal, 2000, pp. 137-143.

[6] Y. Kawai and F. Tomita, “A support system for visually impaired persons to understand three-dimensional visual information using acoustic interface”, IEEE 16th conference on Pattern Recognition, 2002, pp. 974 – 977.

[7] J. Zelek, S. Bromley, D. Aamar and D. Thompson, “A haptic glove as a tactile vision sensory substitution for way finding”, Journal of Visual Impairment and Blindness, 2003, pp. 621–632.

[8] Paulo Costaa, Hugo Fernandesb, Paulo Martinsc, João Barrosod, Leontios J.Hadjileontiadise. “Obstacle detection using stereo imaging to assist the navigation of visually impaired people” , 4th International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2012),82-93.

[9] Ifukube, T., Sasaki, T., Peng, C., “ A blind mobility aid modelled after echolocation of bats” , IEEE Transactions on Biomedical Engineering ,1991, pp. 461-465.

[10] Shoval, S., Borenstein, J., Koren, Y., 1998. The Navbelt - a computerized travel aid for the blind based on mobile robotics technology,IEEE Transactions on Biomedical Engineering ,1998, pp.1376-86.

[11] Fernandes H., Costa, P., Filipe, V., Hadjileontiadis,L., Barroso, J., “Stereo Vision in blind navigation assistance”, Proceedings of the World Automation Congress 2010 – WAC2010, Kobe, Japan, 2010

[12] Aniket Murarka, Mohan Sridharan and Benjamin Kuipers,“Detecting Obstacles and Drop-offs using Stereo and Motion Cues for Safe Local Motion”,The IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008), Nice, France, 2008.