Navigation for Visually Impaired with Object Recognition and Voice Assist

S.Murugan@Prakasam [1], P.Hemalatha [2], P.Vinodhini [3]
[1] AssistantProfessor/CSE, Sri Muthukumaran Institute of Technology/Mangadu/Chennai/Tamil Nadu/India.
[2, 3] Student/CSE, Sri Muthukumaran Institute of Technology/Mangadu/Chennai/Tamil Nadu/India.

Abstract – The visually impaired who have to rely on others for travelling and other activities find difficult in mobility in unknown environment. The paper presents the architecture as well as the implementation of a system that helps the blind person navigate independently. It has additional components to provide refined location about environmental information. The system assist visually impaired people in different ways like camera, voice and ultrasonic based assistance.It provides information not only about obstacles along the travelled path, but also guide the user in selecting the preferred travel path.The idea of the proposed system to make person aware to walk and to detect the obstacle in the path and object recognition.In addition, the system is provided with a convenient and easy navigation aid for unsighted person, which helps in artificial vision to provide the information about the environmental scenario of static and dynamic objects around them.

Index Terms – Path Guiding, Obstacle Detection, Object Recognition, Voice Alert.

1. INTRODUCTION

According to the World Health Organization statistics, approximately 40 million people are blind all over the world. The Blind people faces great problem to recognize the barriers nearby them and to move from one place to another. 39 million are blind and 246 million have less vision. Around 90% of the visually impaired live in low income conditions. 82% of people living with blindness are around 50 and above. In 2015, blind pedestrians were killed in road accident-an increase of 5% from 2014. The biggest problem for blind persons is to detect the unknown environment and sense the movement around them**.** People who suffer from vision loss usually use dogs or walking stick to help them to detect obstacles. The stick cannot scan the platform as such and it just serve as a device to help people from danger in case of situation like traffic and in road crossing. To develop a Blind navigation system with multiple sensing inputs. To guide the Blind persons with dynamic object recognition. A hybrid sensory based blind Navigation system will be provided. To detect the proximity of the objects and recognition objects can be done. Dedicated proximity sensor along with camera interface provides street shop assistance as voice alert will enhance the experience of the blind persons. From the survey of World Health Organization, in the year of among 7 billion human populations 285 million people are blind and which 19 million are children who are below 15 years. Ninety percentages of school children did not pursue their secondary education because of this suffering from vision loss. There are totally 39 million blind people in this technical world. But there are fewer devices available to help them with their daily struggle. Predominantly IOT (Internet of Things) products serve as a major backbone for them. There are IOT products that help them to find objects ahead them. Since early decades, together with the sensors development, many efforts have been made to develop new and sophisticated Electronic Travel Aids (ETAs) able to perceive and represent the surrounding environment. The idea underlying the development of such devices was to overcome human sense limitations, such as blindness. These devices would help blind people to perceive their surrounding environment. It also helps them to navigate around without hurdle, without clashing with world. To develop a user-friendly low cost gadget for safe movement of visually impaired people.

2. RELATED WORK

People who suffer from vision loss usually use dogs or walking stick to help them to detect obstacles. The stick cannot scan the platform as such and it just serve as a device to help people from danger in case of situation like traffic and in road crossing. Secured feeling and confidence could be enormously increased using such devices that give a signal and warning to find the direction of an object less or obstacle less way in an changing environments .Electronic Travel Aids (ETAs) is a device that warns the user with help of some signals either the sound waves or by physical interaction with people such as vibratory patterns.

This system provides an important measure to reduce accidents among blind people in common traffic areas and give away warning to them by creating a great tendency to detect objects and obstacles as blind death has become common due to their inability to see and manage situations in heavy traffic. The existing system helps the visually impaired people to reach their destination by commanding them through voice recognition system via Bluetooth. The headset and walking stick are connected by Bluetooth .As soon as the data is received from the receiver in the Bluetooth headset, it is converted to text using voice recognizer.

3. PROPOSED MODELLING

A hybrid sensory based blind Navigation system will be provided. This system provides an important measure to reduce accidents among blind people in common traffic areas and give away warning to them by creating a great tendency to detect objects and obstacles as blind death has become common due to their inability to see and manage situations in heavy traffic.To detect the proximity of the objects, recognition objects and text to speech conversion can be done. Dedicated proximity sensor along with camera interface provides street shop assistance as voice alert will enhance the experience of the blind persons. Ultrasonic and IR sensors which help in obstacle detection and on hurdle recognition will ring the speaker for different durations to indicate different distances. Proximity sensor is to detect the presence of nearby objects without any physical contact. The system is to provide a convenient and easy navigation aid for unsighted which helps inartificial vision by providing information about the environmental scenario of static and dynamic objects around them. Also an approach to extract and recognize text from scene images effectively using computer vision technology and to convert recognized text into speech so that it can be incorporated with hardware to develop Electronic travel aid for visually impaired people.

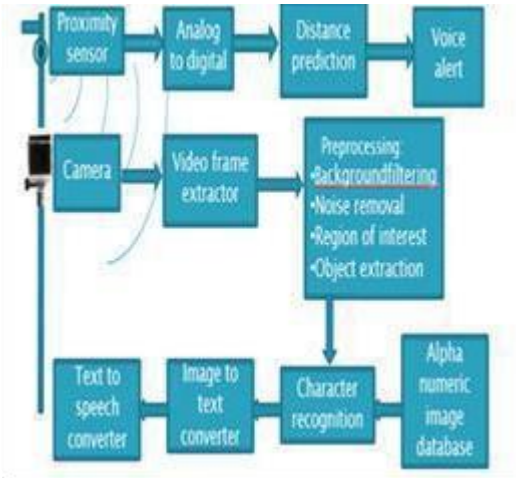

Figure 1: Architecture

4. METHODS

There are three main phases:

1. Proximity sensor for detecting the obstacle.

2. Object Recognition.

3. Text to Speech Converter.

4.1. Proximity Sensor

A proximity sensor is a sensor able to detect the presence of nearby objects without any physical contact. Proximity sensors can have a high reliability and long functional life because of the absence of mechanical parts and lack of physical contact between sensor and the sensed object. Proximity sensor has transmitter and receiver in which transmitter sends ultrasonic waves and the receiver will receive the reflected waves to detect the obstacle.

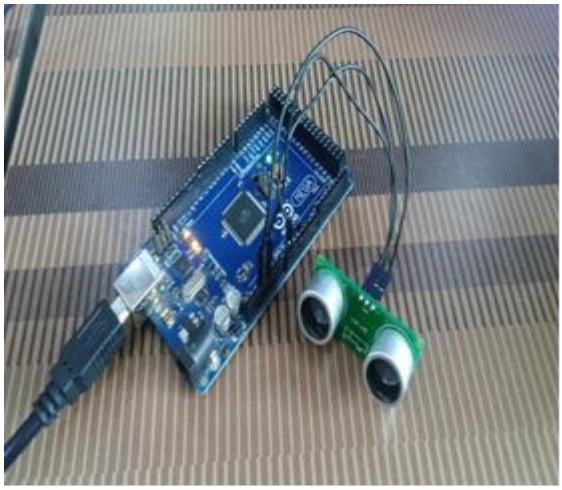

Figure 2: Proximity Sensor

4.2. Object Recognition

Figure 3: Process of Object Recognition

A shop banner image i captured and image Preprocessing (background removal, noise filtering) are done. The preprocessed image is then subjected to text

identification and extraction in image. Extracted text images will be compared to dictionary database to predict the text equivalent of its image.

Figure 4: Noise Removal

4.2.1. Acquisition Stage

It acquires high resolution video from camera. This video will be broken into different frames. Each frame will act as separate image. In this paper, standard images from ICDAR competition are used for testing purpose. In sophisticated system, video from camera is used as input to image processing unit. The challenges occurred due to camera fixation/ position problems like blurring/degradation due to motion of user, perspective distortion due to different angles of the object formed with camera lens axis. So, care must be taken while fixing camera of the system. Acquisition must be proper to produce desired output.

4.2.2. Image Processing

Second stage is Image Processing unit which is a heart of the system. It is subdivided into three secondary stages: Pre-Processing, Processing, Post-Processing.

4.2.2.1. Pre-Processing

As acquired colour image consist of 3 planes (Red, Green, and Blue); it is difficult to process it in quick time. So, it is first converted into Grayscale image. After colour to grayscale conversion, pre-processing stage uses some enhancement techniques to eliminate challenges created by noise, blurring effect and uneven lighting. It is considered that acquired image might be mixed with noise like Salt and pepper noise, Impulse noise etc. or it can be blurred due to motion of camera. Salt and pepper noise can corrupt the image, where the noisy pixels can take only the maximum and minimum values in the dynamic range i.e. black dot on white background (pepper) and white dot on black background (salt) which degrades the text extraction performance of system. To remove such type of noise, standard median filter (SMF), which is a non-linear filter used due to its good de-noising power and computational efficiency. Removal of this type of noise in a system is shown in Figure 4. However, when noise level is more than 50%, edge details will not be preserved by the median filter. So, it is recommended filtering process should preserve the edge details without losing the high frequency components of the image edges.

Sometimes, text extraction becomes troublesome for image captured in dark or uneven lighting. So, application of contrast enhancement is necessary. Histogram Equalization method is used for Contrast enhancement. Figure 5 shows enhanced image using histogram equalization. Another challenge i.e. Blurring can be generally removed using de-blurring techniques like Lucy Richardson algorithm, Blind de-convolution algorithm, Wiener de blurring techniques. Wiener filter is selected which is a natural extension of the inverse filter when noises are present. Figure 5 shows how de-blurring using Wiener filter is effective on text embedded blurred image. From figure, it is observed that binarization after wiener filtering on blurred image produces better result than without applying wiener filtering.

4.2.2.2. Processing

Enhanced Pre-processed image from previous block is forwarded to Processing stage where text detection and extraction is done. Before processing, it is binarized with adaptive thresholding. As a result of literature survey, we used combination of connected component (CC) & region based approach on this Binarized (black and white) image. Applying CC analysis using MATLAB software, areas having text similar patterns with white pixels on dark background, are detected. Using feature extraction algorithm, these detected areas are extracted on separate windows. Figure 6 shows text detection and extraction from Scene as well as Document image. Scene image requires preprocessing whereas Binarized Document image is directly fed to Text detection block. From Figure 4, it is observed that false detections are present due to some amount of noise in scene image including the Real text of image. This false detection reduces accuracy of the system. It also increases processing time and hence, decreases speed of the algorithm.

Figure 5: De-Blurring

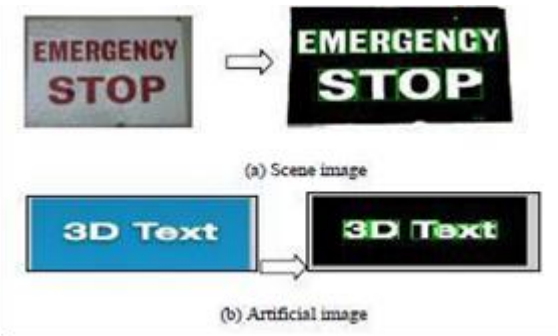

Figure 6: Text

4.2.2.3 Post-Processing

Next step is to recognize text from detected & extracted patterns in previous stage. From previous study, it is known that Recognition requires more sophisticated algorithm. Many approaches have been developed to recognize text from an image Such as linear discriminate analysis (LDA) for single character recognition, Support vector machine (SVM) & Conditional random field (CRF) for distorted multiple character recognition, Stroke Width Transform (SWT) for recognizing text appearing in different angles to overcome perspective distortion problem, etc. Feature learning is one of statistical pattern recognition algorithms are widely applied to computer vision problems and many results obtained with feature learning systems show high performance in recognition tasks. Also this algorithm achieves better results compared to other recognition methods. Hence, Feature Learning method is used for Text Recognition. Feature vector is the simplest form of knowledge representation. For character recognition, Feature vector are number of strokes, number of loops, width to height ratio of character. Figure 5 shows architecture of the recognition system.

4.3. Text to Speech Converter

This block converts the recognize texts into speech. Output is voice which enables blind person to listen to scene instead of read & act accordingly. As there is lots of research has been done on T2S system, many algorithm and softwares are available for it. LabVIEW software is used to design a VI which converts.

5. CONCLUSION AND FUTURE ENHANCEMENT

In this paper, design and implementation of prototype of text detection and recognition system has been discussed. After combining different techniques for Text detection and extraction, it is found that system works faster and better than using single technique for overall system. Text detection followed by recognition using supervised pattern recognition algorithm not only improves accuracy but also increases speed of the system. After successful recognition, text is converted into audio output.

Future Scope: Using this system, in future, by calculating the speed of the blind person in order to measure the distance covered by them to intimate the traffic signal timing and also any text detection and extraction in accurately.

REFERENCE

[1] Roberto Manduchi and James Coughlan, "ComputerVisionWithoutSight", *Communications of the ACM*, pp. 96 104, Vol.55, no.1, Jan.2012.

[2] Qixiang Ye and David Doermann, "Text Detection and Recognition in Imagery: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.1480-1500, Vol.37, No.7, July 2015.

[3] Ho Vu, Duong and Quoc Ngoc, Ly, "A Feature Learning Method for Scene Text Recognition", *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 176 - 180, Ho Chi Minh City, 2012.

[4] Nobuo Ezaki, Marius Bulacu, and Lambert Schomaker ,"Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", *17th Int. Conf. on Pattern Recognition (ICPR 2004), IEEE Computer Society*, pp. 683-686, vol. II, , Cambridge, UK, August 2004.

[5] Andrej Ikica, Peter Peer, "An improved edge profile based method for text detection in images of natural scenes", *IEEE EUROCON-International Conference on Computer as a Tool*, pp.1-4, Lisbon, 2011.

[6] M Sharmila Kumari, Akshatha, "Local Features based Text Detection Techniques in Document Images," *International Conference on Information and Communication Technologies (ICICT)*, pp. 6-11, 2014.

[7] Madhu S. Nair, K. Revathy, and Rao Tatavarti, "Removal of Salt-and Pepper Noise in Images: A New Decision-Based Algorithm", *Proceedings* of *the International Multi-Conference of Engineers and Computer Scientists IMECS*, Volume I, Hong Kong, March 2008.

[8] Sonia George, Noopa Jagdeesh, "A Survey on Text Detection and Recognition from Blurred Images", *International Journal of Advanced, Research Trends in Engineering and Technology(IJARTET)*, Vol. II, Special Issue X, pp. 1180-1184, March 2015.

[9] Yi-Feng Pan, Xinwen Hou, Cheng-Lin Liu, "Text Localization in Natural Scene Images based on Conditional Random Field", *10th International Conference on Document Analysis and Recognition*, pp.6-10, 2009.

[10] Linlin Li, Chew Lim Tan,"Character Recognition under Severe Perspective Distortion", *IEEE 19th International Conference on Pattern Recognition*, pp 1-4, Dec. 2008.

[11] R.C.Gonzalez, R.E.Woods, S.L.Eddins, „Digital Image Processing using MATLAB", *Pearson Education India*, 2009.

[12] S. Se, "Zebra-crossing detection for the partially sighted," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR)*, Hilton Head, SC, Jun. 2000, vol. 2, pp. 211–217.

[13] P. B. L. Meijer. (2003). Vision Technology for the Totally Blind. [Online]. Available: http://www.seeingwithsound.com

[14] "An experimental system for auditory image representations," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 2, pp. 112–121, Feb. 1992.

[15] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-based pedestrian detection," in *Proc. IEEE Intelligent Vehicles Symp.*, Dearborn, MI, Oct. 2000, pp. 215–220.

[16] L. Zhao and C. E. Thorpe, ―Stereo- and neural network-based pedestrian detection,‖ *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 3, pp. 148– 154, Sep. 2000.

[17] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. V. Seelen, ―Walking pedestrian recognition,‖ *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 3, pp. 155–163, Sep. 2000.

[18] U. Franke and S. Heinich, ―Fast obstacle detection for urban traffic situations,‖ *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 3, pp.173–181, Sep. 2002.