Navigating Visually Impaired and Sightless People using Auditory Guidelines

Dr.K.Priyadarshini

Professor, Department of ECE, K. Ramakrishnan College of Engineering, Samayapuram, Trichy, Tamilnadu

Abstract: The proposed work presents a novel strategy for navigating visually impaired and blind people using audio commands .To detect the significant objects in a video shot we start from the theory that moving objects attract attention more than other and are considered prominent. This projected method is evaluated and analysed based on the Optical character recognition which moves towards to detect the streets and shops name boards. Based on the OCR result a Text to speech (TTS) converter algorithm is used to assist the blind people.

Keywords: Optical Character Recognition; Converter Algorithm

I. INTRODUCTION

Digital Image is an image or picture represented digitally i.e., in groups of combinations bits (0 or 1). Digital Image Processing is a growing technology which is help to enhance the quality of the image. The Image processing is performing to extract information from the image digitally by the use of computer algorithm. Digital image processing is a versatile method and also it is very cheaper. Filtering is also including in digital image processing which is helps to blur or sharpen the image. Digital image processing depends upon the computer vision. There are two types of operation which is Low and middle level operation. First, increases the quality of the image to improve the understandability of the users with the help of low level operation. Second, increases the feature extraction and image segmentation with the help of middle level operation. Digital image processing are used in many field such as gamma - ray imaging-ray imaging, image in the ultraviolet band, Imaging in the visible and infrared bands, Imaging in the microwave band, Imaging in the radio band. Smoothing frequency which helps to smoothen the image with some filters such as Ideal low pass filters, Butterworth Low pass filter, Gaussian low pass Filters following the same, we can have filters to sharpen the images such as, Ideal High pass filters, Butterworth high pass filters, Gaussian High pass filters, High-boost filtering and High-frequency Emphasis filtering. Digital images are made of picture elements called pixels. Typically, pixels are organized in an ordered rectangular array. The size of an image is determined by the dimensions of this pixel array. The image width is the number of columns, and the image height is the number of rows in the array. Thus the pixel array is a matrix of M columns x N rows. To refer to a specific pixel within the image matrix, we define its coordinate at x and y.

The coordinate system of image matrices defines x as increasing from left to right and y as increasing from top to bottom. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modeled in the form of multidimensional systems. Image processing basically includes the following three steps:
- Step 1: Importing the image via image acquisition tools.
- Step 2: Analyzing and manipulating the image.
- Step 3: Output which is based on the image analysis.

II. EXISTING SYSTEM

The development and application of technology for orientation and mobility has a long history covering the postwar period. Although some early endeavors envisaged systems that might replace the cane or dog guide, more recent efforts have focused on devices and systems designed to supplement and provide a support system for these basic mobility tools. Mobility aids like walking stick and guide dogs are still used by the blind even today. With the advances of technology, some different types of electronic travel aid have been developed to support the mobility of the blind. Most of the commonly used electronic travel aids use ultrasound. All such devices use the principle of reflection of the high frequency ultrasonic beam, and are available in different models. Sonic Pathfinder, Mowat- Sensor, & Guide-Cane are called clear path indicators or obstacle detectors since the blind can only know whether there is an obstacle in the path ahead. These devices are used to search for obstacles in front of the blind person, and they operate in a manner similar to a flashlight, which has very narrow directivity. Sonic Guide and Nav Belt, however, are called an environment

sensor since it has wide directivity enabling it to search for several obstacles at the same time. The motivation of this work is to develop a portable navigation aid for blind pedestrians.

The most widely used primary mobility aid today is the long cane. This has several limitations such as a range limited to the length of the cane, typically one pace ahead of the user, difficulties detecting overhanging obstacles, and difficulties storing in public places. Here , the suggested navigation system involves a microcontroller with speech output. It is a self-contained portable electronic unit. It can supply the blind person with assistance about walking routes by using spoken words to point out what decisions to make. On the other hand, and in order to overcome the imperfections of existing electronic travel aids, the proposed method of measuring distance travelled in this system, is to use the acceleration of a moving body which in this case is the blind person. An accelerometer, followed by two integrators is used to measure a distance travelled by the blind. This technique is considered in inertial navigation systems and suffers from drift problems caused by the double integration and offset of the accelerometer which are overcome by the footswitch. When this footswitch is closed, the acceleration and the velocity are known to be equal to zero and this can be used to apply a correction. In addition, to help blind or visually impaired travellers to navigate safely and quickly among obstacles and other hazards faced by blind pedestrians, an obstacle detection system using ultrasonic sensors and vibrators has been added to this aid.

The proposed obstacle detection system consists then in sensing the surrounding environment via sonar sensors and sending vibro-tactile feedback to the user of the position of the closest obstacles in range. [1] Dense Conditional Random Field (CRF) , a powerful graphical model to globally capture the contextual information. [2] Presented a novel spatiotemporal saliency detection method using Gradient flow field and energy optimisation . [3] Fully convolutional networks for semantic segmentation, Liang-Chieh Chen, George Papandreou, Senior Member, IEEE, Iasonas Kokkinos, Member, IEEE, Kevin Murphy, and Alan L. Fully convolutional " networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning.[4] Spatiotemporal saliency detection for video sequences based on random walk with restart, Hansang Kim, Youngbae Kim, Jae-Young Sim. A novel saliency detection algorithm for video sequences based on the random walk with restart (RWR) is proposed in this paper.[5] Exploit features extracted from deep convolutional neural networks trained on object recognition datasets to improve tracking accuracy and robustness. he fingerprint, finger-vein and finger knuckle patterns can be simultaneously acquired and employed for more reliable bio-metrics identification. Speech is also one of the biometric method used now a days [6,14]. But if the speech is with different emotions then it is tough to identify the attack. Since recognition of emotion [7,8-9] itself is a tough from the speech. In this context, the constrained imaging requirement associated with the acquisition of finger-vein images can add to the total cost and user inconvenience while integrating them with fingerprint based systems. However, the simultaneous acquisition of fingerprint and finger knuckle images can be achieved without any additional inconvenience to users, also at lower cost, with simple addition of an external imaging camera to the existing (slap) fingerprint devices which can simultaneously acquire finger dorsal images and synchronizes the acquisition with external software. Therefore it is important to investigate the uniqueness and stability in the pieces of information which can be recovered from the finger dorsal images. In [10,12,13], the high performance analysis of the half pel and quarter pel interpolation for a single frame is done . The multimedia communication is efficient and standard video coding compression. The first frame can improve the resolution process and second frame can use the location of quarter pixel. In [11], presents the block matching algorithm based on the optimization and differential evolution, Two optimization technique can be using the method, partial swarm optimization (PSO) and differential evolution(DE) .The proposed method the high PSNR and compared with PSO algorithm. In addition, there are varieties of recorded forensic images in which only the finger dorsal patterns are available to establish the identity of a suspect. Automated or forensic identification of knuckle.

Forensic identification is the task of determining whether or not observed evidence arose from a known source. It is useful to associate probabilities with identification/exclusion opinions, either for presentation in court or to evaluate the discriminative power of a given set of attributes. At present, in most forensic domains outside of DNA evidence, it is not possible to make such a statement since the necessary probability distributions cannot be computed with reasonable accuracy, although the probabilistic approach itself is well-understood. In principle, it involves determining a likelihood ratio (LR) [15,16] – the ratio of the joint probability of the evidence and source under the identification hypothesis (that the evidence came from the source) and under the exclusion hypothesis (that the evidence did not arise from the source). Evaluating the joint probability is computationally intractable when the number of variables is even moderately large. It is also statistically infeasible since the number of parameters to be determined from the data is exponential with the number of variables. We generalize this approach to more complex data such as vectors and graphs, which makes LR estimation computationally tractable [9,12]. In paper [12] A modified UMHexagonS algorithm is employed in which search size of diamond is increased and a search pattern is skipped off and the experimental results of the algorithm employs improved PSNR with just a minimum bit rate increase with the maintenance of the image quality .[13] A real-time uncompressed video diffusion system, where Unequal bit Allocation (UBA) approach are adopted end-to-end

III. PROPOSED SYSTEM

Loss of sight is a problem that plagues millions of people everywhere. Blind people face many types of hurdles in performing every day routine works. Even in their own homes they must exhibit efforts to navigate from one place to another and to locate objects. According to the World Health Organization (WHO), 253 million people live with visual impairment, 36 million of which are blind and 217 million people have moderate to severe vision impairment. The conventional methods adopted like cane helps in avoiding the obstacles in their way but they do not help them identify and locate the objects. Hence, assistance is required for the blind that helps him/her in locating objects in an indoor environment. In this context, the current work proposes a novel solution for assisting the blind. The proposed system intends to assist the blind by taking voice commands to detect objects using image processing and provide audio output to navigate and reach the required object. The system also intends to recognize the necessary sign boards like washroom.

The video is captured using a camera which is then divided into a sequence of frames. Object detection is done using Haar cascade classifiers and color based object detection technique. Haar Cascade Algorithm: The Open CV library in Python has functions specifically to detect objects. It delivers software packages that are used to train classifiers for their object detection system, called Haar Training. Haar-like features: Object Detection using Haar feature based cascade classifiers is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. The algorithm extracts images using a lot of positive and negative images. A Haar-like feature can be considered as a template of several white and black rectangles interconnected. The features used are different size and rectangular. The main Haar-like features are shown in Fig. 1. The features value for the given mask is calculated as the weighted sum of intensity of the pixel intensities covered by the whole mask. But all the features extracted will not be useful for the purpose at hand. For example, in this work, objects like washroom, switch board, banana are relevant. At first, images with object of interest (positive image) and images void of object of concern (negative image) are captured. Images are resized and converted to gray scale. Text files of positive and negative images are generated which contain the coordinates of the positive and negative training samples along with their names in a specified order. Text files so created
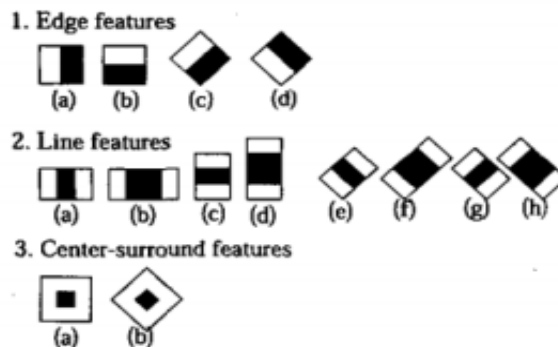

Figure 1.Feature prototypes of simple Haar-like centre surround

are used to obtain a vector file (.vec). This file contains compact information of positive instances of objects and negative image such as background. The training of classifiers is done using the vector file. After the training is completed, the output is stored in a directory in the form of text files. Finally, these files are converted into a single xml file which is the final classifier. The same classifier is used in python code to detect as well as track the object. Commands like 'left', 'right' and 'center' given via headsets aid the blind man to reach the object. Color based object detection: Some colored objects like a green bottle; a red ball is detected and tracked using color based object detection technique. At first, the upper and lower boundaries of color of the object are defined in HSV color. Video is captured and the same in converted to frames. These frames are converted into HSV color format. Then a filter is created which will mask the green color. The image obtained will be a binary image and it is eroded and dilated to remove noise. Then the largest contour in this binary image is found and then the coordinates and radius of a circle enclosing this contour is obtained.

IV. PROCEDURAL VIEW

It uses image processing as its primary technique to identify objects and sign boards. Video is captured by the Pi camera, from which the frames are extracted. The frames are preprocessed for better results. Image processing algorithms for object detection are applied on these frames and the object is detected. After the object is located, audio messages from headsets are given to the user to notify him/her about the location of the object. Obstacle detection is

done using ultrasonic sensors. Any obstacle encountered in the path is notified to the blind person by producing a beep sound from the buzzer. The proposed method can capture objects even in low lighting and bad weather conditions. Since it uses visual feed for navigation, more accuracy is obtained. They can easily target objects with the navigating vibration sensors.
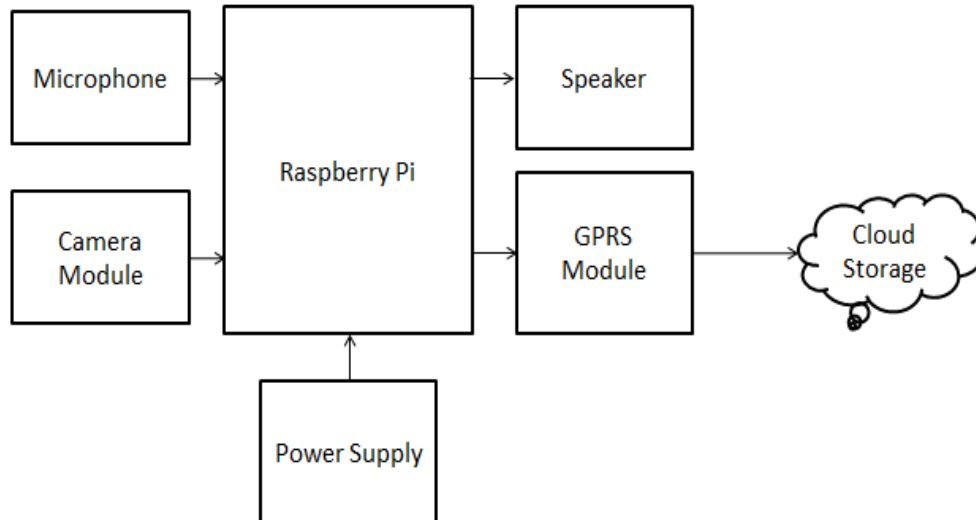
V.   BLOCK DIAGRAM

Figure.2 Overall proposed system

5.1 HARDWARE REQUIREMENTS

- Raspberry pi zero
- Speaker
- Microphone
- Camera module
- GPRS Module

5.2 SOFTWARE REQUIREMENTS

- Python 3.7.2
- OpenCV

The proposed system focuses on detection of objects. The system is made wearable and is portable. The system is mounted on the chest of the person. The Pi camera connected to Raspberry Pi captures the video of the scene and this is converted into frames by the processor. The user needs to provide speech commands when he/she is in need of the object. The speech input is recognized and the system guides the user to the object via audio output. Fig 2. shows an object being detected and guiding the user whether the object is to his right, left and center respectively.

VI.   RESULTS AND CONCLUSION

This proposed work presents a novel technique for assisting visually impaired people. The system designed has a simple architecture and makes it user friendly thus, making the subject independent in his/her home. The system also aims at helping blind to navigate in his/her surroundings by detecting obstacles, locate his basic necessities, read sign boards and texts. Preliminary experiments show promising results as the user can freely navigate in his surroundings safely. The system is made much more user friendly by accepting speech as the input to access his basic necessities.

REFERENCES

[1]. Ziad O. Abu-Faraj, Paul Ibrahim, Elie Jabbour and Anthony Ghaoui, "Design and Development of a Prototype Rehabilitative Shoes and Spectacles for the Blind", IEEE Int. Conf. Bio Medical Engineering and Informatics, 2012, 795-799.

[2]. Giva Andriana Mutiara, Gita Indah Hapsari and Ramanta Rijalul, "Smart Guide Extension for Blind Cane", IEEE Int. Conf. Information and Communication Technologies, 2016

[3]. G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and Sazali Yaacob, "A Stereo Image Processing System for Visually Impaired", Int. Journal of Computer, Electrical, Automation, Control and Information Engineering, vol.2, No.8, 2008, 2794-2803.

[4]. Rui Jiang, Qian Lin Li, "Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio", Proc. International Conference on Computer Vision, 2015.

[5]. M. Hiromoto, H. Sugano, and R. Miyamoto, "Partially Parallel Architecture for Ada Boost-Based Detection With Haar-Like Features" , IEEE Trans. Circuits and Systems for Video Technology, vol. 19, 2009, pp. 41-52.

[6]. Revathi AJeyalakshmi, C Thenmozhi, K "Person authentication using speech as a biometric against play back attacks", Multimedia tools and applications, 2018.

[7]. Revathi, A Jeyalakshmi, C Karthick M 'Speaker Independent Emotion recognition using Combined Feature and Spectral Analysis' International journal of control theory and applications, Vol.10(6), pp 483-492, 2017.

[8]. Jeyalakshmi, C Murugeswari, B Karthick, M 'Recognition of Emotions in Berlin speech: A HTK based approach for speaker and Text independent emotion recognition' Pakistan journal of Bio-technology, Vol.14(1), pp 63-69,2017.

[9]. Revathi A Jeyalakshmi, C 2018, "Emotions recognition: different sets of features and models", International Journal of Speech Technology, Springer, https://doi.org/10.1007/s10772-018-9533-6.

[10]. K.Priyadarshini, D.Jackuline Moni "A high performance VLSI architecture with half pel and quarter pel interpolation for a single frame," in Proc International science press, pp. 673- 680, 2016.

[11]. K.Priyadarshini, D.Jackuline Moni "Analysis of block matching algorithm based on particle swarm optimization and differential evolution". International Journal of applied engineering research.

[12]. K.Priyadarshini, D.Jackuline Moni "A Novel Approach To Improved Unsymmetrical –Cross Multi –Hexagon Grid Search Algorithm" International Journal of Applied Engineering Research Volume 9, Number 24 (2014) pp. 29141-29153.

[13]. K. Mohana Chandricka and. K. Priyadarshini "Uncompressed Video Streaming in Wireless Channel without Interpolation using Search Algorithms". International Journal of Advanced Scientific Research & Development (IJASRD), 06 (03/I), 2019, pp. 61 – 67. https://doi.org/10.26836/ijasrd/2019/v6/i3/60309 .

[14]. Revathi, A, Jeyalakshmi, C, Muruganantham, T "Perceptual Features based Rapid and Robust Language Identification System for Various Indian Classical Languages", In: Hemanth D., Smys S. (eds) Computational Vision and Bio Inspired Computing. Springer Lecture Notes in Computational Vision and Biomechanics book series (LNCVB), vol. 28, pp.291-305, 2018.

[15]. Revathi, A Jeyalakshmi, C "Comparative Analysis on the Use of Features and Models for validating Language Identification System", Proceedings of International Conference on Inventive Computing and Informatics (ICICI 2017), pp.693-698, Nov.2017.

[16]. Karthick, M Muruganantham, T Jeyalakshmi, C Revathi, A 'A high performance HTK based language identification system for various Indian classical languages' Pakistan journal of Bio-technology, Vol.14(1), pp 77-81, 2017.

Biography

Dr. K. Priyadarshini the author, is a Professor, of Department of ECE in K. Ramakrishnan college of Engineering, Samayapuram, Tiruchirapalli, Tamilnadu. She completed his B.E at V.L.B Janakiammal college of Engineering & Technology, under Bharathiar University , M.E (Communication Systems) at Jayaram college of Engineering & Technology affiliated to Anna University, Chennai, M.B.A at IGNOU & pursued her Ph.D at Karunya Institute of Technology & Sciences, Coimbatore in the area of Video Processing. Her research interests are Image processing, Motion Estimation, Video surveillance.