Smart Assistance System for the Visually Impaired

Kote Shubham*, Pawale Ankit* , Rajguru Ajinkya* , Shelke Om*

*Department of Computer Engineering, Sinhgad Technical Educational Society's, Sinhgad Institute of Technology and Science

Abstract- Physical movement is one of the biggest challenges for the visually impaired. People with complete blindness or low vision often have a difficult time in self-navigating unfamiliar environments. Traveling or simply walking down a crowded street may pose great difficulty. So, many people with low vision tend to bring a sighted friend or family member for assistance. It also becomes difficult for them to keep a track of their routine environments .This system proposes conglomeration of technologies like Image processing,Speech processing etc, so the problems faced by blind people can be reduced to certain extent. Object recognition methods in computer vision , Image processing, Text to Speech conversion can be embedded in a single object : SMART GLASSES (spectacles) .

Index Terms-Blind navigation, Image processing, Voice processing, Object recognition, object detection, Face recognition.

I. INTRODUCTION

Of the 37 million people across the globe who are blind, over 15 million are from India.India is now home to the world's largest number of blind people. India needs 2.5 lakh donated eyes every year, the country's 109 eye banks (five in Delhi) manage to collect a maximum of just 25,000 eyes, 30% of which can't be used. Majority of the people cannot afford such treatments. To be categorised as blind, there is a total loss of vision. Blindness cannot be corrected by simple visual aids such as glasses. For the indigents blindness is a drawback. So this paper puts forward a system to aid the visually challenged. The Assistor is a device which is a passive type intelligent stick that focuses on aiding the visually challenged people to move around from one place to another without having to worry about anything.

As a result, blind people usually rely on other sensory information in order to avoid obstacles and to navigate . For example, the motion of dynamic obstacles generates noise allowing visually impaired people to determine the approximate position using their auditory senses. The additional use of tactile senses is required for precise obstacle localisation. For this purpose a white cane is commonly used by blind people , which has two main disadvantages. It is relatively short and the detection occurs only by making contact with the obstacle which could sometimes might be dangerous. Another popular navigation tool for visually impaired individuals is a guide dog. Compared to white canes, dog guides are able to detect obstacles as well as steering around them, however they are expensive and only have a very limited working life.

However, many obstacle detection and avoidance systems have been proposed during the last decade to help blind people navigate in known or unknown, indoor and outdoor environments. This navigation can primarily be categorised as vision replacement, vision enhancement and vision substitution. Vision replacement systems provide the visual cortex of the human brain with the necessary information either directly or via the optic nerve. Vision enhancement and vision substitution systems have similar working principles with regard to environment detection process, however, each provides the environmental information differently. Vision enhancement presents the information in a visual manner, whereas vision substitution typically uses tactual or auditory perception or a combination of the two.

We propose a Smart system to help such visually impaired people in their basic activities of life.
Our System mainly consists of two components:

1) Wireless Camera
2) Android Device

Our main objective is to simplify the system for the users and make interaction between the two entities as easy as possible.The entire project is dependent on the Smartphone App and its reliability. It performs all the computation and calculations. A separate database is designed, where the definition of the objects are found. In the system level we could say that the novelty lies in the real-time application working on the Smartphone.

II. LITERATURE SURVEY

There are a lot of devices which assist the visually challenged for navigation indoor and outdoor. All these devices rely mainly on Global Positioning System (GPS) alone, to navigate around.

1. A Stereo Image Processing System for Visually Impaired.

The above paper proposes a system utilises stereo vision, image processing methodology and a sonification procedure to support blind navigation. The developed system includes a wearable computer, stereo cameras as vision sensor and stereo earphones, all moulded in a helmet.

Limitations:
- Size of hardware required is extensive and voluminous.
- Musical stereo sound for the blind's understanding of the scene in front (No use of voice commands).

2. Blind Path Obstacle Detector using Smartphone Camera and Line Laser Emitter.

Two things are prominently used in this paper i.e Mobile camera and laser. The laser and the mobile is kept at static distance. The image is captured from the camera and along with it the laser is also observed.Using the static distance and the angle between the laser point and the camera the distance is measured.

Limitations:
- Distance between the camera and the laser must be constant.
- May not work efficiently on shiny surface as laser intensity may decrease.

3. Abandoned object detection via temporal consistency modelling and back-tracing verification for visual surveillance. This paper presents effective approach for detecting abandoned luggage in surveillance video. Here the camera is fixed at a position. Hence, the dataset of images captured is static. Therefore if any irregularity is observed in the surveillance the objects are detected.

Limitations:
- The surveillance is static
- Background should be known beforehand.

III. METHODS

A. Text-to-Speech

This module comprises of image and speech processing. The main aim of this module is to acquire a 3D world real image of any text constraints area and to convert this image into text followed by providing audio output using speech processing. We are implementing a dynamic system that makes use of Google API's for conversion of Text to Speech dynamically provided that good internet connectivity is present. 1)If a book or any writeup is held in front of the camera frame capture the image of the text to be read by the user by using the camera.

For instance, if a book or any writeup is held in front of the camera frame capture the image of the text to be read by the user by using the camera.Send the data to the android device via bluetooth. Match each and every letter and provide the output to the user i.e the name of the

recognised face if available in the dataset through the ear piece.

B. Object Recognition

Object Recognition is a process in which Real world objects are identified using Image processing. Object Recognition is a important process which will aid visually impaired persons to locate their frequently used day to day objects. It becomes a tedious task for visually impaired person to locate these objects. Our system gives sort of visual aid that dynamically identifies objects and locates them.

Our algorithm analyses the position of the object, size of the object, shape of the object, etc. For Example, a visually impaired person is sitting on his study table. He has multiple objects in front of him such as Water Bottle, Walking Stick, Fruits ,etc. So our System will help him locate all his objects.

C. Face recognition

Some face recognition algorithms identify facial features by extracting landmarks, or features, from an image of the subject's face. For example, an algorithm may analyse the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw.

These features are then used to search for other images with matching features. Other algorithms normalise a gallery of face images and then compress the face data, only saving the data in the image that is useful for face recognition. A probe image is then compared with the face data. One of the earliest successful system is based on template matching technique applied to a set of salient facial features, providing a sort of compressed face representation.

Recognition algorithms can be divided into two main approaches, geometric, which looks at distinguishing features, or photometric, which is a statistical approach that distills an image into values and compares the values with templates to eliminate variances.

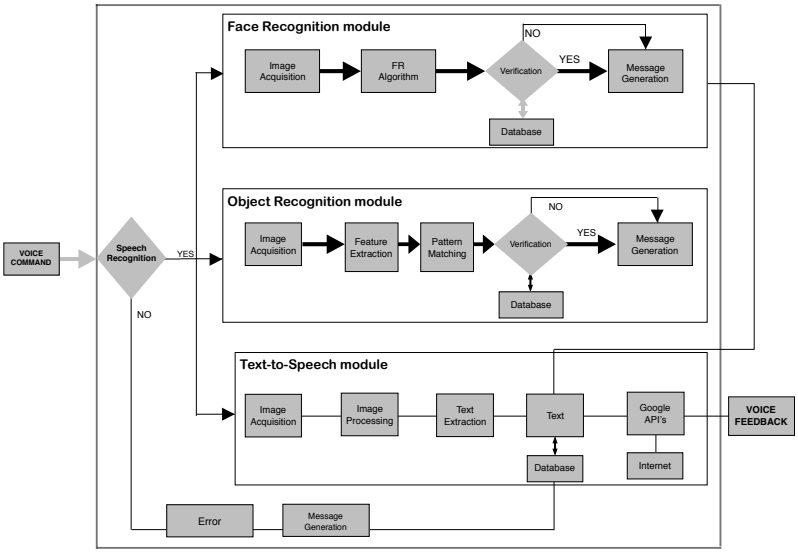


Fig 1:System Architecture

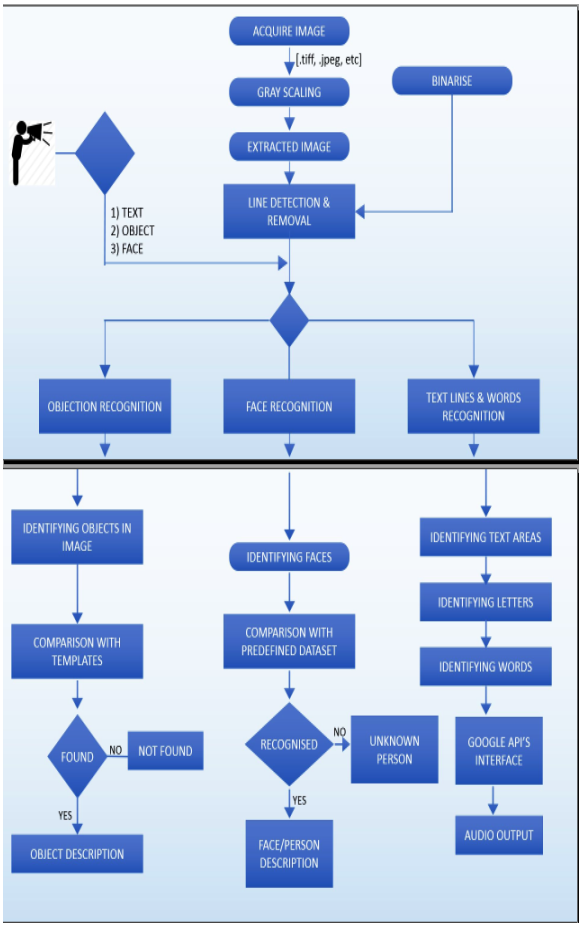
IV. ALGORITHMS

5.1.1 Face Recognition

1.Prepare a dataset of known/recognised faces.
2. Reading this dataset in our program.
3. Calculate the eigenfaces from the dataset (Training set), keeping only m images that correspond to highest eigenvalues. These m images define the face space.
4. Calculate corresponding distribution in m-dimensional weight space for each known individual, by projecting their face images in face space.
5. Loading any image format (.bmp, .jpg, .png) from given source. Detecting the faces present in the images using the CascadeClassifier in opencv.
6.We will use Face Recogniser for face recognition.
7. Calculate set of weights based on input image and the m eigenfaces by projecting the input image onto each eigenface.
8. Determine if image is known or unknown by checking to see if the image is sufficiently close to face space.
9. Text to Speech conversion using known description of person and output through earpiece.

5.1.2 Object recognition

1.Prepare a dataset of known/recognised objects.
2. Reading this dataset in our program.
3. Loading any image format (.bmp, .jpg, .png) from given source. Detecting the objects present in the images using Edge detection, line detection, and pattern detection algorithms.
4. Extracted objects are than compared with the know dataset.
5. Decision is based on this comparison if the object is recognised or not.
6. Once the object is recognised its respective description is extracted .
7. Text to Speech conversion using known description of object and output is provided through earpiece.



5.1.3 Text-To-Speech

1. Loading any image format (bmp, jpg, png) from given source. Then convert the image to grayscale and binarise it using the threshold value.
2. Detecting image features like resolution and inversion. So that we can finally convert it into a straightened image for further processing.
3. Lines detection and removing. This step is required to improve page layout analysis, to achieve better recognition quality for underlined text, to detect tables, etc.
4. Page layout analysis. In this step we are trying to identify the text zones present in the image.
5. So that only that portion is used for recognition and rest of the region is left out.

6. Detection of text lines and words. Here we also need to take care of different font sizes and small spaces between words.
7. Recognition of characters. This is the main algorithm ; an image of every character must be converted to appropriate character code.
8. Saving results to selected output format, for instance, searchable PDF, DOC, RTF, TXT.
9. Text to Speech conversion using google API's and output through earpiece.

V. CONCLUSION

In this paper, a system which allows visually impaired individuals to detect and avoid obstacles was implemented as an android application. The obstacle detector application provides a high detection rate of up to 100% on selected environments. The main limitations of this system were flooring materials which have extremely weak light reflection and obstacles with a colour similar to the laser light. In order to improve the robustness of this system a powerful line laser module could be used, allowing an improved laser line detection. A cross laser module, as opposed to a line laser module, should be investigated. The proposed system could be also applied to smart phones that have two back cameras, thus enabling us to additionally measure a depth map of the environment. This application could be enhanced with GPS information, a common feature to most smart phones. Additionally, if the processing performance is improved it will allow the processing of a higher frame resolution, and thus allow more precise acquisition of the local environment. This can be achieved by limiting the search of the template matching algorithm, since the laser light is moving in a predefined range.

VI. ACKNOWLEDGEMENT

Special thanks goes to our guide for their constant guidance and support in our project. Her valuable guidelines and suggestions helped us in completing this particular task.

VII.REFERENCES

Blind Path Obstacle Detector using Smartphone Camera and Line Laser Emitter
Rimon Saffoury, Peter Blank*, Julian Sessner**, Benjamin H. Groh*, Christine F. Martindale*, Eva Dorschky*, Joerg Franke** and Bjoern M. Eskofier**

Autonomous Walking Stick For The Blind Using Echolocation And Image Processing .
Akhilesh Krishnan, Deepakraj G, Nishanth N, Dr.K.M.Anandkumar

A Stereo Image Processing System for Visually Impaired .
G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and Sazali Yaacob

Abandoned Object Detection via Temporal Consistency Modeling and Back-Tracing Verification for Visual Surveillance.
Kevin Lin, Shen-Chi Chen, Chu-Song Chen, Daw-Tung Lin, Senior Member, IEEE, and Yi-Ping Hung.

Smart vision for the blind people .
R.Mohanapriya,U.Nirmala,C.Pearlin Priscilla