Abstract — In the current scenario, eye related problems are seen in high numbers both in India and also all over the world. Some people are born with reduced sight while others may acquire an eye condition later in life. Others may lose their sight in an accident, or because of a disease like diabetes or arthritis. This paper demonstrates the prototype of visual aid for blind and partially sighted. It uses Raspberry pi kit with ARM v7 processor for Digital Image Processing and the kit interface with a HD camera to capture high quality images. And we use headphones to guide the user through the instructions. The main advantage of this system is, more accuracy can be achieved in detecting small objects even in unknown places.

Keywords — Raspberry pi kit, ARM v7 processor, HD camera.

I. INTRODUCTION

The World Health Organization estimates that 285 million people are visually impaired worldwide (39 million are blind and 246 million have low vision) [1]. With 7.8 million blind people in India, the country accounts for 20 per cent of the 39 million blind population across the globe, of which 62 per cent are on account of cataract, 19.7 per cent refractive error, 5.8 per cent glaucoma and one per cent corneal blindness. Experts said that of the total blind population in the country, 4.7 per cent accounted for diabetic retinopathy and age related macular degeneration. It was also disclosed that a total of 285 million visually impaired were also present in the country, who had some form of impairment. At least 20 per cent of young people with vision impairment have additional disabilities and/or special educational needs and a further 30 per cent have very complex needs.

There are various existing methods to aid the blind to move through the obstacles but are inefficient at certain circumstances and have many drawbacks. White canes and guide dogs [2] are really helping for navigational assistance, but they have some practical restrictions as well. The white cane is only partially effective as it does not detect objects above the knee height, and does not provide cues in sufficient time to avoid a collision in a populated area. As for the guide dog, unfortunately not all visually impaired individuals have access to dogs due to an ongoing shortage of properly trained dogs. Navigation in unfamiliar spaces is a problem for the visually impaired [3], but their learning is relatively rapid. So applications that support navigation in unfamiliar places are very helpful [4]. Integration of current technology such as position recognition, obstacle detection, and embedded systems accommodates the design of feedback systems to help the visually impaired navigate more easily.

This paper presents an assistive device for blind and partially sighted which help the visually impaired to navigate through the unknown places and also to overcome challenges of dependence. The paper is organized as follows. Section II presents previous works on navigation assistive devices for the blind, Section III presents the design and implementation. Section IV presents Digital Image Processing Algorithm, and Section V presents conclusion and future works.

II. PREVIOUS WORKS

Over the past three decades, research has been conducted to design new navigation devices for the visually impaired. Shoval et al [5] proposed a navigation belt comprising of an array of ultrasonic sensor to detect obstacles, but it is not an ideal method for operation in dense and noisy environment. Bousbia-Salah [6] proposed a method of detecting obstacles on the ground through an ultrasonic sensor integrated on the white cane and the user’s shoulders. The hand-held sonar device does not provide navigation assistance, and cannot detect drop-offs such as curbs, steps etc.

A laser cane that uses optical triangulation with three laser diodes. The first laser points at the ground detecting a drop in elevation, the second points straight in front of the user parallel to the ground, and the third points straight ahead at an angle of 45 degree from the ground to protect the user from overhanging obstacles. But the laser cane provides no navigational assistance, triggers a tone to indicate a drop off even for a small change in ground level. Farrah [7] proposed the virtual reality technology to capture images of the house using cameras, and uses this information for indoors navigation. But this virtual reality technology requires extensive image processing capabilities, and does not work in new or busy outdoor environments.

III. DESIGN AND IMPLEMENTATION

i. Overview of components

We are using digital image processing technique to process the image captured and guide the user through audio instructions. The Fig. 1 shows the overview of components used in the implementation of navigation...
assistive device. It includes Raspberry pi kit, Camera, Power bank, Headphone and IR control.

![Fig.1 Overview of components](image1)

### Raspberry pi kit

We are using RASPBERRY PI 2 B kit comprises of ARM v7 processor, HDMI port, 2 USB ports, power socket and have the following technical specifications.

**TABLE 1 SPECIFICATIONS OF RASPBERRY PI 2 B**

<table>
<thead>
<tr>
<th>Chipset</th>
<th>Broadcom BCM2836 SoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Architecture</td>
<td>Quad-core ARM Cortex-A7</td>
</tr>
<tr>
<td>CPU</td>
<td>900 MHz</td>
</tr>
<tr>
<td>GPU</td>
<td>Dual Core Video Core IV® Multimedia Co-Processor Provides Open GL ES 2.0, hardware-accelerated Open VG, and 1080p30 H.264 high-profile decode Capable of 1Gpixel/s, 1.5Gtexel/s or 24GFLOPs with texture filtering and DMA infrastructure</td>
</tr>
<tr>
<td>Memory</td>
<td>1GB LPDDR2</td>
</tr>
<tr>
<td>Operating System</td>
<td>Boots from Micro SD card, running a version of the Linux operating system</td>
</tr>
<tr>
<td>Dimensions</td>
<td>85 x 56 x 17mm</td>
</tr>
<tr>
<td>Power</td>
<td>Micro USB socket 5V, 2A</td>
</tr>
</tbody>
</table>

![Fig.2 chipset](image2)

The Fig. 2 and Fig. 3 show the chipset and pin diagram of GPIO.

### Power bank

We use a power bank to give standalone power supply for the Raspberry pi kit. The voltage and current specification of the power bank should be 5V and 2A respectively.

### Headphone

Normal mobile headset can be used by inserted it into the audio jack of the kit to provide instructions to the user for navigation.

### IR control

IR sensor based remote control is used to orderly turn ON/OFF the Raspberry pi kit to avoid any loss of data that stored in the SD card.
ii. Working

The camera will be used to capture consecutive high quality images of the area in front of the user, which is then stored in the Memory of the kit. The images are then resized to a nominal value to reduce the time required to operate on the matrix. The first priority of the algorithm is to detect the people in front of the user. This is done by detecting the faces of the people in the image. Viola Jones algorithm is used to detect the faces. If the faces are detected, then the voice command is given to the user using the headphone. The position of the people can be detected by analyzing the position of the pixel distance from the reference point, which is middle and both the corners.

IV. IMAGE PROCESSING ALGORITHM

A. Face Detection

Viola–Jones algorithm-The basic principle of the algorithm is to scan a sub-window capable of detecting faces across a given input image. The standard image processing approach would be to rescale the input image to different sizes and then run the fixed size detector through these images. This approach turns out to be rather time consuming due to the calculation of the different size images. Contrary to the standard approach Viola–Jones rescale the detector instead of the input image and run the detector many times through the image – each time with a different size. At first one might suspect both approaches to be equally time consuming, but Viola–Jones have devised a scale invariant detector that requires the same number of calculations whatever the size. This detector is constructed using a so-called integral image and some simple rectangular features reminiscent of Haar wavelets.

Feature types and evaluation

The characteristics of Viola–Jones algorithm which make it a good detection algorithm are:

- Robust – very high detection rate (true-positive rate) & very low false-positive rate always.
- Real time – For practical applications at least 2 frames per second must be processed.
- Face detection only (not recognition) - The goal is to distinguish faces from non-faces (detection is the first step in the recognition process).
4. Cascading Classifiers

The features sought by the detection framework universally involve the sums of image pixels within rectangular areas. As such, they bear some resemblance to Haar basis functions, which have been used previously in the realm of image-based object detection [10]. However, since the features used by Viola and Jones all rely on more than one rectangular area, they are generally more complex. The figure on the right illustrates the four different types of features used in the framework. The value of any given feature is the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles. Rectangular features of this sort are primitive when compared to alternatives such as steerable filters. Although they are sensitive to vertical and horizontal features, their feedback is considerably coarser.

Advantages of Viola-Jones algorithm
- Extremely fast feature computation
- Efficient feature selection
- Scale and location invariant detector
- Instead of scaling the image itself (e.g. pyramid-filters), we scale the features.
- Such a generic detection scheme can be trained for detection of other types of objects (e.g. cars, hands)

Disadvantages of Viola-Jones algorithm
- Detector is most effective only on frontal images of faces
- It can hardly cope with 45° face rotation both around the vertical and horizontal axis.
- Sensitive to lighting conditions
- We might get multiple detections of the same face, due to overlapping sub-windows.

B. Object Recognition

Bag of Features – This method is one that represents images as order less collections of local features. The name comes from the Bag of Words representation used in textual information retrieval. With Bag of Words, one represents a document as a normalized histogram of word counts. Commonly, one counts all the words from a dictionary that appear in the document. This dictionary may exclude certain non-informative words such as articles (like “the”), and it may have a single term to represent a set of synonyms. The term vector that represents the document is a sparse vector where each element is a term in the dictionary and the value of that element is the number of times the term appears in the document divided by the total number of dictionary words in the document (and thus, it is also a normalized histogram over the terms). The term vector is the Bag of Words document representation – called a “bag” because all ordering of the words in the document have been lost. The Bag of Features image representation is analogous. A visual vocabulary is constructed to represent the dictionary by clustering features extracted from a set of training images. The image features represent local areas of the image, just as words are local features of a document. Clustering is required so that a discrete vocabulary can be generated from millions (or billions) of local features sampled from the training data. Each feature cluster is a visual word. Given a novel image, features are detected and assigned to their nearest matching terms (cluster centers) from the visual vocabulary. The term vector is then simply the normalized histogram of the quantized features detected in the image.

![Fig. 8 Process for Bag of Features Image Representation](image)

The procedure for generating a Bag of Features image representation is shown in Fig. 8 and summarized as follows:

1. **Build Vocabulary**
   - Extract features from all images in a training set.
   - Vector quantize, or cluster, these features into a “visual vocabulary,” where each cluster represents a “visual word” or “term.” In some works, the vocabulary is called the “visual codebook.” Terms in the vocabulary are the codes in the codebook.

2. **Assign Terms and Generate Term Vector**
   - Assign Terms: Extract features from a novel image. Use Nearest Neighbors or a related strategy to assign the features to the closest terms in the vocabulary.
   - Generate Term Vector: Record the counts of each term that appears in the image to create a normalized histogram representing a “term vector.” This term vector is the Bag of Features representation of the image.

V. CONCLUSION

We have presented a paper to help the blind and partially sighted to recognize the obstacles like steeps, slopes, staircase, vehicles etc., to navigate safely and also to identify the people in front of them by face detection technique. Here we uses Raspberry pi kit with ARM v7 processor along with compatible HD camera for the
image processing. The kit along with the power bank can be placed in the user’s pocket as it is compact and camera can be fixed in the shirt so as to capture and process the image in front of him. The main advantage is, it is cost effective. If it is made in large scale, the cost will get reduce further. In future, we can expect it to be implementing in android app.

REFERENCES


[8] Rapid object detection using a boosted cascade of simple features.


[15] “Efficient Object Motion Detection Based on RGB-D Image” by Yi-Hua Wang, Mind-Hwa Sheu Dept. of Electronic Engineering, Natl. Yunlin University of Science and Technology, Taiwan, Chi-Chia Sun Dept of Electrical engineering, Natl. Formosa University, Taiwan. 2015 International Conference on Consumer Electronics-Taiwan (ICCE-TW).


AUTHORS


Guide: Mr. M. Karthikeyan. M.E.( PhD), HOD, EEE Dept., Vel Tech.
email id: mkarthikeyan@veltechengg.com
Cell: 9952052439