# A Perceptual Field of Vision, using Image Processing

## Virtual Eye

Pruthvi. S, Shama.M.S, Hitesh. V. Harithas, Sandesh. S. Chiploonkar

Department of Electronics and Communication Engineering
BNM Institute of Technology
Bangalore, Karnataka, India
pruthvi.shiva1997@gmail.com, starryshama@gmail.com, hharithas@gmail.com, sandesh3196@gmail.com

*Abstract* - **Currently, estimated facts state that there are more than 285 million visually impaired people around the globe, of which 39 million are blind and the others have a low vision [1]. Approximately 90% of people suffering from blindness are from low-income backgrounds [2]. The main aim of this paper is to provide an efficient visual platform to enhance the perception of the surroundings for a visually impaired user. This is achieved by using the concept of real-time image and video processing. This data is analyzed and compared along with the database which consists of pre-stored data that aids in recognition of the captured image.**

**A head mount camera is used to capture an image on a real-time basis whenever desired. The camera is placed to provide a maximum field of vision and to eliminate the blind spot. The captured image is then processed and compared with the information stored in the database, providing an audio output indicating the desired information. Audio output is provided through bone conduction headphones which communicate audio signals directly with the inner ear. This keeps the outer ear free to be sensitive to the surroundings. A distress alert mechanism is also included as a safety measure to the blind at times of danger. It helps in sending messages containing distress alert signal which contains current location of the user.**

***Keywords: computer vision, image processing, virtual vision, database, audio processing, bone conduction, distress alert.***

## I. Introduction

The main aim of this paper is to develop a device that can provide an efficient visual platform to enhance the perception of the surroundings, for a visually impaired user. The Virtual Eye serves as a simple and efficient method for providing a glimpse of the world through audio or haptic feedback as per the situation under test. A user-friendly device is designed with minimal wirework and mounting accessories. Virtual eye offers an excellent solution for the people with low vision through a low-cost aid. It acts as a warning system providing a reliable software interface with the hardware components. It is used to provide an enhanced and futuristic solution to the common problems faced by the visually impaired community, to reduce the burden of the blind by computing various test case scenarios through image processing and make him independent of the cane and to develop a product, which is both economical and user-friendly in terms of use.

### A. *Digital Image Processing:*

Digital Image Processing is involved in processing a set of images or video frames using mathematical operations. The output of this includes the characteristics and parameters of the image. Image Processing involves applications of any form of signal processing techniques, treating the image as two-dimensional signal. Computer Vision is widely associated to Image Processing, based on certain algorithms for decryption of the physical information stored in an image or video input.

### B. *Python and OpenCV.*

Python is a flexible and multi-purpose coding environment which is a high-level interpretation language. Created and brought to usage by Guido van Rossum in 1991, python has been known for its readability of code. Python provides a free and expressive platform for writing codes. With the aid of tabs and spaces, minimum and optimal code length can be achieved. Python houses several libraries such as the SciPy, NumPy and MatplotLib that enhance the applications of code to perform a multitude of specifically assigned tasks with minimum computing time. The Raspberry Pi single-board computer project has adopted Python as its main user-programming language [3].

### C. *Bone Conduction*

In bone conduction, sound to the ear is delivered by making use of a bone conductor vibrator that rest on the area, just behind the ear. Vibrations transmitted by the bone conductor propagate through the bones of the skull those results in a movement of the fluid in the inner ear. These hair cells in the inner ear are thereby triggered to produce certain electrical signals. These electrical signals are interpreted as sound by the brain [4].

The sound delivery does not depend on the outer ear to collect and clear sound vibration down to the ear drum. It also bypasses the middle ear system needed to transmit vibrations from the eardrum through to the inner ear.

One of the basic requirements to ensure the working of the device is to hold the bone conductor firmly against the head. Bone conduction aids can also be built into hats. Bone Conduction paves way for the visually impaired to make efficient use of the Virtual Eye as the bone conduction technology provides an advantage of no wirework.

## II. Related Works

Henry Candra, Mitchell Yuwono, Rifai Chai, Hung T. Nguyen, and Steven Su [5] describe the usage of a standard and established face detection algorithm in association with a specialised feature descriptor called the E-HOG (Edge-

Histogram of Oriented Gradients). The performance of face detection portrayed with the adaption of this algorithm (96.4%) lays the foundation for developing an algorithm that aids in the modelling of an optimal face detection feature for the visually impaired. An elaborated explanation on the target features for classification of faces (the eyes and the mouth) is provided. Use of an adaptive filter provides optimal detection with minimal computation. The importance of a HOG classifier and a State Vector Machine (SVM) was of utmost importance for the development of an original detection algorithm based on the requirements of the visually impaired.

Syed Tahir Hussain Rizvi, M. Junaid Asif and Husnain Ashfaq [6] address the need for an alternate alerting mechanism, apart from the standard audio output for the visually impaired community that arises in environments such as market places or traffic junctions. Here, there is a need to leave the ears of the blind person free so as that he can respond to stimulus. The mechanism of an alerting system using vibrations through a vibratory motor known as haptic feedback is portrayed.

Dr.M.Geetha, Sangeetha.B, Priyadarshini.T, Sanjana.S [7] provide an insight on the importance of wireless communication through GSM/GPRS and GPS technology in the development of an anti-theft system for vehicles. The transmission of alert messages through the GSM module is employed. The system uses AT commands for transmission of messages to the owner. A GPS module installed is used for the tracking and the co-ordinates are sent through the GSM module. This system provides an alerting and tracking mechanism that are essential in realising the goal of safety at times of distress.

### A. *Motivation:*

Virtual Eye aims at catering the needs of the visually impaired community with the goal of achieving a visual perception of surroundings, thereby boosting self-reliance among the visually impaired community. The realisation of the above mentioned goal is made possible by providing multiple features addressing the basic needs in the life of a user. The motivation for the design and development of each feature is described.

The usage of a GSM and GPS technology is of prime importance in the development of a distress alert mechanism for the visually impaired user. The user is able to access two control buttons to: (a) transmit an alert message (using a GSM module) with his current location (using a GPS module) attached at the time of distress or (b) transmit his current location when asked by his contact through an SMS accordingly.

The motive of virtual eye is to imbibe self-reliance to a visually impaired user. The framework of face detection and recognition algorithm plays a vital role towards this motive. It helps the user in identifying faces of relatives, friends and family provided their data is already present in the database.

Haptic feedback is used to develop an optimal obstacle avoidance system. Sensors are required to cover a maximum field of vision and provide vibrations to the user indicating the position of the obstacle (left, right or ahead).

## III. Types of Blindness

The term blindness is characterized by either complete or nearly complete vision loss, based on which there are various types [8]:

### A. *Total Blindness:*

Complete Blindness is associated with total loss of vision and will have no perception of light. It is featured with central visual acuity of less than 20/400 with normal visual sight.

### B. *Legally Blind:*

Visually impaired fall under the category of legally blind when their visual acuity is less than 20/200, or if the near field of vision is less than 20 degrees.

### C. *Night Blindness:*

Night blindness also known as nyctalopia, is the visual impairment that occurs in dim light or during night. The factors which cause night blindness are namely cataracts, vitamin A deficiency or retinal disease wherein a tunnel vision is created due to collection of dark pigments in retina.

## IV. Problems Faced by The Blind

### A. *Crossing roads:*

Crossing roads or junctions can be tedious for any normal person. The visually impaired have to bare the additional burden of depending on their cane while simultaneously listening to the surrounding noise for danger. Natural instincts suggest them to walk by, in the absence of any approaching noise. There are numerous situations where approaching noise maybe deceiving or cannot be relied upon. Also, presence of nearby stationary objects is not considered which may be the cause for tripping [9].

### B. *Obstacles of greater height:*

The use of a cane aids in the detection of obstacles in proximity to the ground. The risk of colliding with an obstacle at a height considerably above ground level is high. The field of sensing obstacles is restricted only to small height, which is one of the major concerns for the visually imapired user.

### C. *Risk of cane breakage:*

A cane is a hand-held device used primarily to help in the navigation. In areas with heavy traffic, it is possible that a vehicle may run over the cane and break. At this stage the visually impaired person is left helpless and vulnerable.

### D. *Negative effects of Range in existing smart canes:*

The error associated with high range smart canes is of practical nature. In a situation where an output is produced due to an obstacle existing far away from the user, the user may confuse this to be from a nearby object. Also, in place with high number of obstacles, the output from the cane becomes continuous which misleads the user [9].

*E. Cannot rely only on audio output of the device due to presence of external sound:*

An audio output alarming system helps the visually impaired person in a low noise environment. In case of an environment with high noise from multiple sources (example crowded places, markets, junctions), it is natural instinct to judge incoming danger by listening to the surrounding noise. The high volume of the audio output restricts the user from listening to the surrounding noise.

## V. BLOCK DIAGRAM AND WORKING

### IMAGE CAPTURING SYSTEMS

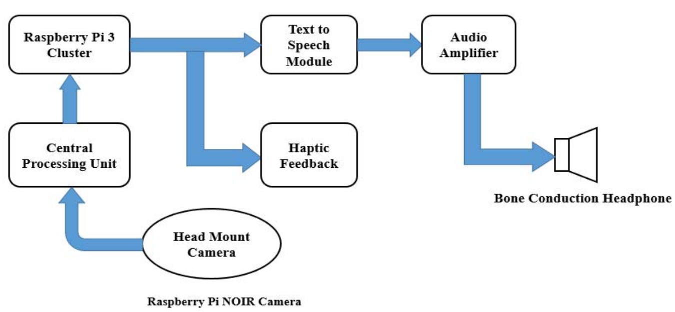



Fig. 1 Block diagram of image capturing system

*A. Head Mount Camera:*

A Raspberry Pi Non-Infra-Red camera is used to capture the real-time image/video of the surrounding objects and transmit the data to the CPU as shown in Fig. 1. This image capturing system is integrated to a wearable device. The camera used has a viewing angle of 120 steradian.

*B. Raspberry Pi Cluster:*

A cluster of two Raspberry Pi 3 Model B behaves as the Central Processing Unit and is used to process the image obtained from the camera in real time. This system is mounted at the belt region of the user and acts the heart of the system to provide various functionalities.

*C. Text to Speech:*

The Text To Speech module is used to convert the text data stored in the database for a particular image to audio format. It is one of the modes of outputs provided to the visually impaired user.

*D. Audio Amplifier:*

The output of the audio signal from text to speech module is amplified before feeding it into an audio output device, in this case, a bone conduction headphone (volume control is provided to the user for ease of usage).

*E. Haptic feedback:*

It is an alternate means of providing an output that involves a vibration motor to alert the user, through a vibratory motion by analyzing the surrounding scenes for any hindrances. This mode of output can be used in obstacle filled environments, to easily traverse across them.

*F. Bone Conduction Headphones:*

The audio signal is amplified by means of an audio amplifier and is output by means of a speaker or Headphone as per the needs and necessity of the user. It plays a vital role in surroundings such as crowded places, busy streets where the outer ears are left open to sense impending danger as it is human tendency to react to stimuli.

### DISTRESS ALERT SYSTEM

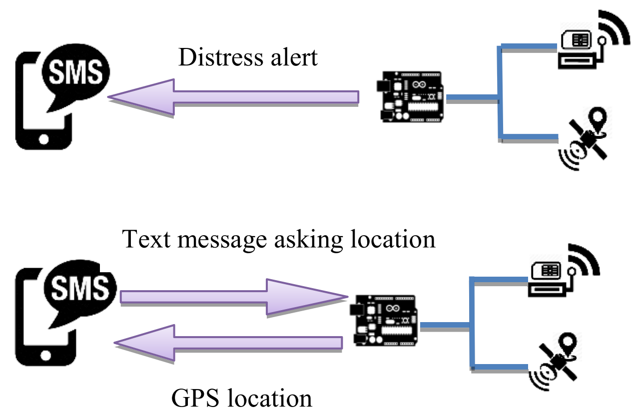



Fig. 2 Distress alert transmission and reception

*A. Mobile Phone:*

It is used as a means to provide a communication interface. It serves by sending and receiving messages between the two communicating parties (the visually impaired user and a pre stored contact). An SMS with a notification buzz is received on the pre stored contact's phone to alert him about the situation. The message communicated contains "Distress Alert" and the current location of a visually impaired user.

*B. Arduino Microcontroller:*

Arduino NANO is a microcontroller that acts as the brain of the system which controls and manages the activity of GSM and GPS module. It offers an array of digital and analog configurable input and output pins as required by the application. The power for the same is obtained by the power system layer residing below.

*C. GSM + GPRS module:*

Global System for Mobile Communications + General Packet Radio Service, SIM900A is a 2G based module that is used for providing communicating medium. It communicates at a baud rate of 9600. It can be programmed to make and receive calls, send and read messages. This is achieved by using "Attention Commands" (AT commands).It routes the message written in the code to a destination number specified (pre stored contact) at the push of a button at the visually impaired user's end.

*D. GPS module:*

Global Positioning System, Neo 6m is a GPS module that is used to extract and notify the current location of the user. This latitude and longitude are concatenated along with the Google maps link and is sent as a message using the GSM module.

## VI. Design: The Four Layer Approach

### A. Introduction

The virtual eye offers a multi feature package as described above and therefore requires a practical realization of the system with a specific design. The intended system is modeled according to a multi-layer oriented design as shown in Fig. 3. The major segments of the overall system are classified and put into separate layers with appropriate facility for input power connections, interconnection between layers and connections for the output peripheral devices. Virtual Eye accommodates four layers each for a specific task with pre-defined objectives.

### B. Description:

The four layers of virtual eye are namely the power system layer, distress alert layer, R-pi cluster layer and the obstacle avoidance layer. The objective of such a design is to obtain the outputs using a modular approach, where separate tasks are assigned to each layer. A control mechanism is provided by means of two separate remote controls with buttons dedicated to perform subtasks of its associated layers.

- Layer 1: Power system layer:

  This is the bottom most layer in the system. The layer consists of 6 Li-ion batteries arranged in series and parallel configuration based on the required voltage values defined to drive the sub systems in the upper layers. An electrical provision for charging the batteries is provided. The output from the batteries is given to the buck converter. A buck converter is used to regulate the output voltage to 5V. Therefore, the overall objective of this layer is to provide optimal power to the above layers using battery management.

- Layer 2: Distress alert layer.

  It consists of an Arduino nano as the controller to drive a GSM module (Sim 900A) and a GPS module (Neo 6m). The input power to the arduino nano is provided by the power system below. The GPIO pins of the Nano are used to interface GSM and GPS module by connecting appropriate pins to Tx and Rx. The software code written for the nano includes a standard distress alert message. However, it is essential to provide a control mechanism for the user to exploit this feature only at the time of need. Therefore, a remote control with a button is connected to this layer which when pressed activates the functioning of the distress alert mechanism. A GPS module is used to provide exact location of the user (latitude and longitude points) concatenated with Google maps to provide a map like view to the pre stored contact. This information is sent as a message to the contacts number.

- Layer 3: R-pi cluster layer:

  It is the heart of face detection and recognition feature provided by virtual eye. The objective of this layer is to aid the system in capturing an image from the camera, detect and recognize a pre stored face. There is also provision made to register a new face in the database that can be used for recognition in the future. This layer houses two raspberry pi modules working as a cluster for faster processing of data. An IR camera is connected to the cluster and used for real time image/video capturing. The remote control for this layer comprises of four buttons with specific tasks, they are, creating a database of images of a particular face, training using a cascade trainer, recognition and shutting down function. The user is constantly notified through an audio message regarding the function chosen.

- Layer 4: Obstacle avoidance layer:

  It primarily focuses on alerting the user of nearby obstacles with the use of haptic feedback. It consists of an Arduino Nano the controller to control three ultra-sonic sensors placed at 120º from each other. This is done to ensure a wider field of vision to detect obstacles. The output of the sensors is given to two vibration motors. The code is written such that when an obstacle is detected to the left, the vibratory sensor at the right vibrates indicating the user to move in that direction and vice versa. In a case where obstacles are detected to the left and the right, both the vibratory motors vibrate indicating the user to change his orientation to move further.


Fig. 3 Multi-layer oriented design

## VII. Multifeature Package

### A. Face Detection and Recognition

Face detection is a mechanism through which the dataset containing facial features is created. The dataset contents are trained using the trainer program, and the accuracy of training is identified. It extracts the user ID of individual images and trains the processor from the predefined set of finite images. The trained data is stored in a file with extension .yml. This file is later used in the prediction algorithm to facilitate the detector stage. The algorithm outputs confidentiality values along with the user ID.

### B. *Distress Alert*

At a time of emergency or when the user is in distress and is in need of immediate help he can make use of this feature and send an SMS to pre- stored contacts along with his location details.

The visually impaired user has an option of sending a message of distress with the press of a push button. It also provides a feature of sending a message containing only the current location to the concerned recipient. This function is performed using a GSM SIM900A module, a GPS NEO 6M module and an Arduino Nano microcontroller.

### C. *Obstacle avoidance*

Three ultrasonic sensors are arranged at an angle of 120º from each other to cover the posterior viewing angle. The sensors emit short and high frequency pulses of sound at regular intervals. When these pulses hit the object in the line of sight, an echo signal is reflected back to the sensor. Using the time taken for the echo signal to reach the sensor, the distance is calculated using the concept similar to that of sonar or radar. Any obstacle in front of a visually impaired user can be sensed and notified through haptic feedback with the help of mini vibration motors. Various pulse patterns are assigned for obstacles at different distances and their relative position in front of the user.

## VII. EXPERIMENTAL RESULTS

### 1. FACE DETECTION AND FACE RECOGNITION

The face recognition process is carried out in three steps, they are:

1. Dataset Creation: Contains gray scale images of faces captured for each user id.
2. Trainer: Training phase which calculates the cumulative accuracy of the faces in data set and stores it in a file (file extension .yml) for each user id by extracting random features from each image.
3. Detector: The detection stage of an individual user. Face recognized is given as an audio output when the confidence value falls within the threshold.

### A. *Dataset Creation:*

Initially, the data set is created which contains facial features extracted from the face of an individual user. Features such as head, eyes, and lips are extracted and stored in the database of images. Only the cropped image of the face is stored in gray scale, this helps in removing irrelevant features and details from the captured image. The dataset of images stored are as shown in Fig. 4.

### B. *Trainer:*

The microprocessor is trained for the dataset of images. It extracts the user id of individual images and trains the processor from a predefined set of images. The trained data is stored in a file with file extension .yml. This file is later used in the prediction algorithm to facilitate the detector stage. Fig. 5 shows trainind data stored in the system.

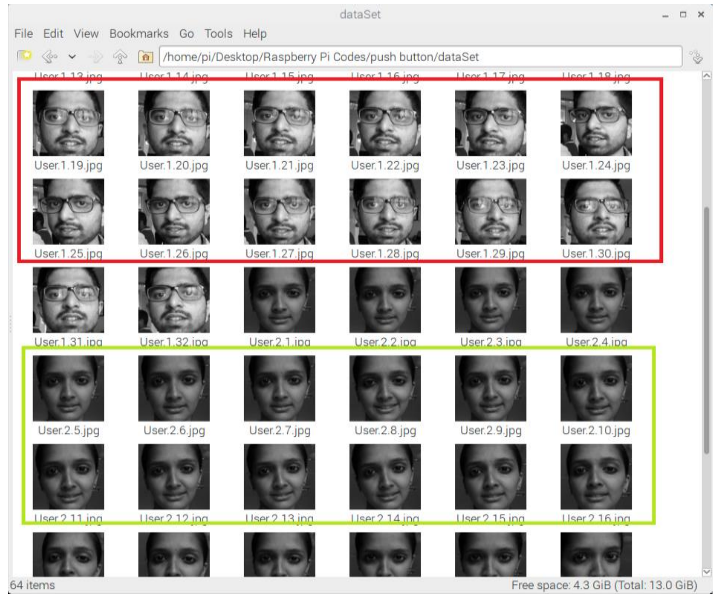
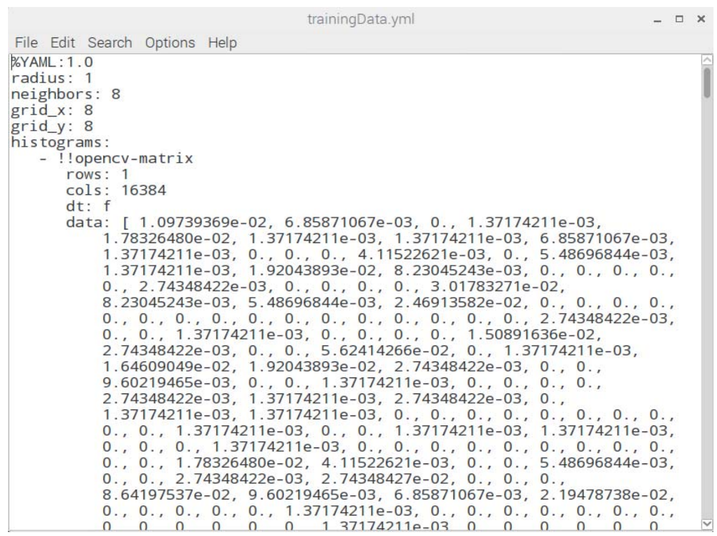



Fig. 4 Dataset Stored


trainingData.yml
File Edit Search Options Help

```
%YAML:1.0
radius: 1
neighbors: 8
grid_x: 8
grid_y: 8
histograms:
   - !!opencv-matrix
      rows: 1
      cols: 16384
      dt: f
      data: [ 1.09739369e-02, 6.85871067e-03, 0., 1.37174211e-03,
          1.78326480e-02, 1.37174211e-03, 1.37174211e-03, 6.85871067e-03,
          1.37174211e-03, 0., 0., 0., 4.11522621e-03, 0., 5.48696844e-03,
          1.37174211e-03, 1.92043893e-02, 8.23045243e-03, 0., 0., 0., 0.,
          0., 2.74348422e-03, 0., 0., 0., 0., 3.01783271e-02,
          8.23045243e-03, 5.48696844e-03, 2.46913582e-02, 0., 0., 0., 0.,
          0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 2.74348422e-03,
          0., 0., 1.37174211e-03, 0., 0., 0., 0., 1.50891636e-02,
          2.74348422e-03, 0., 0., 5.62414266e-02, 0., 1.37174211e-03,
          1.64609049e-02, 1.92043893e-02, 2.74348422e-03, 0., 0.,
          9.60219465e-03, 0., 0., 1.37174211e-03, 0., 0., 0., 0.,
          2.74348422e-03, 1.37174211e-03, 2.74348422e-03, 0.,
          1.37174211e-03, 1.37174211e-03, 0., 0., 0., 0., 0., 0., 0., 0.,
          0., 0., 1.37174211e-03, 0., 0., 1.37174211e-03, 1.37174211e-03,
          0., 0., 0., 1.37174211e-03, 0., 0., 0., 0., 0., 0., 0., 0., 0.,
          0., 0., 1.78326480e-02, 4.11522621e-03, 0., 0., 5.48696844e-03,
          0., 0., 2.74348422e-03, 2.74348427e-02, 0., 0., 0.,
          8.64197537e-02, 9.60219465e-03, 6.85871067e-03, 2.19478738e-02,
          0., 0., 0., 0., 0., 1.37174211e-03, 0., 0., 0., 0., 0., 0., 0.,
```


Fig. 5 Training data stored as .yml file

### C. *Detector:*

The trained data from the trainer stage is taken to extract information for the prediction algorithm. The prediction algorithm outputs confidentiality values along with the user id. Confidentiality values range from 0 to 255, with 0 representing 100% accuracy and 255 representing 0% accuracy. From the training algorithm that we have designed a system which has a confidentiality value ranging from 35 to 45, we are able to achieve accuracy in the range of 82.4% to 86.3%. This accuracy can the enhanced by increasing the number of sample images stored in the dataset. Fig. 6(a) and 6(b) shows the output of detection phase.

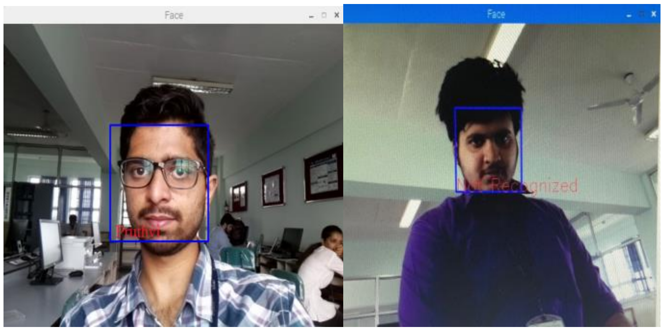


Fig. 6(a) Face Recognized as "Pruthvi" (b) Face "Not Recognised"

Fig. 6(a) and 6(b) shows the face detection feature. The first image shows the person who is detected and recognized i.e. the features of the face have been extracted and stored in the database are compared and the face is recognized by means of indication of the name. The second image is the one showing the face which isn't recognized that is the face features have not been stored in the database.

### 2. Distress Alert

The visually impaired person in any kind of confusion and distress such as if he is lost or if he is injured at a place has an option of sending a message of distress with the press of a push button. When the button is pressed, few saved contacts (monitoring end) get the message that the visually impaired person is in distress. On the press of the button for the second time, the location is sent to the monitoring end following the request to the visually impaired human.

This function is performed using a GSM SIM900A module and an Arduino UNO processor connected with respect to the transmitter and receiver ports. The Distress alert operation is shown below with respect to the user end (visually impaired human) and the monitoring end.

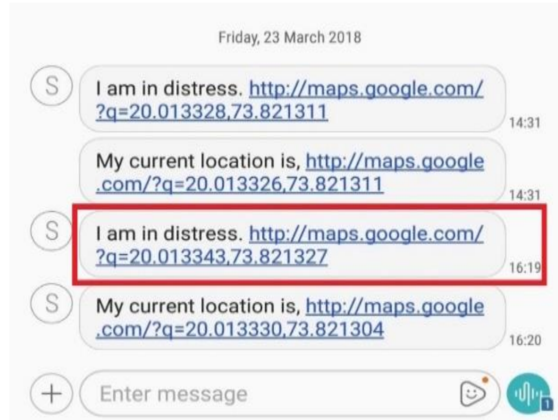
i. At the user end:

The visually impaired person in distress presses the push button. The GSM module sends the message of distress to the pre-stored contacts. At the monitoring end, the user gets a notification in the form of a buzzer beep. The location is in the form of coordinates (latitude and longitude) is obtained with the help of Neo6M GPS module and sent using SIM900A module.

ii. At the monitoring end:

The monitoring end such as the user's relative or services like hospital receives the distress message from the user. The monitoring end receives the location in the form of coordinates and concatenated with the google maps link which helps the concerned third party in tracking the location.

Fig 7(a) shows distress message sent to the receiver, 7(b) shows the location in google maps.



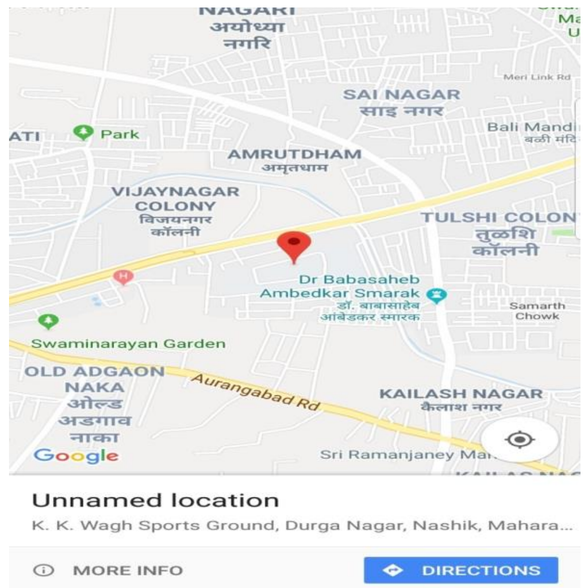
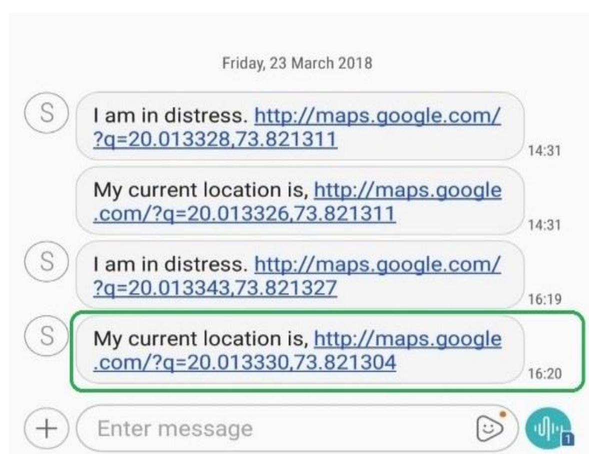




Fig. 7(a) Distress Alert Message received along with a google maps link opened as shown in 7(b).

Fig 8(a) shows the current location of the user when a concerned third party send "Where are you?", 8(b) shows the location in google maps.



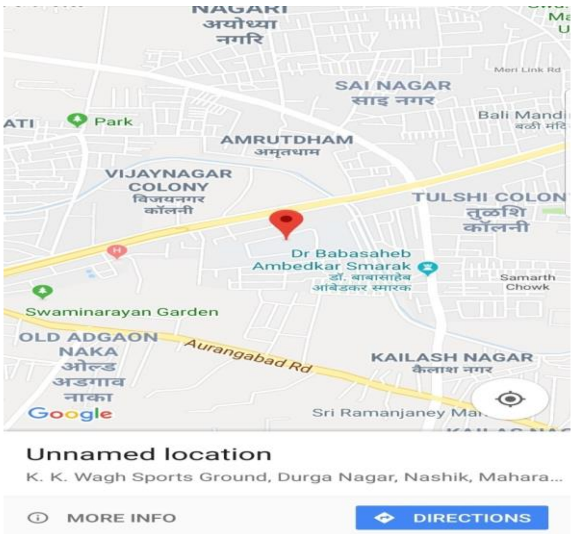



Fig. 8(a) Current Location Message received along with a Google maps link opened as in 8(b)

### 3. Obstacle Avoidance

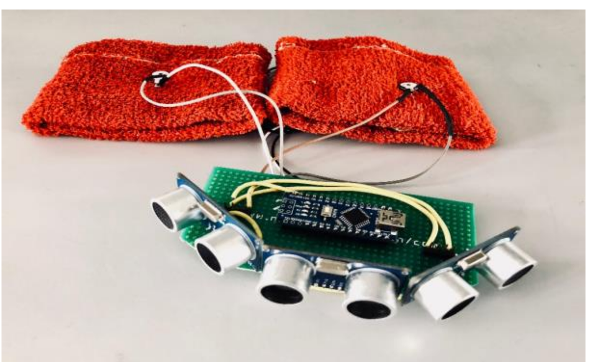


Fig. 9 Ultrasonic sensor arrangement along with min vibration motor

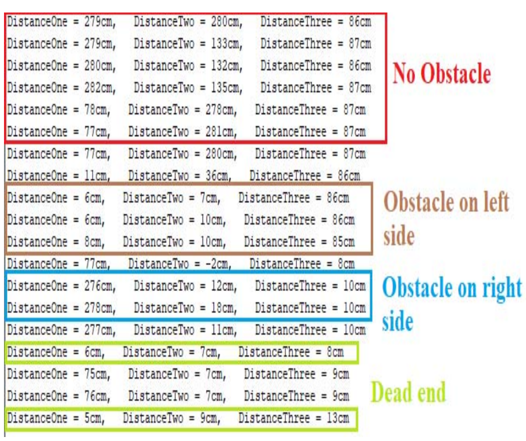



Fig. 10 Obstacle avoidance output on serial monitor

Fig. 9 and Fig. 10 show the arrangement of ultrasonic sensor at 120º from each other along with mini vibration motor, and the serial monitor readout respectively.

## VIII. Conclusion And Future Scope

Virtual Eye aims at providing an edge over the current problems faced by the visually impaired user, based on the concept of "adding life to years". An audio representation of the world is sketched to the user. Technical Additions to this product will increase the chances of the blind participating in competitions. It provides an overall improvement in everyday activities, by widening the scope to pursue their career. The Virtual Eye can be modified accordingly to aid the blind in undertaking professional courses with ease and benefit them by improving their status. The product is hence expected to strike on a large scale and serve the purpose of the visually challenged all over the world.

## References

[1] "Blindness and visual impairment", time 10:25, 09/02/2018, Available: http://www.who.int/mediacentre/factsheets/fs282/en/

[2] "Real-time Local Navigation for the Blind: Detection of Lateral Doors and Sound Interface" – article by M. Moreno, S. Shahrabadi, J. Jose, J.M.H. du Burf, J.M.F. Rodrigues, *4th International Conference on Software Development for Enhancing Accessibility and Fighting Info – exclusion (DSAI 2012), pp. 74-82*.

[3] Python (Programming Language), time 15:45, 27/03/2018, Available: en.wikipedia.org/wiki/Python_(programming_language)

[4] Are bone conduction headphones good enough yet? , time 13:43, 22/03/2018, Available: www.theverge.com/2016/10/24/13383616/bone-conduction-headphones-best-pair-aftershokz

[5] "Classification of Facial-Emotion Expression in the Application of using Viola-Jones and Edge-Histogram of Oriented Gradient" (Henry Candra,Mitchell Yuwono, Rifai Chai, Hung T. Nguyen, and Steven Su), 2016, *pp 423-426*.

[6] "Visual Impairment Aid Using Haptic and Sound Feedback" (Syed Tahir Hussain Rizvi, M. Junaid Asif and Husnain Ashfaq), *2017 International Conference on Communication, Computing ND Digital Systems (C-CODE), pp 175-178*.

[7] "Anti-theft and Tracking Mechanism for Vehicles using GSM and GPS" by (Dr.M.Geetha, Sangeetha.B, Priyadarshini.T, Sanjana.S), *2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), pp 252-255*.

[8] Different types of blindness, time 10:45 08/01/2018, Available: www.livestrong.com/article/83641-eyes-send-images-brain/

[9] "Challenges That Blind People Face", time 13:45, 10/01/2018, Available:www.livestrong.com/article/241936-challenges-that-blind-people-face/