Navigation System for Blind - Third Eye

Samartha Koharwal, Samer Bani Awwad, Aparna Vyakaranam

Abstract--- Visual impairments people with severe condition are unable to move independently. In this fast moving world, these people are generally left underprivileged. Few methods have been used to help them and provide them with some level of mobility comfort. Conventional methods such as trained dogs or a cane are not reliable enough in providing sufficient information of possible hindrances. Moreover, training and managing dogs is challenging task. There are some guidance systems which use RFID technology. However, this technology cannot be used in an outdoor open area. In this paper, an AI based system titled "Navigation System for Blind - Third Eye" is proposed. In order to support blind and visually impaired people's mobility indoor and outdoor, this work proposes a simple electronic guidance embedded vision system which is configurable and efficient. The system utilizes three types of devices including IR sensor, sonar sensor and camera. A microcontroller processes the reflected signals from all devices in order to classify front obstacle. This system can be fasten to a hat or to a pen-sized hand mini stick. The system provides affordable and reliable solution and also helps the impaired people to be highly self-dependent.

Keywords: IR, RFID,There,This system ,Conventional

I. INTRODUCTION

As derived from "World Health Organization" report and fact sheet updated on October 2017 on visual impairment, the estimated number of people live with vision impairment is about 253 million; 36 million are totally blind while 217 million suffer from moderate to severe vision impairment. Globally, the main cause of vision loss is the chronic eye diseases while the top two causes of visual impairment are in-corrected refractive errors and un-operated cataract. In this fast moving world, visually impaired people are left behind and not treated equally. To help them and provide them with some level of comfort, many solutions and techniques have been tried and developed. One of these techniques is called orientation and mobility. In this technique, a specialist helps the visually impaired and blind people and trains them to move on their own. They are trained to depend on their other remaining senses to move independently and safely. Another method is through using a guide dogs. In this method, the dogs are trained specially to support the movement of the blind people. The dogs navigate around the obstacles as an alert to the user to change his way. However, it is difficult for visually impaired and blind people to understand the complex direction provided by these dogs. Additionally the cost of these trained dogs is very high.

Nowadays, different types of canes are also being used such as the white cane, the smart cane, and the laser cane. However, these tools also have several constraints. The cane is unhandy since it has long length which also makes it difficult to keep in public places. Additionally, the cane has limitations in recognizing obstacles.

Many techniques have been developed recently to enhance the blind people's mobility. These techniques are developed based on signal processing and sensor technology. However, these tools tend to fail in a crowded area or in an area where there are high volumes of electronic waves.

In order to support blind and visually impaired people's mobility indoor and outdoor, this work proposes a simple electronic guidance embedded vision system which is configurable and efficient. The system utilizes three types of devices including IR sensor, sonar sensor and camera. A raspberry pi 3 model B microcontroller processes the reflected signals from all devices in order to classify an obstacle. The proposed guidance system is able to determine the obstacle distance, in addition to material and shape characteristics of the obstacle. Furthermore, the system can name some of the detected objects. Moreover, neither cane nor other marked tool are needed to be carried by the user. Like other systems, the proposed system can be fasten to a hat or to a pen-sized hand mini stick. It has high immunity to both ambient light and object's color.

II. RELATED WORK

This section of related work gives a general overview of Artificial Intelligence, including its definition, history, advantages, disadvantages, complications, types and deployment methods. Artificial Intelligence is "The science and engineering of making intelligent machines" (John McCarthy, 1956). In addition to computational intelligence, there have been other names which had been proposed for artificial intelligence as well, such as synthetic intelligence or computational rationality. The Artificial Intelligence term also describes a property of machines and programs. We are surrounded by artificial intelligence and in some way or the other we happen to interact with it. These interactions with machines are limited by their ability to understand and process the information. There are basically four types of AI:

Type 1: Purely Reactive

This is the most basic type of AI in which computer only processes one situation/environment and acts on it. It does not have any model, design or even the concept of the wider world. It specializes only in one area (Futurism, 2018).

Type 2: Limited Memory

This AI allows machines to consider past information and apply it to current solutions which are preprogrammed representations of the world. It has just enough information to make a proper decision and execute appropriate actions (Futurism, 2018).

Revised Manuscript Received on March 10, 2019.

Samartha Koharwal, Department of Computer Engineering and Computer Science, Manipal International University, Malaysia.

Samer Bani Awwad, Department of Computer Engineering and Computer Science, Manipal International University, Malaysia.

Aparna Vyakaranam, Department of Computer Engineering and Computer Science, Manipal International University, Malaysia.

Type 3: Theory of Mind

This AI has the capacity to understand thoughts and emotions. It can understand feelings, motives, intentions and expectations and has the capability to interact socially. It is yet to be built (Futurism, 2018).

Type 4: Self-Aware

This types of AI is aware of themselves. It analyzes its internal states, predict the feelings of others, and can make abstractions and inferences. They are the future of machines (Futurism, 2018).

Complications in AI

There are a lot of complications when trying to create an intelligent system. Older AI or a simple AI is just a list of conditions for when to take what action based on those conditions. This is arguably not intelligence, and developing the true intelligence involves an understanding of input and how this input relates to the output (Zhihong Zeng, 2009). Fuzzy logic, which is designed after studying human's excellent ability gives an approximate answer without any real values. However fuzzy logic has many complications (Lotfi A. Zadeh, 1999). Conceptually, computer processes works in bits. Hence, processing words or concepts is consider a complicated task for the computer.

There are complications with image processing as well such as recognizing different photos locations from photos on the internet because they have a lot of variation. Computer cannot simply model the world from these photos because an average internet has immense variations. Image processing requires the data to be consistent. Rendering a 3D image is more complicated as the data has to be correct and more accurate with least variations (Noah Snavely, Steven M. Seitz, Richard Szeliski, 2008). Computer cannot simply detect an image. Finally, using all the set fields of AI to develop strong AI is very complicated. Designing and developing a system which has sentiments and conscience would require us to fully understand how the human brain works which we do not as of now.

Natural Language Processing

Human language is very complex and subjective at all times. The current form of communication with a machine is through input devices like mouse, keyboard or verbal commands. This is totally different from how humans communicate with each other.

This basic issue of accurately representing ideas with images or words, is creating significant challenges for natural language processing (Deb Roy, 2005).

Frameworks with natural language processing would have the abilities to express and decipher dialects at human interpreter levels, comprehend the distinction between a blue water and blue sky, and process commands like "give me that purple thing down there" into physical activity (Deb Roy, 2005).

Using current natural language processing techniques, machine can recognize spoken words. Furthermore, it can even do basic level of translations. However, comparing to fully understand the complexities of human conversation, this is still primitive. Machines cannot summarize a novel as of now (Lotfi A. Zadeh, 1999) or process any data or conversation that requires moderate conceptual knowledge. Devillers' research in detection of emotions from human conversations (Laurence Devillers, Laurence Vidrascu, Lori Lamel, 2005) gives much wider view and helps in understanding the meaning behind different pronunciation and words. Havasi's Open Mind Sense (OMCS) project provides a better understanding of fundamental concepts that words represent (Catherine Havasi, 2009).

Knowledge Management and Fuzzy Logic

Havasi's ConceptNet is a large graph of simple concepts that link to related concepts (Catherine Havasi, 2009). It is basically a part of OMCS project. It also includes AnalogySpace, a process and method, which analyzes data and get a sense out of it (Catherine Havasi, 2009). The graph allows system to learn from new inputs and make interpretations with existing data. The small part of graph is shown in Figure 1. The figure shows a small part of the ConceptNet graph showing all the connections between "cake" and all related concepts.

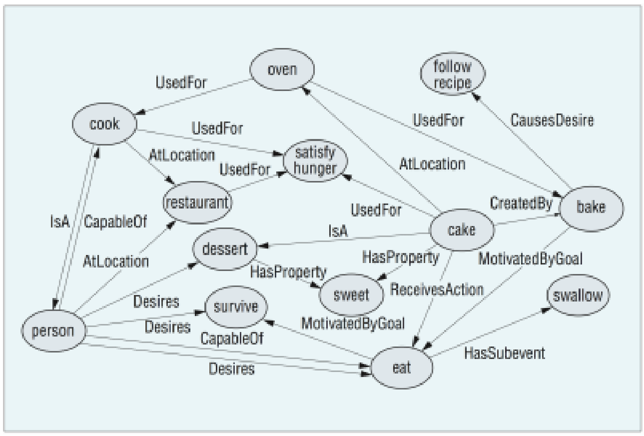

Figure 1: ConceptNet Graph

Humans have this unexplained ability to make approximations without using any real numbers or computations (Lotfi A. Zadeh, 1999). Zadeh shows that fuzzy logic is the perfect way a computer can mimic the ability of approximation we humans naturally have (Lotfi A. Zadeh, 1999). Through computations with imprecise numbers and using perception instead of measurement allows for fuzzy logic. Fuzzy logic is useful when the computer needs to process something which is not precise or the data is not enough and when threshold for error is high (Lotfi A. Zadeh, 1999).

Image Processing and Computer Vision

Human-centered design is trying to move away from current method where a machine simply responds to given commands from keyboard, mouse or simple verbal commands. This requires an increased ability to process images and perceive information (Zhihong Zeng, 2009). Most existing image processing software utilizes 2D spatial analysis for facial recognition by looking for geometric shapes and edges in the face (Zhihong Zeng, 2009). However, research suggests that most accurate behavioral

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

judgments of human action comes from analyzing both facial expression and body language (Zhihong Zeng, 2009).

It might now be dealing with image processing of humans, but Snavely has developed an algorithm which helps machine create 3D models of areas from a series of 2D images (Noah Snavely, Steven M Seitz, Richard Szeliski, 2008). Researchers aim to release this algorithm "into the wild" to analyze all of the images on the internet to create a 3D model of the earth. However, the algorithm is not robust enough. As of now, it is unable to handle the variability of photos on the internet to consistently identify the location of the image it is processing.

Anyhow, if the location is fixed and the images used are from google image search of a specific location, an accurate 3D model is generated (Noah Snavely. Steven M. Seitz, Richard Szeliski). System with the ability to generate 3D model of their environment from a series of 2D images will be a great improvement to the current methods of computer vision. This ability of the system is illustrated in Figure 2. The figure shows all of the images returned by a Google Image search "Yosimite Half Dome" through Snavely's algorithm, an accurate 3D model was generated (right ref. Fig 2). Then, a famous photo by Ansel Adams (left ref. Fig 2) was analyzed with the 3D model and the model showed its best prediction for where Adams was when he took the photo of

Figure 2: Half Dome Prediction

III. METHODOLOGY

This project proposes the design of portable AI based "guidance system for blind –Third Eye", which benefits the visually impaired community and also helps in their day to day mobility. Third eye provides the visually impaired community a new way to visualize the world by explaining them about their surroundings. The whole system is controlled by 'Raspberry Pi' microcontroller.

Third eye harnesses the maximum capabilities of Raspberry Pi microcontroller which has enough potential to up hold the system with one advantage being the inbuilt graphic card. The prototype uses various sensors such as IR Sensors, Sonar Sensors and a Camera module which helps the system to gather the required data. Additionally, text to speech module is used to talk to the user.

Python forms the heart of the system. It is used to program the whole system which helps the raspberry pi microcontroller to communicate with all the sensors. It then processes the collected data and converts it into information which is finally delivered to the end user. The IR sensor is used to map the object's shape and size while the sonar sensors get the data about the distance of the object at regular intervals. Camera module plays an important role as it takes the pictures which is then processed using image processing technique to properly visualize the object. All the information is processed and converted to text which is then fed into a text to speech module. The text to speech module delivers this information to the end user in his/her ear using headphones.

System Design

This system is developed using various technologies, which work together in order to run Third Eye smoothly.

The system design consider the following hardware components:

- Raspberry Pi 3
- IR sensors
- Sonar sensors
- Camera module
- Headset

The system design considers the following software components:

- Python programming language

Raspberry Pi operating system and programming platform

- OpenCV
- Text to Speech Module

Raspberry Pi

The Raspberry Pi is a low cost, portable computer that can use a normal TV or a computer monitor as its display and standard keyboard and standard mouse as input devices. It supports various programming languages such as python, java, scratch etc. It can perform any task that one would expect from a desktop computer. It can be used to browse the internet and to play high definition video. Additionally, it can be used to make spreadsheets, word processing, and playing games.

Furthermore, the raspberry pi has something which regular desktop computers does not have. It has the ability to interact with the outside world. It has been used in a lot of projects of all kinds, from music machines to weather stations. It has been used also to create tweeting birdhouses with infra-red cameras. Raspberry pi microcontroller is shown in Figure 3.

Figure 3: Raspberry Pi Microcontroller

Infrared Sensor

IR sensors are small microchips with a photocell that are designed to either emit or receive infrared light and in some cases sensor does both. They are used in detection of remote

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

objects, it can be anything, TV is just one example. Inside any IR emitting device, there is a matching IR led, which emits IR pulses to tell the other device what action to do. Human eyes cannot perceive IR light, that's why it takes more effort to test the IR sensors. IR sensors are specially filtering IR light; they cannot detect visible light. The demodulator inside the IR sensor looks for modulated IR signal at 38 KHz. Normal IR will not be detected, it has to blink at 38 KHz. IR sensors are digital, so either they sense 38 KHz IR signal and give low output (0V) or they do not; output (3-5V).

This system uses the IR sensor to detect the presence of an object at a certain distance. This is possible through using a pair of IR sensors (Transmitter and Receiver), the transmitter emits IR and the receiver receives the light bounced back from the object. IR sensor is shown in Figure 4.

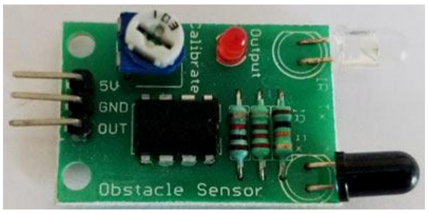

Figure 4: Infrared (IR) Sensor

Sonar Sensor

Sonar sensor is a device that can measure the distance to an object using sound waves. It emits the sound wave at a specific frequency and the sensor waits for that wave to bounce back from the object. By perceiving the time interval between emission and receiving, it calculates the distance between the sonar sensor and the object.

However, the sonar sensor might not detect some objects. This is because sometimes wave might not bounce back properly to the sensor and it can go somewhere else. In this situation, sonar sensor will not be able to predict the distance accurately. Sonar sensor is shown in Figure 5.

Figure 5: Sonar Sensor

Camera Module

This module takes pictures and sends them to the microcontroller. Frequency of taking images can be altered depending on the program and usage. This module has Sony IMX219 8-megapixel sensor. It can be used also to take high definition videos. It even supports low-light clicking and is capable of 1080p30, 720p60 and VGA90 video modes. Camera module is shown in Figure 6.

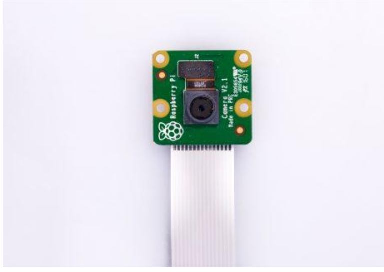

Figure 6: Camera Module

Use case diagram

A use case diagram is a diagram that describes the relationships among use cases and actors within a system. Use case diagram provides an overview of all or part of the system or organization usage requirements in the form of essential model or a business model. It also allows one to communicate the scope of a development project.

In this system, there is only actor named User. The user will be able to use all the functions or inputs of this system. The following Figure 7 shows the use case diagram comprising of the above-mentioned actor and its use.

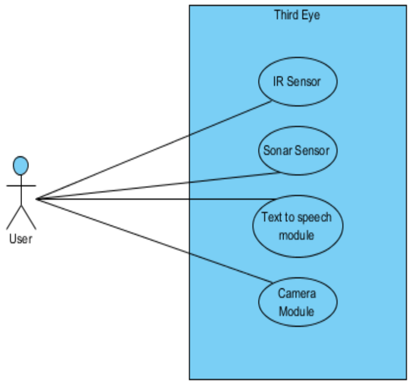

Figure 7: Use Case Diagram

System Architecture

A system architecture is the conceptual model of the project that defines the structure, behavior and functionalities. This section focuses on the system architecture and explains about the functionality of each segment of the system. It also expresses how sensors communicate with each other and work together to give the desired output. The system architecture is shown in Figure 8.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

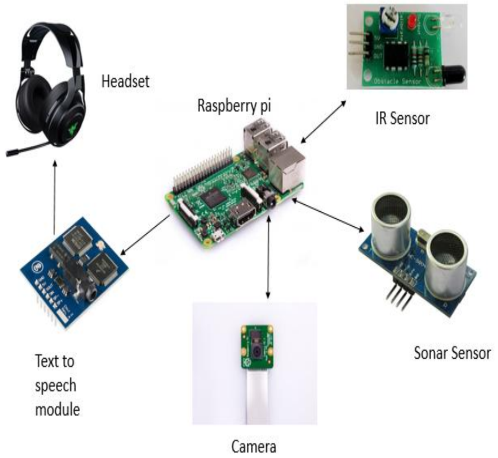

Figure 8: System Architecture

Figure 9 shows the flowchart for the Third Eye system. The flowchart shows the sequence of steps and decisions required to perform the Third Eye system process.

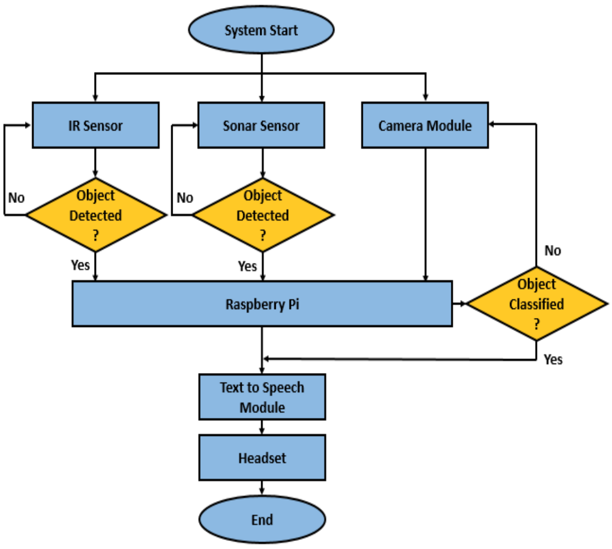

Figure 9: Flowchart of Overall System

The camera is considered the eye of this system. It continuously captures still images which are sent to the raspberry pi microcontroller. The raspberry pi microcontroller takes those images, processes them using artificial intelligence image processing algorithms and generates a matching response which is sent to text to speech module. OpenCV framework is used to develop the AI model which gives Cafeemodel for the raspberry pi. Python program is used to get the input through camera and handle the interaction between the model and the input. Then, it shows the output on the screen as a text. Text to speech module takes the text as an input and converts it to an audio output which goes directly into the user's ears.

Infrared sensors are placed in a way to match the angle of vision in a human. They are used to calculate the distance between the user and an object located at a far distance. Sonar sensors are placed at an angle of 30 degrees from the vertical plane facing downwards. This inclination of 30 degrees helps sonar sensors to scan for the objects placed on ground near the user. Hence, they can be used to calculate the distance from an object lying on the ground adjacent to the legs of the user. Both of these sensors send the data to raspberry pi which in turn processes the data and sends the desired output to the text to speech module. This again takes text as an input and converts it to an audio response which reaches the user through headset. This full process keeps going simultaneously and continuously until the user decides to switch off the device.

IV. RESULTS AND DISCUSSION

The Third Eye – AI based Navigation System for Blind is developed and integrated as per the design documented in this section. The system is implemented using raspberry pi, sonar sensors, infrared sensor and camera as the hardware base, python is used to program the hardware and OpenCv is used to setup and run the object detection model with python.

The system has 3 separated components, sonar sensor, infrared sensor and camera which work simultaneously. Python controls the sonar and infrared sensors. For object detection, OpenCv provides framework for python to control the camera and perform object detection.

Speech Synthesis on Raspberry Pi

Flite (festival-lite) is a small text to speech synthesis engine. It is fast and open source engine developed at Carnegie Mellon University (CMU). It was primarily designed for small embedded machines and large servers. It was designed as an alternative to Festival for voices built using the FestVox suite of voice building tools.

Implementation

The system is composed of three different components namely: sonar sensor, infrared sensor and camera. All the data from all three components is sent to raspberry pi which processes it and gives the output to the user. The structure, implementation, and execution is similar for sonar and infrared sensors. Camera works on a different structure and algorithm. The features of the system include:

- Detect the distance of the object (sonar and infrared sensors)
- Detect the direction of the object based on the sensor that detect the object.
- Detect the type of the object (using object detection algorithm)
- Convey the output to the user through headphones.

The following sections elaborate the complete process.

Detect the distance of the object (Sonar Sensor)

In order to get the distance of the object from different directions (left or right), two sonar sensors are placed in each direction. Each sonar sensor gets the distance using sound-waves after the wave hits any object and bounces back. It is capable of calculating distance every second. However, in this work, the system is configured to calculate the distance every three seconds, due to the simultaneous running of other components. Sonar sensor in the circuit is shown in the Figure 10.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Figure 10: Sonar Sensor

The output of the sonar sensor can be seen in real time on the screen with the actual distance value from the object. The output is shown in Figure 11.

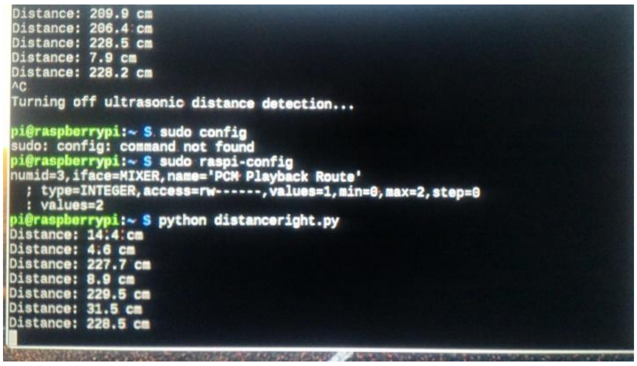

Figure 11: Sonar Sensor Output

The above output screen shows the distance of the user from the object in every three seconds. At the same time, this information is being processed by the raspberry pi and being transmitted to the user through headphones. This guides the user about the position of the object.

Variation in the distance

The sonar sensors are always not accurate as there is no ideal condition in an actual world. There might be a variation between the detected distance and the actual distance. These variations can be reduced by using better and accurate sensors. Figure 12 shows the variation in the distance.

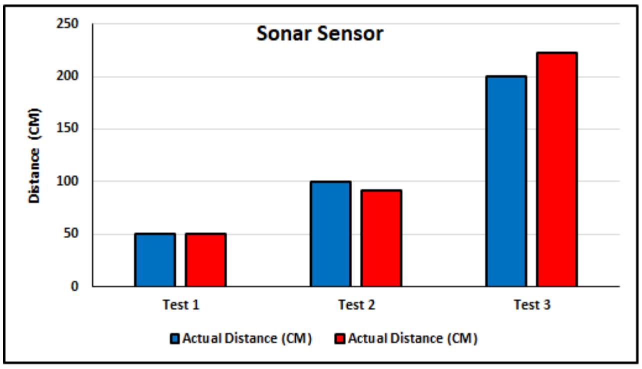

Figure 12: Actual versus Detected Distance through Sonar Sensors

Detect the distance of the object (Infrared Sensor)

Infrared sensor is placed in the forward direction to get the distance of the object which is in front of the user. It uses infrared to detect the distance unlike sonar sensor which uses sound-waves. Thus, infrared sensor is more precise in obtaining the distance. However, infrared sensor comes with its own limitation such as the range is not long. In addition, if the object is made of glass, sometimes the object may not be detected. The infrared connected to the circuit is shown in Figure 13.

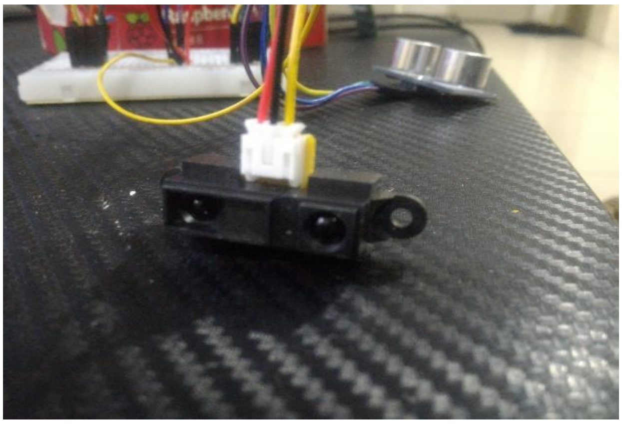

Figure 13: Infrared Sensor

The output of infrared sensor is calculated in real time. It conveys the distance to the person through the headset. Different infrared sensors come with variable range. The infrared has been tested using three ranges; 5, 15 and 30 cms.

Variation in the sensor

The infrared sensor is more accurate than sonar sensor as it detects object using a beam of infrared light. The beam generally travels in the straight line. However, the variations in detected distance may still occur. Figure 14 shows the variation in the distance.

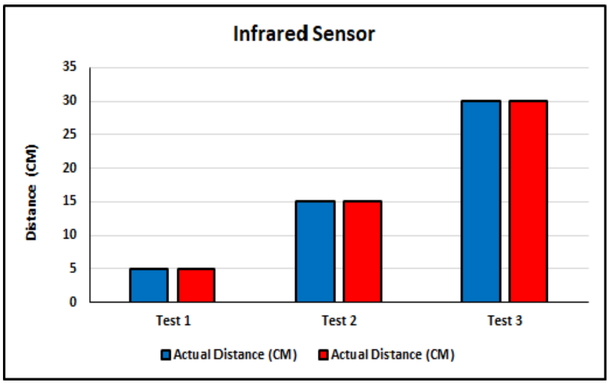

Figure 14: Infrared Sensor Graph

Detect the direction of the object

Detecting the distance of the object can only help the user partially, direction should also be known. In order to get the direction, two sonar sensors are used and placed in left and right directions. To cover the whole 180 degrees eyes vision, one infrared sensor is also placed in the forward direction. These three sensors provide full coverage of the surroundings for the user with directions. This helps raspberry pi to guide the user to walk freely without any collision with any object.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Detect what is the object (using object detection algorithm)

This is one of the most important parts of the system. In order to detect what is the object, the system needs object detection algorithm to run on. This system uses OpenCv for the model and the backend and python to connect to the model and to run everything in the front end. The algorithm works on 'caffemodel' which is locally stored on the raspberry pi. When the program is running, python connects with the model and starts taking the video. Then, the objects captured in the video are compared with the model and the closest match is shown on the screen and sent to the headset. However, raspberry pi 3 does not have high power GPU and CPU, thus the frame rate is not high. Currently, the system can detect up to nine types of objects which are sofa, table, chair, bottle, TV-monitor, aero-plane, train, person and dog. The system is compatible with raspberry pi camera module as well as USB cameras. Figure 15 shows the camera in the circuit.

Figure 15: Camera in the Circuit

The output of object detection can be seen in real time on the screen with the object name as shown in Figure 16.

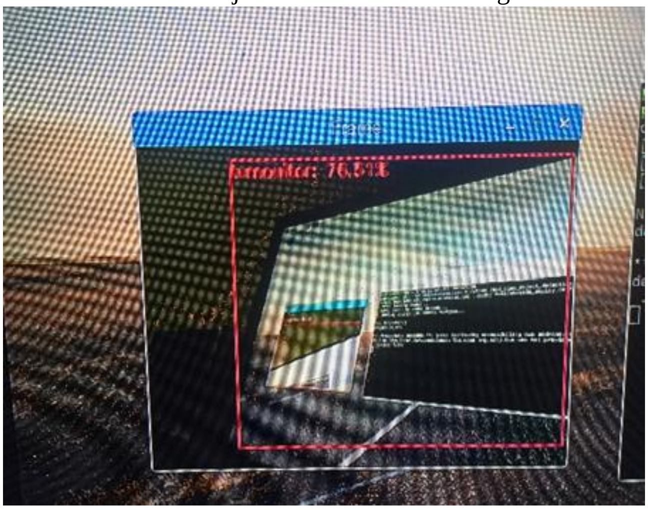

Figure 16: Object Detection Output

Circuit Implementation

Circuit of the system is designed and implemented in such a way that all the components work simultaneously without interfering with each other. Current to sonar sensors is controlled using resistors and all the components are carefully placed to make it as small as possible and to provide handiness. The actual circuit is shown in Figure 17.

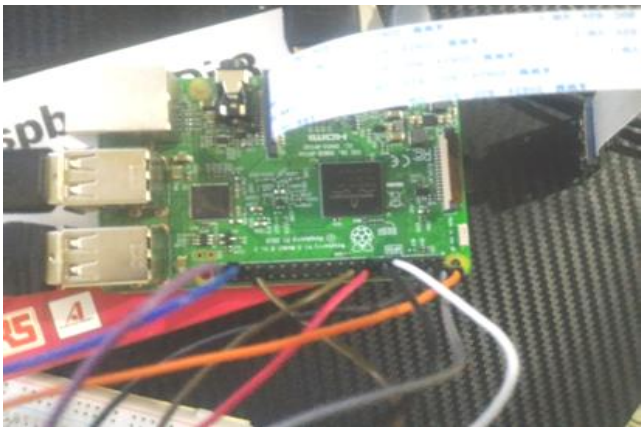

Figure 17: Raspberry Pi Connections

V. CONCLUSION AND FUTURE ENHANCEMENT

In today's world, disability of any kind for any person can be hard and it is the same case with blindness. Blind people are generally left underprivileged. It is very difficult to give a vision to a blind person. In this paper, a new AI based system called "Navigation System for Blind - Third Eye" to control the navigation of a blind person has been proposed and developed. This AI based system offers a simple electronic guidance embedded vision system which is configurable and efficient. The system helps blind and visually impaired people to be highly self-dependent by assisting their mobility regardless of where they are; outdoor or indoor. Results show that all the sensors work properly and give accurate readings, though the range of the prototype sensors is not high. Object detection algorithm utilizes 100% CPU which makes the raspberry pi hot and thus, in future system, two raspberry pi are recommended to be used; one for object detection and one for all the sensors.

REFERENCES

1. Laurence Devillers, Laurence Vidrascu, Lori Lamel, "Challenges in real–life emotion annotation and machine learning based detection," Neural Networks, 1st ed., Elsevier, 2005, pp. 407 – 422.
2. Terrence Fong, Illah Nourbakhsh, Kerstin Dautenhahn, "A survey of socially interactive robots," Robotics and Autonomous Systems, 1st ed., Elsevier, 2003, pp. 143 – 166.
3. Catherine Havasi, et al., "Digital Intuition: Applying Common Sense Using Dimensionality Reduction," Human – Level Intelligence, 1st ed., IEEE Computer Society, 2009, pp. 24 – 35.
4. Deb Roy, Ehud Reiter, "Connecting language to the world," Artificial Intelligence, 1st ed., Elsevier, 2005, pp. 1 – 12.
5. Noel Sharkey, "The Ethical Frontiers of Robotics," Science, 1st ed., AAAS, 2008, pp. 1800 – 1801.
6. Noah Snavely, Steven M. Seitz, Richard Szeliski, "Modeling the World from Internet Photo Collections," Int J Comput Vis, 1st ed., Springer Science + Business Media, LLC, 2008, pp. 189 – 210.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

7. Lotfi A. Zadeh, "From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions," IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications, 1st ed., IEEE Computer Society, 1999, pp. 105 – 119.
8. Zhihong Zeng, et al., "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1st ed., IEEE Computer Society, 2009, pp. 39 – 58.
9. Futurism. (2018). Types of AI: From Reactive to Self-Aware [INFOGRAPHIC]. [online] Available at:
10. http://www.futurism.com/images/types-of-ai-from-reactive-to-self-aware-infographic/
11. Raspberry Pi. (2018). Raspberry Pi - Teach, Learn, and Make with Raspberry Pi. [online] Available at: http://www.raspberrypi.org.
12. Digital Image Processing - Kenneth R. Castleman, Prentice-Hall, 1996.