Insight System for Visually Impaired Persons: Assistive Computer Vision and AI-Based Environmental Understanding for Independent Mobility and Healthcare Support

Prof. Giulia Conti[1]
[1] University of Melbourne, Department of Assistive Technologies and Biomedical Engineering, Melbourne, Australia

ABSTRACT

The world health organization (WHO) announces that the percentage of about 80% of visual impairment is either preventable or curable with treatment, but 20% of peoples are suffering due to this disability, so we decided to help them with the existing and future technologies with this paper. This paper proposes the idea of voice-based assistive for the visually impaired person. It involves text extraction from scanned image using Optical Character Recognition (OCR) and converting the text to speech by the Google Text-to-Speech (gTTS) tool. This is for blind peoples to recognize the products in the real world by extracting the text on image and converting it into speech. The proposed method is carried out by using Raspberry Pi and portability is achieved by using a battery backup. The proposed system has various modules like Ultrasonic sensor for obstacle detection and makes alert as voice, Personal assistant using Machine Learning, Chat-Bot for Android users and an IOT enabled location tracker inside the Chat-Bot. This system helps millions of peoples who are suffering from blindness.

KEYWORDS: Ultrasonic Sensor, OCR, gTTS tool, Personal Assistant, Machine Learning, Chat-Bot, IOT enabled Location tracker.

1. INTRODUCTION

Vision is an important factor in human life. The visually impaired peoples are increasing every year due to health- related issues and many other reasons. The most difficulty faced by a visually challenged person is difficulty in reading the text and detecting the object in front of them. Various developments in the field of technology helps the blind person to developing camera-based and voice-based system alongwith computer vision tools with the existing useful products such as an ultrasonic sensor, Google Text to Speech tool, Optical Character Recognition System, etc.

Insight is a prototype for blind peoples to do their activities without the help of others. The device can be worn on the head of a blind user. The proposed system is developed by using Raspberry Pi and the portability is achieved by using a battery backup. Thus, the user could use thisdevice anywhere and able to use anytime. The device has the first login using a username and password. After login, the various modules suchas the ultrasonic sensor, GPStracker, chat-bot begins to start their processing. The Ultrasonic sensor starts working when there are any obstacles in front of the blind user and alerts them using the voice with distance. Text extraction from an image is carried out by using Optical Character Recognition (OCR). It is the tool for which the text from an image is then converted to digital format, from any scanned document or a photo of a document or a scene-photo. The digital image is converted to text using the Tesseract OCR engine it will helps in detecting the white spaces outline etc...It conjointly checks the standard of recognized text. In this system, the conversion of text to speech output is done by using an e-Speak algorithm.Here we use Google text to speechtool.

The personal assistant based on the latest technology machine learning in the system helps the blind person to know about the wanted information such as date, time, location etc. Machine learning is not only a single technique or atechnology.It is the field of computational science which it incorporates numerous technologies to create systems that can learn based on the data'susing their environment and can make predictions, actions when compares with the new situations. The personal assistant enables the microphone for taking inputs and outputs through the attached headphone. The chat-bot an android based application in the system helps the person to communicate with their relatives or friends. The text message received is then converted into voice at the user side also theinput will be given a voice, then it will be converted as text and then send to the relatives or friends

of the blind user. The IOT enabled GPS tracker in the system helps their relatives or friends to track the location using thechat-bot.

2. PROPOSED SYSTEM

This system helps the blind person to browse the text from any quite notes and provides him with higher support through voice. The system is designed to alert the blind user if any obstacle is in front of him. Thus, the user

safety could beensured. Also, an android based chat-Bot helps the person to communicate with their friends or relatives also they could track him using the same application. The proposed system has several modules such as Camera, Ultrasonic Sensor, Chat-Bot, GPS, Microphone, Personal Assistant, Headphone, LCD, and Touch Sensor. The text recognition from an image taken by using a camera and recognizes the text from the image using an Optical Character Recognition (OCR). Conversion of the recognized text file to voice output by using the gTTS tool. The captured image is first converted to grayscale and then filtered using a Gaussian filter to reduce the noise in the image. Here we use an adaptive Gaussian thresholding is used to reduce the noise foundon that image. The filtered image is then converted to digital format. Then the digitized image is cropped so that the portions of the image with no characters are removed. The cropped frame is loaded to the Tesseract OCR so as to perform text recognition. The output of the Tesseract OCR will be a text file which will be the input of the gTTS tool. The Ultrasonic sensor emits sound waves at high frequency that the humans couldn't hear.Then the sound will gets reply back to the origin, thenit will calculate the distance based on the time required. Then the system will alert the blind user by voice with the distance. The Chat-Bot an android based application in the system helps the person to communicate with their relatives or friends. The text message received is then converted into voice at the user side and the input will be given a voice, then it will be converted as text and then send to the relatives or friends of the blinduser.
Fig 2.1 shows the system design of the proposed system.

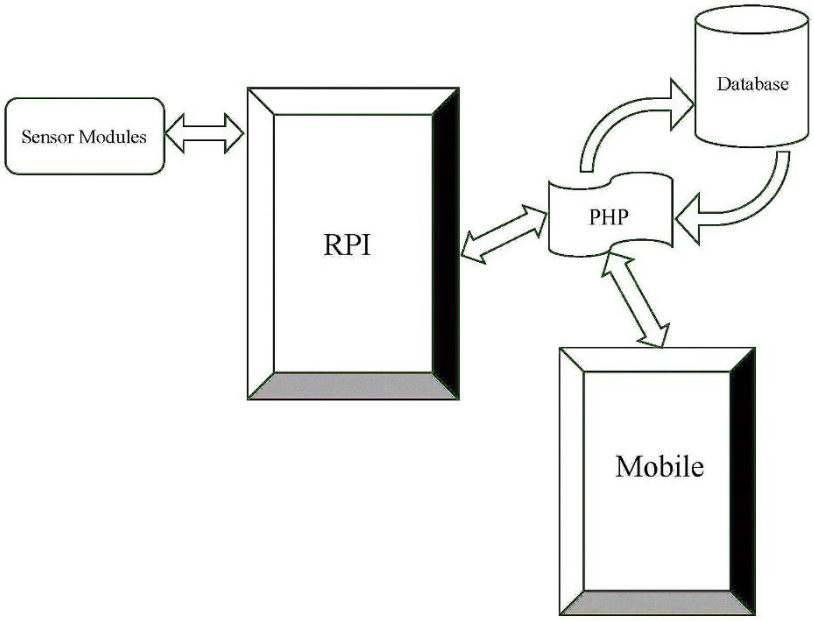

Figure 2.1

An IOT system consists of sensors or devices which will communicate with the cloud through connectivity. Once the data get hits into the cloud server then the software processes. After thatit willdecidewhich actions to be performedsuch as, sends an alert to the user, automatically adjusts the sensors based on the user actions, etc. The GPS tracker in the system helps their relatives or friends to track the location using the Chat-Bot byusing IOT server. This helps in saving the track history of the blind person for their safety. The microphone is a type of transducer which converts energy from one form to another form. Microphones convert sound waves into the

audio signal. Here we use the microphone for taking input for Personal Assistant. The Personal Assistant based on the latest technology Machine Learning helps the user to communicate with the device and to know the attributes like time and date, text information, etc. Machine learning is not only a single technique ortechnology.It is the field of computational science which it incorporates numerous technologies to create systems that can learn based on the data'susing their environment and can make predictions, actions when compares with the new situations. The Headphone converts the electrical signals to sound waves for the output to the user from the system. The LCD is used as another output of the system. It is not for the blind user, itis for the system diagnosis. The Touch sensor is similar to that of a simple switch. It is used for the function selections and for different purposes like a trigger for image capture, friend selection in chat-bot, to enable Personal Assistant and systemdiagnosis.

The system is portable, which is achieved by providing a battery backup. Thus, the user could carry this device anywhere.

3. LITERATURE SURVEY

Most of the existing systems for the guidance of visually impaired people are limited for few functions only. Most systems are built in MATLAB platform and use laptops so they are difficult to carry.

A camera based text reading is suggested in the paper[1] to help the visually impaired person read the text prese nt in the captured images. This paper process face detection when a person enters into the frame by using the mode control. The proposed idea involves the extraction of text from the scanned image using Tesseract Optical Character Recognition (OCR) and the conversion of a text by an e - Speak tool(google text-to-speech), a process that allows visually impaired people to read the text from the captured image. The proposed hardware is designed using Raspberry Pi and by the use of battery the portability could have achieved.

From this paper,we have taken the concept of how to perform the text extraction from an image and how toprocess.

The paper [2] proposes an Ultrasonic Navigation based system to help the Visually Impaired person. In two modes, hurdle detection mode and fixed mode, the system can be operated. Using Arduino board, we use an ultrasonic sensor and water sensor in the hurdle detection mode to avoid obstacles. In this mode, the system uses the proposed system to detect solid and liquid obstacles and bring corresponding messages via Bluetooth to the blind person through voice. The fixed mode provides the information and guidance to walk from one place to another safely by assigning a fixed route in the proposed system from source to the destination location. The system also provides instructions to the blind to travel through various places using the GPS navigation system. An Android application is used to send messages via Bluetooth. This system provides complete guidance and protection to a blind person in different situations. From this paper we have taken the concept of how to perform the GPS tracking and its methods to done.

The paper [3] proposes the deals with the obstacles in the path and their distance from the user are detected using an Ultrasonic Sensor and an Arduino board connected to a smartphone or PC. The distance at which the Ultrasonic Sensorsdetect the obstacles accurately. The classification of objects which could be determined using a smartphone or a PC. The correct or accurate feature selection is the big task because of the large amount of feature selection. This will cause detection slower. These processes will help in combining the results from Ultrasonic Sensors and Image Processing technique to alarm the user about the distance from it and the type of an obstacle. The combination of Ultrasonic Sensors and CV algorithms used for detecting obstacles in the user's path to enhance the efficiencyof the system that it will detect the glass, doors, windows, etc. If the user uses only a CV system for Object Detection. Then the images do not indicate the presence of glassesor mirrors in the user's path and there is a possibility of the user to collide with it. In such a case, the Ultrasonic Sensor reply back ultrasonicwaves from a mirror or glass door,thus the user will get alertabout the presence of an obstacle in the path. The Ultrasonic Sensor provides information about the distance andthenumber of nearest objects has been detected during their path on the basisof a number of Ultrasonic Sensors used in it. The image processing algorithms which helps in detecting the texts and the objects from captured images. In the Object Detection system, theyuse the combination of two sensors. One is Ultrasonic Sensors and another is the Camera (smartphone or webcam) to capture images for the further process to determine the information about the object.The user will get alerted

about the type of object in theuser's path through voice. The Ultrasonic Sensor gives accurate ObjectDetection resultsalso compares the different Object Detection techniques to determine the bestperforming algorithm which giving maximum accuracy for Object Detection and further process.

From this paper, we have taken the concept of how the obstacle detection and their various functions will perform.

4. HARDWARE IMPLEMENTATION

The various hardware component used in the device is Raspberry Pi, ultrasonic sensor, camera, touch keys, USB microphone. Raspberry Pi is a mini computer which can be programmed for particular tasks. The Raspberry Pi 3 Model B is the third generation Raspberry Pi which is used in the system development. The Raspberry Pi 3 Model B has a powerful processor, 10x faster than the first-generation Raspberry Pi. Raspberry Pi works in an open source platform. It has 1GB RAM ,64 Bit CPU, 4 x USB ports, 4 pole Stereo output and Composite video port ,10/100 Base Ethernet socket, CSI camera port for connecting the Raspberry Pi camera, DSI display port is used for connecting with the Raspberry Pi touch screen display. A Micro SD port is usedfor loading your operating system and storing data. Micro USB power source helps in the running of Raspberry Pi. USB ports available on

this board are used to connect the camera with raspberry pi. Three GPIO pins are used, for capturing an image, for more control and for shutting down the system respectively. The board is operated in such a way that the code starts executing when it is powered ON. The audio output is available through the audiojack.

A compactable camera on the device is used for image capturing. It has autofocusing capability with a resolution of 1280 X 720 which is capable of capturing good quality images. The USB powered camera is used in orderto connect it with the Raspberry Pi board. The captured images are sent to Raspberry Pi and all the image processing was done. The voice output is available through the audio jack. It can be heardusing theheadphone.

An Ultrasonic sensor measures the distance by using ultrasonic waves. The sensor head emits an ultrasonic wave and receives the wave reflected back from an obstacle. Ultrasonic Sensor measures the distance to the target by measuring the time between the emission and reception of the waves. An optical sensor has one transmitter and one receiver, whereas an ultrasonic sensor which it uses a single ultrasonic element for both theemission and reception. In a reflective model ultrasonic sensor, it consists of a single oscillator which it emits and receives the ultrasonic waves alternately. This enables the miniaturization of the sensor head. Ultrasonic sensors are a reliable and cost-effective solution for distance detection, calculation, and obstacledetection.

The Touch sensor works similar to that of a simple switch. If there is contact with the surface of the touch sensor occurs then the circuit will be closed inside the sensor and there is a flow of current will happen. When the contact is released then the circuit will get opened and no current flows. So, the touch sensor willhelp the user to communicate with the system using touch.
Figure 4.1 shows various components connected to the Raspberry Pi.

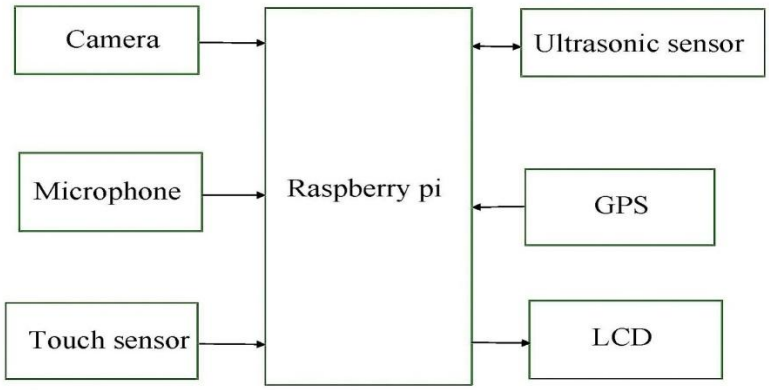

Figure 4.1

The system willstart processing whenever a user presses a button specially designed in the user interface. The camera is placed on that device helps in captures an image of the scene. Then the search for text areas willbe performed using several different methods by exploiting edge, color, and morphological information. If there any text areas were foundat the initial input image, then the camera will zoom in to obtain more detailed images of eachtext areas. Then the characters are recognized and read out to the blind person using a Headphone.

5. SOFTWARE IMPLEMENTATION

Raspberry pi works with Raspbian OS which is derived from the Debian operating system. We use several software's for the working of this prototype.

Raspbian OS is Raspberry Pi's Debian - based operating system.The other versions of Raspbian OS are RaspbianStretch and Raspbian Jessie. In 2015, the raspberry foundation introduces the Raspbian OS as the primary raspberry series operating system. Mike Thompson and Peter Green are the developers of Raspbian OS as an independent project. The operating system is still under in developing and the support will be also available. Raspbian is highly optimized for the low-performance ARMCPUs.

Python-Tesseract is that the tool for python that it'll helps in acknowledge and browse the text foundonpictures. Python-Tesseract might be a wrapper for Google's Tesseract-OCR Engine. It is also useful as a complete invocation script for Tesseract, as it will browse all image varieties supported by the Python Imaging Library along with jpeg, png, gif, bmp, tiff, and others, while Tesseract - OCR supports tiff and bmp only by default. In addition, if used as a script, Python-Tesseract can print the recognized

text rather than writing it to a file.

gTTS (Google Text-to-Speech), a Python library which helps in converting the text to speech andthe Command Line Interface (CLI) tool helps in interface with Google Translates text-to-speech API. It will writes the file from an audio file and then stores as digital format. It options versatile pre-processing and tokenizing, furthermore as automatic retrieval of supported languages.

OpenCV-Python is a python library which helps in resolve computer vision issues. it's a cross-platform computer vision library as an open source that began as a research project at Intel in 1999 by Gary Bradsky, and therefore the 1st unharness came to go into 2000. It can be programmed by using several programming languages like C, C++, Python, Java, etc. This will help in writing more powerful image process functions and high-level algorithms. Open CV helps a wide range of image processing applications and it can be used simply also it is easy to handle.

Android Studio is the only and officially integrated development environment (IDE) for Google's Android operating system. It is developedby the JetBrains' IntelliJ IDEA software and is specially designed for Android development. It is available todownload on several OS platforms such as windows, Linux, macOS. We use Android studio for developing the Chat-Bot application for an android user. The chat-bot will help in communication with their colleaguesor Friends or Relatives. And the Chat-Bot will have gps navigation system.

Visual Studio Code editor is a programming platform developed by Microsoft for Windows, Linux, and macOS. It will help with debugging support, embedded Git control, highlighting syntax, good techniques for code completion, snippets, and code refactoring. It can also be customizableallowing the users tochange the editor's theme, keyboard shortcuts, and preferences of the editor. Visual studio is afree and open source code editor. The compiled digitized data are freeware for private or commercial use. We use visual studio code Editor for the development of Graphical User Interface (GUI) and several important developments of the proposed system using thepython-PyQt5.

PyQt5 is a collectionof Python bindings for Qt v5. It will help in developing the graphical user interface for a particular application. PyQt5 is implemented in C++ itself because of the enhancement of the functionality of that applications. In our project the GUI is designed using this PyQt5.

Spyder is used to edit the python code in Raspberry Pi. It is a powerful tool for coding or programming the

device using Python. It provides many built-in integrations with many popular scientific packages such as NumPy, SciPy, Pandas, IPython, QtConsole, Matplotlib, SymPy, etc. It has much more built-in features also its functions could be unlocked even further via first- and third-party plugins. It can also be used as a PyQt5 extension library for the advanced editing processes.

The given flowcharts (fig 5.1 and fig 5.2) refers to the flow ofexecution of the proposed system. The explanation is as follows.
The system first initializes all the components of the proposed system and starts checking the conditions. When the condition (count++) occurs, it will start running. When the condition (count==1) then the image will be captured. When the condition (count==3) then the image will get converted to text and reads out. When the condition (PA button==high) then the microphone will starts working and reads the input as voice and converts it to the text format. Then this text will get compared with the trained data set and gives the response through theheadphone.

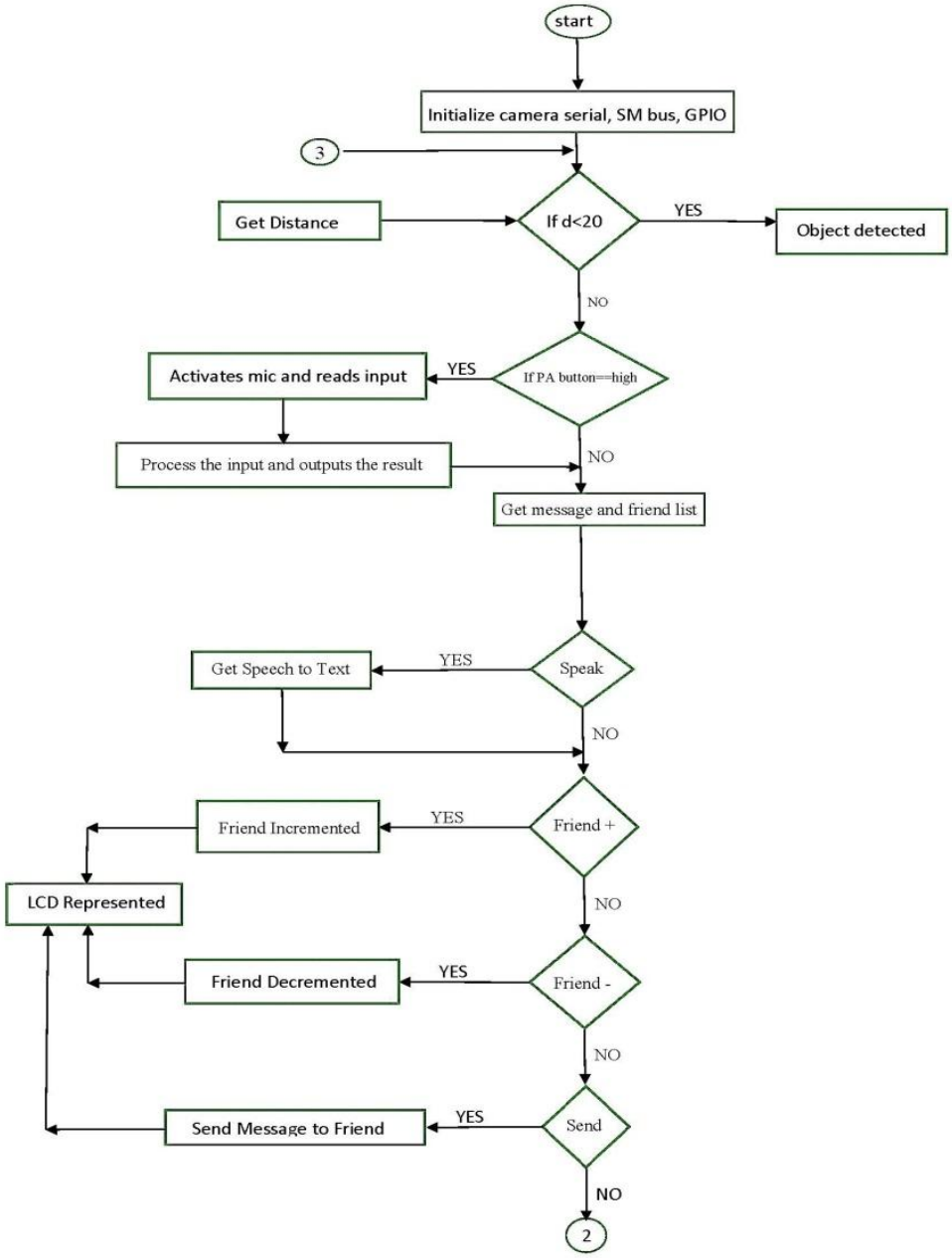
start
Initialize camera serial, SM bus, GPIO
3
Get Distance
If d<20
YES
Object detected
NO
Activates mic and reads input
YES
If PA button==high
Process the input and outputs the result
NO
Get message and friend list
Get Speech to Text
YES
Speak
NO
Friend Incremented
YES
Friend +
LCD Represented
NO
Friend Decremented
YES
Friend -
NO
Send Message to Friend
YES
Send
NO
2

Figure-5.1

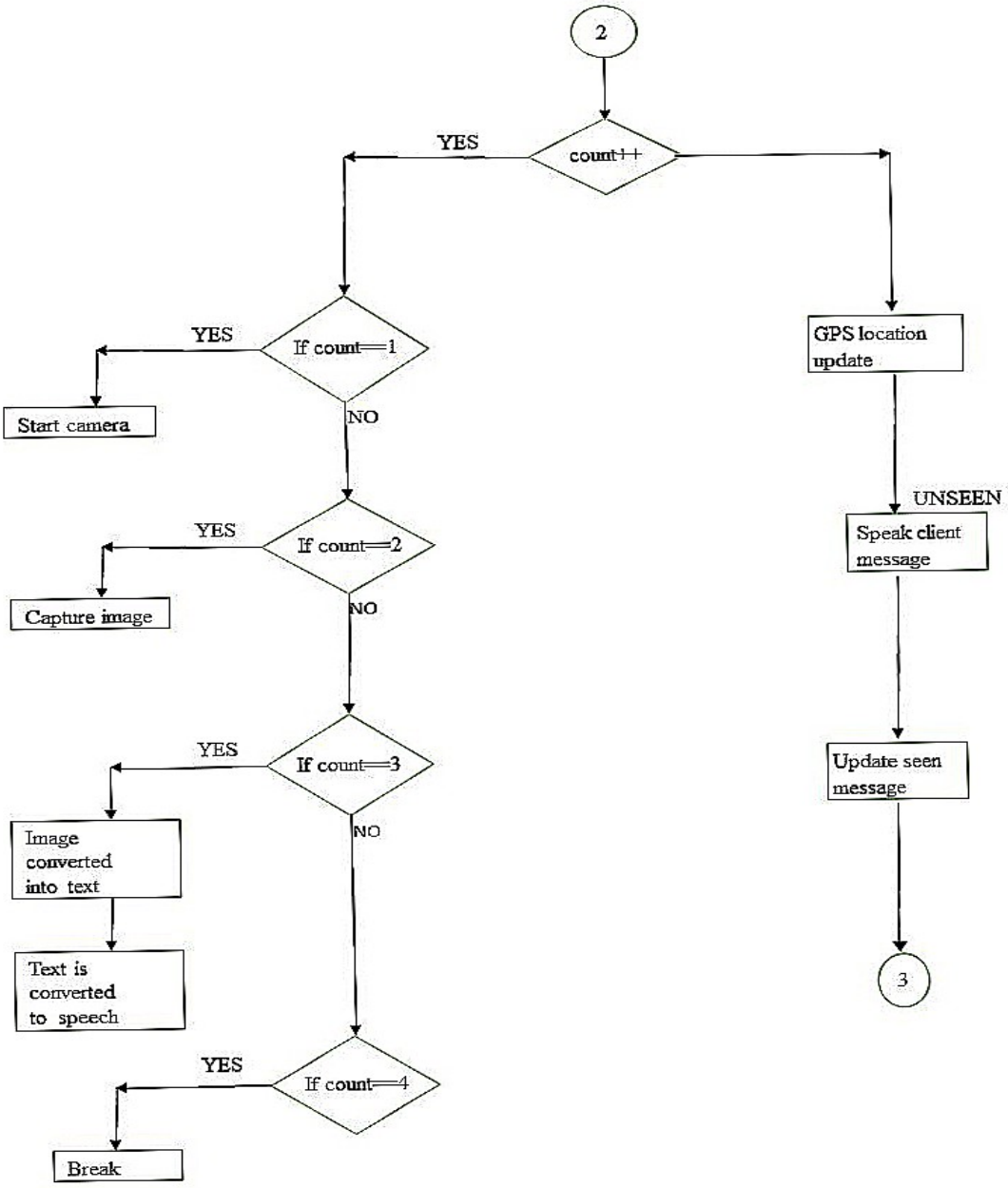

Figure 5.2

While thecondition "speak" achieves then the microphone enableand collect your voice, then it converts into text and stored locally. When the condition (Friend +), then the user could select their friend in ascending order. When the condition (Friend -), then the user will get the previous friend. Using the last two conditions theuser couldselect any of their friends and send the message to a friend while the condition (send) satisfies. The Ultrasonic sensor activates and finds the distance ofthe object whenthe condition (d<20) satisfies.The GPS location data suchas the user's current latitude and longitude will get uploaded to the IOT server using the GPS module and then it will get saved in the server.

\

6. OUTPUT

Insight is a portable device which can wear on the head. The battery backup is used, so the user could carrythe device anywhere and able to use at any time. The device helps the visually impaired person in the following ways. Text extractionfrom the scanned imageusing Optical Character Recognition (OCR) and convert the text to voice as

an output using the Google Text- to-Speech (gTTS) tool. Obstacle detection by Ultrasonic Sensor and gives an alert about the danger and thedistancefrom the user to the obstacle in front of the blind person.Anandroid based Chat-Bot application help in communicating withtheirfriends and relatives. Inside the Chat-Bot, an IOT enabled location tracker is included. It will store the location history of the blinduser, sothe androidusercould find the person wherehe is at the instant.Personal Assistant based on Machine Learning helps in the user interaction with the device and it will respond to what he wants to know at the moment.

7. CONCLUSION

The proposed system is a voice-based system which helps the visually impaired person to recognize the real-world products. The device is designed in such a way that, they could carry anywhere they go so portability is solved. This system helps millions of peoples who are suffering from blindness. The most difficulty faced by a visually challenged person is difficulty in reading the text and detecting the object in front of them.Various developments in the field of technology help the blind person by developing camera-based and voice-based system combined with computer vision tools with the existing beneficial products such as an ultrasonic sensor, Google Text toSpeech tool, Optical Character Recognition System, etc. The main aim of this system is to act as a security guard and helps the blind to be aware of their surroundings and guardian can able to trace them by the help of GPS and IoT. The system will be an interactive one to the user byadding the Personal Assistant to this prototype. It is based on the latest technology Machine Learning, helps the user to communicate with the device and to know the attributes like time and date, text information, etc. An android based Chat- Bot helps the user to communicate with their friends, colleagues, friends, etc. by converting the speech to text and viceversa.

REFERENCES

[1] Rajesh M, Bindhu K. Rajan, Ajay Roy, Almaria Thomas K, Ancy Thomas, BincyTharakan T, Dinesh., “Text recognition and face detection aid for visually impaired person using raspberry pi” International Conference on circuits Power and Computing Technologies, 2017.

[2] Reshma Vijay Jawale, Madhavi Vijay Kadam, RavinaShantaramGaikawad, Lakshmi SudhaKondaka, “Ultrasonic Navigation based Blind Aid for The Visually Impaired”, IEEE International Conference on Power, Control, Signals and Instrumentation Engineering.

[3] Navya Amin and Markus Borschbach, “Quality of Obstacle Distance Measurement using Ultrasonic Sensor and Precision of Two Computer Vision-based Obstacle Detection Approaches”, International Conference on Smart Sensors and Systems,2015.

[4] Ms.Rupali,D Dharmale,Dr.P.V.Ingole,"TextDetectionand Recognition with Speech Output for Visually Challenged Person",vol. 5, Issue 1, January2016.

[5] Nagaraja, L., et al. "Vision based text recognition using raspberry PI." National Conference on Power Systems, Industrial Automation (NCPSIA2015).

[6]] Rajkumar N, Anand M.G, Barathiraja N, "Portable Camera Based Product Label Reading For Blind People.",IJETT, Vol. 10 Number 11 - Apr2014.

[7] Boris Epshtein, EyalOfek, Yonatan Wexler, "DetectingText in NaturalSceneswithStrokeWidthTransform.".

[8] Ezaki, Nobuo, et al. "Improved text-detection methods for a camera- based text reading system for visually impaired persons." Eighth International Conference on Document Analysis and Recognition (ICDAR’05). IEEE,2005.

[9] Ray Smith, “An Overview of the Tesseract OCREngine.".

[10] https://en.wikipedia.org/wiki/Optical_character_recognition.