SMART ASSISTANCE FOR BLIND PEOPLE USING RASPBERRY PI 3

A PROJECT REPORT

Submitted by

AJIT BABU.V **(211614106015)**

ANESTBALARUBASINGH.J.A **(211614106022)**

BATCHU ADITYA **(211614106038)**

DHANUSHKODI.G **(211614106049)**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

CHENNAI – 602105

ANNA UNIVERSITY: CHENNAI 600 025

APRIL 2018

SMART ASSISTANCE FOR BLIND PEOPLE USING RASPBERRY PI 3

A PROJECT REPORT

Submitted by

AJIT BABU.V **(211614106015)**

ANESTBALARUBASINGH.J.A **(211614106022)**

BATCHU ADITYA **(211614106038)**

DHANUSHKODI.G **(211614106049)**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

CHENNAI – 602105

ANNA UNIVERSITY: CHENNAI 600 025

APRIL 2018

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"SMART ASSISTANCE FOR BLIND PEOPLE USING RASPBERRY PI 3"** is the bonafide work of **"AJIT BABU.V (211614106015), ANEST BALA RUBASINGH.J.A (211614106022), BATCHU ADITYA (211614106038), DHANUSHKODI.G (211614106049)"** who carried out the project work under my supervision.

SIGNATURE

Dr.M.PALANIVELAN, M.E., Ph.D.

HEAD OF THE DEPARTMENT

Professor & Head,
Department of Electronics and
Communication Engineering,
Rajalakshmi Engineering College,
Thandalam, Chennai – 602 105.

SIGNATURE

Mrs.T.HELAN VIDHYA, M.E.

SUPERVISOR

Assistant Professor,
Department of Electronics and
Communication Engineering,
Rajalakshmi Engineering College,
Thandalam, Chennai – 602 105.

Submitted to Project and Viva-voce Examination held on _____

Internal Examiner **External Examiner**

ACKNOWLEDGEMENT

We thank Almighty God for enabling us to complete our project work. Our sincere thanks to our beloved chairperson **Dr. (Mrs.)Thangam Meganathan, M.A., M. Phil., Ph.D.**, for her enthusiastic motivation which is a lot in completing this project.

We like to thank our Principal **Dr.S.N.Murugesan, M.E., Ph.D.,** for his kind support and the facilities provided to carry out our work. We would like to thank **Dr.M.Palanivelan, M.E.,Ph.D.,** Head of the Department, Electronics and Communication Engineering, for giving the opportunity and facilities to complete our work in time.

We take this opportunity to express our profound gratitude and deep regards to our Supervisor **Mrs.T.Helan Vidhya, M.E.,** Assistant Professor, Department of Electronics and Communication Engineering, for her exemplary guidance, monitoring and constant encouragement throughout the course of this project.

We express our gratitude to the Project Coordinators **Mrs.B.Thilakavathi**, **Mrs.A.Asha, Dr.S.Chitra** and **Dr.K.Senthil Kumar** for their kind co-operation in doing our project work .We also thank our parents and classmates for their moral support and valuable suggestions in the project.

ABSTRACT

Object detection is technology of computer vision and image processing which deals with detection of instances of semantic objects of particular classes in digitalized images and videos. Outline of this paper is a system for detecting the type of obstacle; converting text into audio for reading purpose and detecting the kind of obstacle for blind people using Raspberry pi3.This technique helps the blind people to see the world virtually with an audio output. Optical Character Recognition (OCR) algorithm can be used to translate print version or hand-written text into audio output using Raspberry Pi. Ultrasonic sensors are attached with motor which detects the presence of obstacle in 180°. In this proposed model Raspberry pi3 is the medium which provides interface between Ultrasonic sensor and camera, which is implemented to provide the required output to the blind user that can be listened through headphones.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARM	Advanced RISC Machines
HDMI	High Definition Multimedia Interface
MICR	Magnetic Ink Character Recognition
OCR	Optical Recognition System
ROI	Region Of Interest
SSD	Solid State Drive
TTS	Text-to-Speech
USB	Universal Serial Bus
VGA	Video Graphics Array

CHAPTER 1

INTRODUCTION

Of the 314 million visually impaired people worldwide, 45 million are blind. Recent developments in computer vision, digital cameras and portable computers make it feasible to assist these individuals by developing camera based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems. Reading is obviously essential in today's society. Printed text is everywhere in the form of reports, receipts, bank statements, restaurant menus, classroom handouts, product packages, instructions on medicine bottles, etc.

The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and foster economic and social self-sufficiency. Today, there are already a few systems that have some promise for portable use, but they cannot handle product labeling. For example, portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products through speech and Braille.

1.1 PURPOSE

The main purpose of Optical Character Recognition (OCR) system based on a grid infrastructure is to perform Document Image Analysis, document processing of electronic document formats converted from paper formats more effectively and efficiently. This improves the accuracy of recognizing the characters during document processing compared to various existing available character recognition methods. Here OCR technique derives the meaning of the

characters, their font properties from their bit-mapped images. The primary objective is to speed up the process of character recognition in document processing. As a result the system can process huge number of documents with-in less time and hence saves the time. Since our character recognition is based on a grid infrastructure, it aims to recognize multiple heterogeneous characters that belong to different universal languages with different font properties and alignments.

1.2 PROJECT SCOPE

The scope of our product Optical Character Recognition on a grid infrastructure is to provide an efficient and enhanced software tool for the users to perform Document Image Analysis, document processing by reading and recognizing the characters in research, academic, governmental and business organizations that are having large pool of documented, scanned images. Irrespective of the size of documents and the type of characters in documents, the product is recognizing them, searching them and processing them faster according to the needs of the environment.

1.3 EXISTING SYSTEM

In existing approach it is a method to design a Text to Speech conversion module by the use of Mat lab by simple matrix operations. Initially by the use of microphone some similar sounding words are recorded using a record program in the Matlab window and recorded sounds are saved in ".wave" format in the directory. The recorded sounds are then sampled and the sampled values are taken and separated into their constituent phonetics. The separated syllables are then concatenated to reconstruct the desired words. By the use of various Mat lab commands i.e. wave read, subplot etc. the waves are sampled and extracted to get

the desired result. This method is simple to implement and involves much lesser use of memory spaces. We develop a framework for reconstructing images that are sparse in an appropriate transform domain from polychromatic Computed Tomography (CT) measurements under the blind scenario where the material of the inspected object and incident-energy spectrum are unknown.

The existing navigation systems for the blind people require a precise GPS maps. This make them unusable in region where there are no GPS maps, they are not sufficiently accurate. Algorithm for GPS navigation for the visually impaired along a GPS track, which describe the path as a sequence of waypoints is proposed. The natural voice navigation, adaptive to the velocity and accuracy of the GPS data, start of the navigation from any waypoint, correlation of the direction of movement if it is necessary, return the user to the route if deviation is deviated, work with and without electronic compass, detection of the movement of the user in the opposite direction.

Problem Statement:

Traditional methods like Braille exist using which the blind people have to trace and read text, which is very slow and not very practical. Existing OCR systems are not automatic and require full-fledged computers to run and hence are not effective. Reader Mobile runs on a cell phone and allows the user to read mail, receipts, fliers, and many other documents

Disadvantage:

- Braille system is very slow and not very practical.
- Existing OCR systems are not automatic.
- IR sensor is used.

1.4 PROPOSED SYSTEM

In proposed system we have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle.

In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision. By using ultrasonic sensor, we will measure the distance between the blind people and obstacle then the distance will be played through ear phones. Optical Character Recognition is the electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether form a scanned document, a photo of a document, a scene-photo or form subtitle text superimposed on an image. It is widely used as a form of information entry from printed paper. It is a method of digitizing printed texts so they can be electronically edited, searched, stored, more compact, displayed on-line and used in machine process such as cognitive computing, machine translation ,text-to-text, key data and text mining. OCR is a field of research in pattern recognition,

artificial intelligence and computer vision. Early versions needed to be trained with images of each character, and worked on one font at a time. Advanced systems capable of producing a high degree of recognition accuracy for most fonts are now common, and with support for a variety of digital image file format inputs. Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual component.

Advantage:

- It is inexpensive.
- Flexible for blind people .
- To convert printed books to digital text.
- Effective to handle.

CHAPTER 2

LITERATURE SURVEY

The paper titled **"An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking"** by E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri (2005) contains feasibility of an electromagnetic sensor to assist the autonomous walking of visually impaired and blind users is demonstrated in this paper. It is known that people affected by visual diseases usually walk assisted by some supports, among which the white cane is the most common. Our idea consists in applying a microwave radar on the traditional white cane making aware the user about the presence of an obstacle in a wider and safer range. Compared to the already existing Electronic Travel Aids devices, the proposed system exhibits better performance, noise tolerance and reduced dimensions. In the following, the latest developments of this research activity are presented, with special concern for the miniaturization of circuit board and antennas. A laboratory prototype has been designed and realized and the first test results of obstacle detection are hereby shown to demonstrate the effectiveness of the system.

The paper titled **"Text to Speech for the Visually Impaired"** by Mrs.Shilpa Reddy K, MounikaS.k,Pooja K , SahanaN (2016) states that perusing is fundamental in day by day life for everybody. Outwardly debilitated people can read just by utilization of unique applications by them like Braille dialect. The disadvantage of this framework is that each item does not give the content in Braille. In this paper, they have proposed an assistive content perusing system to help outwardly impeded people to peruse writings from different questions in their day by day lives. At first, we catch the picture of the required, pre-handling is performed on it. Pre-handling incorporates steps like dark scale and binarization,

question of intrigue acknowledgment. In the proposed framework, we are making the utilization of OTSU calculation to change over the dim scale picture into binarized one. The content districts from the caught picture are then separated and perceived by utilizing Optical Character Acknowledgment Programming (OCAP). The principle calculation in OCR to be specific MODI is utilized here. This extricated content of different textual styles and sizes then can be perceived independently and afterward consolidated in a word giving its yield as sound utilizing Text-to-discourse utilizing the SAPI libraries.

The paper titled **"A Reading aid for the Blind People using OCR and OpenCV"** by MallapaD.Gurav, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I. Byakod,ShivaleelaKanade (2017) presents that Optical Character Recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It coverts images of typed or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In this research these images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition, Artificial intelligence and computer vision. In this research, as the recognition process is done using OCR the character code in text files are processed using Raspberry pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document Image Analysis (DIA) we use information in grid format in virtual digital library's design and construction. This research mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry pi. Raspberry pi features a Broadcom System On a Chip (SOC) which includes ARM

compatible CPU and an on chip graphics processing unit GPU. It promotes Python programming as main programming language.

The paper titled **"Smart Stick for Blind People with Live Video Feed"** by SuchitaWankhade, MrunaliBichukale, Shruti Desai, ShraddhaKamthe, Archana Borate (2012) states that the ability and capability of vision to human being is an important factor of our life ,but some person whose unable and have lack of vision because they are visually impaired this paper we introduce one smart system which is nothing but smart devise which become helpful for that visually impaired people because of it that person can be detect obstacle with help of it blind stick, also that blind person get feel for when mobile phone get vibration alert or give some oral voice message to that person. Adaptability and flexibility for blind person can be define as ability to displace with conviction and welfare to his domain but it is not happened without science and education of technology we introduced one system which is make useful for blind people with the help of it those people can detect obstacle in front of them and prevent her/himself and familiar or unfamiliar person can also following and track that person from home also ,we get information related to current place of that person for that purpose we build one android application with the help of GPS (Graphic Positioning System).We can locate that blind person another features like grabbing live video feed (capturing video) ,voice Message all this features are build for smart device .

The main goal of that paper to introduce one smart system which become useful to the blind person by giving one smart stick, which capture obstacle in front of that person due to these new features. Hence we did one smart device for

visually impaired people which help them to move in there surrounding environment with mobility and confidence. GPS is use for find current location of person this project is implemented for blind people they can detect obstacle in local area network means only within range by buying IP address we can use that technologies any part of world.

The paper titled **"Object Detection Methodologies for Blind People"** by Miss. Kirti P. Bhure, Mrs. J. D. Dhande (2017) states that vision is the most important sense. Image plays vital role in the human perception of the surrounding environment. However there are visually impaired people, industry has created a variety of computer vision products and services by developing new electronic technologies for the blind in order to overcome the difficulties. Digital image processing is the field which processes the digital image by using digital computer. An increasing interest in developing technologies attempts to help visually impaired people in their daily lives. It is shown that the object identification is the difficult task for visually impaired people. Although there are many applications that can be used for this task, there are still limitations that require more improving. For this reason, this paper provides the survey and an analysis of various evaluations for the technologies that used in the object identification task. For the visually impaired the idea of sensory substitution can be used.

CHAPTER 3

PROPOSED SYSTEM

In proposed system we have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle. In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision. By using ultrasonic sensor, we will measure the distance between the blind people and obstacle then the distance will be played through ear phones. Optical Character Recognition is the electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether form a scanned document, a photo of a document, a scene-photo or form subtitle text superimposed on an image. It is widely used as a form of information entry from printed paper. It is a method of digitizing printed texts so they can be electronically edited, searched ,stored, more compact, displayed on-line and used in machine process such as cognitive computing, machine translation ,text-to-text, key data and text mining. OCR is a field of research in pattern recognition,

artificial intelligence and computer vision. Early versions needed to be trained with images of each character, and worked on one font at a time. Advanced systems capable of producing a high degree of recognition accuracy for most fonts are now common, and with support for a variety of digital image file format inputs. Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual component.

3.1 BLOCK DIAGRAM

We have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person.

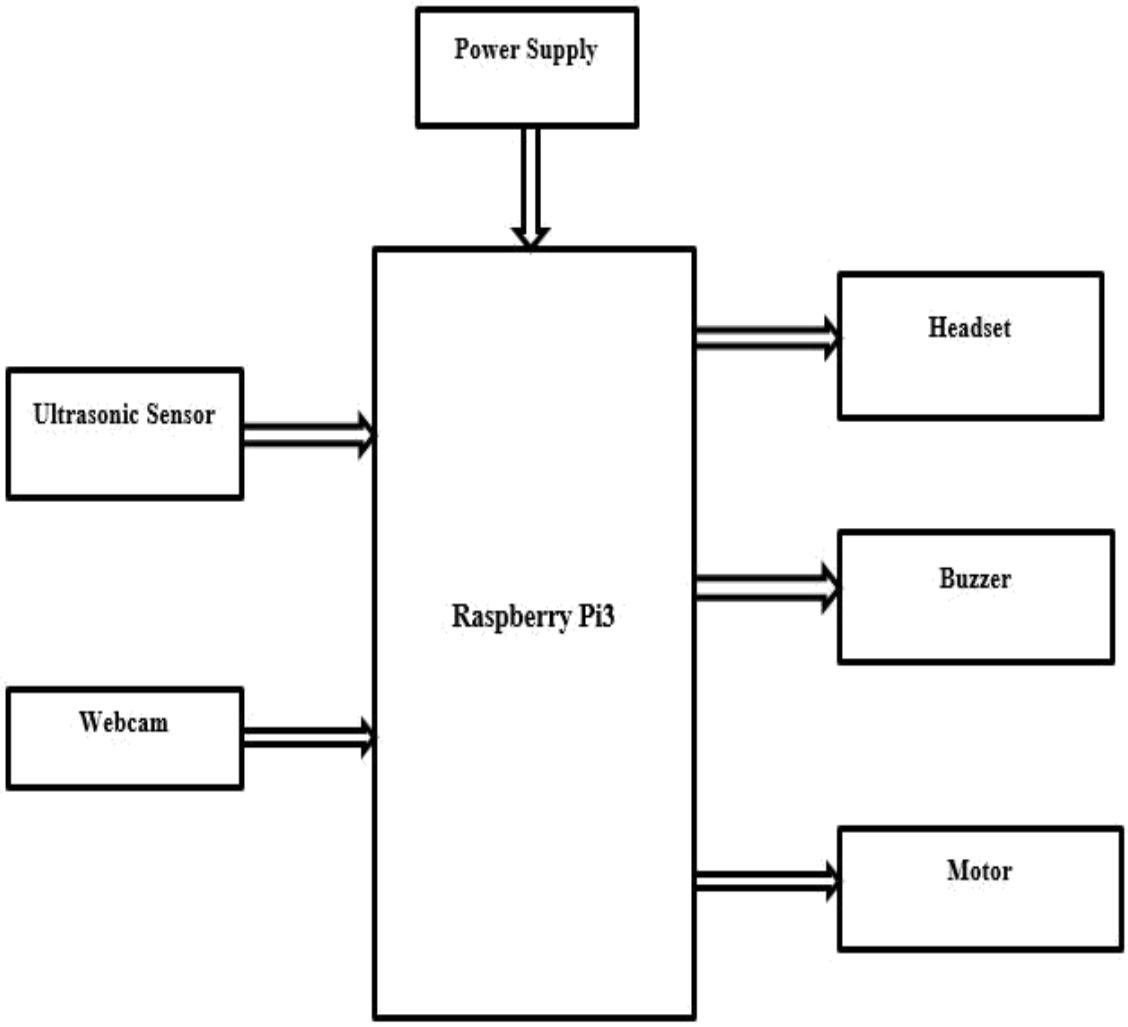

Figure 3.1 Block Diagram

Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle.

In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-

assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision

When our application starts running it first check all the devices and resources which it needs are available or not. After that it checks the connection with the devices and gives control to the user. The GUI for the user has the following options. An optional label is used for displaying the image taken from the camera. A status box is for representing the detected data from the image. The capture button is to detect the data from the image. The detect button is to detect the human from the video streaming in front of the camera. The audio jack port is the output port here. The Raspberry board comes with integrated peripherals like USB, ADC and Serial etc. On this board we are installing Linux operating system with necessary drivers for all peripheral devices.

3.2 SYSTEM ARCHITECTURE

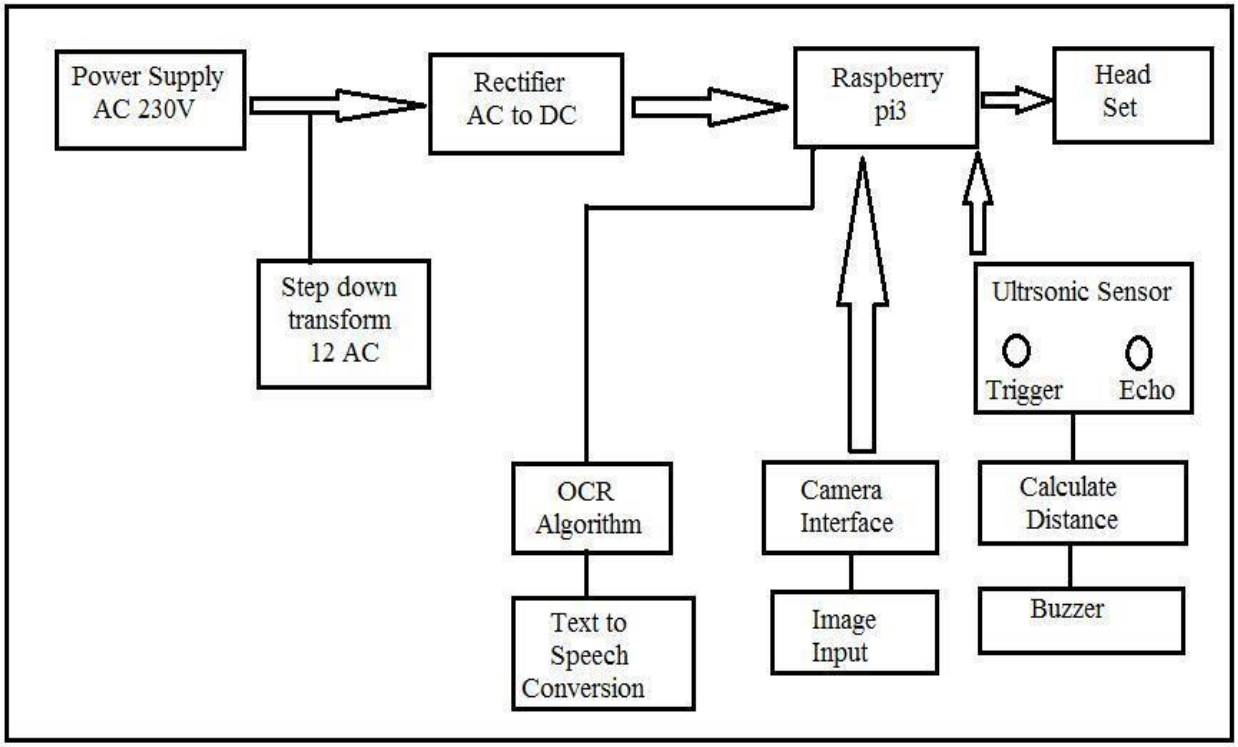

Figure 3.2 System Architecture

We have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle.

In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision.

When our application starts running it first check all the devices and resources which it needs are available or not. After that it checks the connection with the devices and gives control to the user. The GUI for the user has the following options. An optional label is used for displaying the image taken from the camera. A status box is for representing the detected data from the image. The capture button is to detect the data from the image. The detect button is to detect the human from the video streaming in front of the camera. The audio jack port is the output port here. The Raspberry board comes with integrated peripherals like

USB, ADC and Serial etc. On this board we are installing Linux operating system with necessary drivers for all peripheral devices.

Use Case Diagram:

When capture button is clicked this system captures the product image placed in front of the web camera which is connected to ARM microcontroller through USB .After selecting the process button the captured label image undergoes Optical Character Recognition(OCR) Technology. OCR technology allows the conversion of scanned images of printed text or symbols (such as a page from a book) into text or information that can be understood or edited using a computer program.

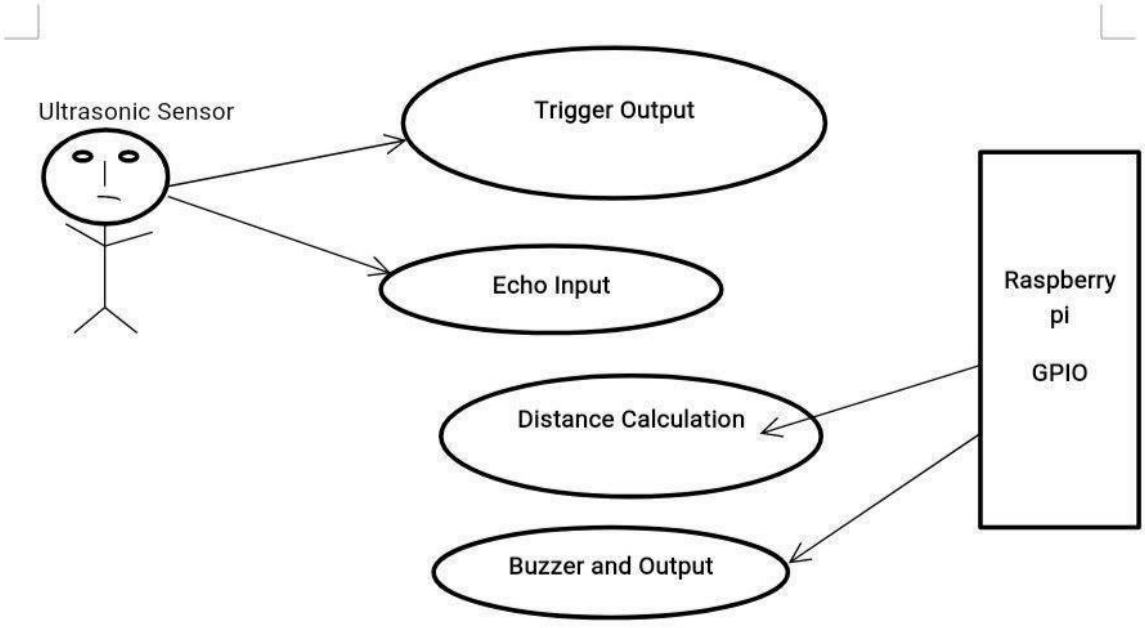

Figure 3.3 Case Diagram for Ultrasonic Sensor

The most familiar example is the ability to scan a paper document into a computer where it can then be edited in popular word processors such as Microsoft Word. However, there are many other uses for OCR technology, including as a

component of larger systems which require recognition capability, such as the number plate recognition systems, or as tools involved in creating resources for SALT development from print based texts. In our system for OCR technology we are using TESSERACT library. Using Flite library the data will be converted to audio.

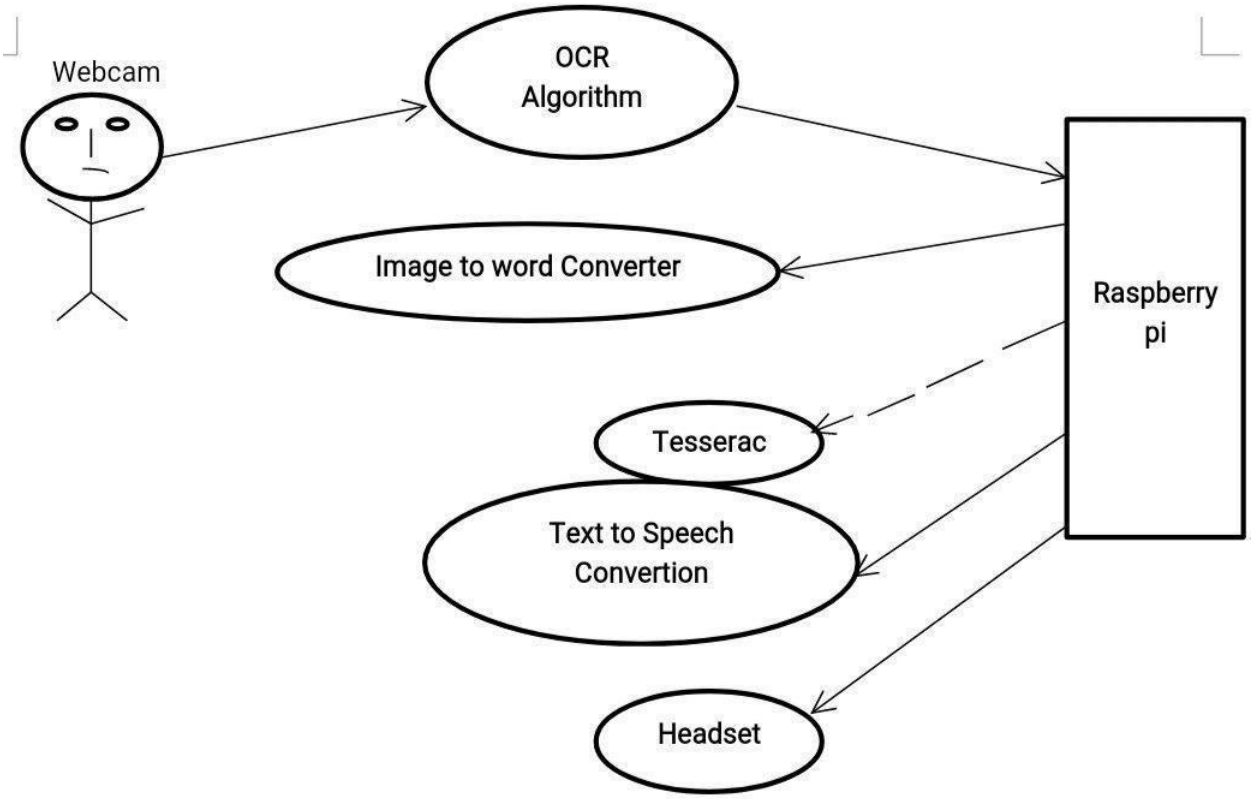

Figure 3.4 Case Diagram for OCR Algorithm

Camera acts as main vision in detecting the label image of the product or board then image is processed internally and separates label from image by using open CV library and finally identifies the product and identified product name is pronounced through voice. Now it identifies received label image is converted to text by using tesseract library. Once the identified label name is converted to text and converted text is displayed on display unit connected to controller. Now

converted text should be converted to voice to hear label name as voice through ear phones connected to audio jack port using flite library.

Modules:

- Image capturing and pre-processing.
- Automatic text extraction.
- Text recognition and audio output.

3.3 IMAGE CAPTURING AND PRE-PROCESSING

The video is captured by using web-cam and the frames from the video is segregated and undergone to the pre-processing. First, get the objects continuously from the camera and adapted to process. Once the object of interest is extracted from the camera image and it converted into gray image. Use haar cascade classifier for recognizing the character from the object. The work with a cascade classifier includes two major stages: training and detection. For training need a set of samples. There are two types of samples: positive and negative.

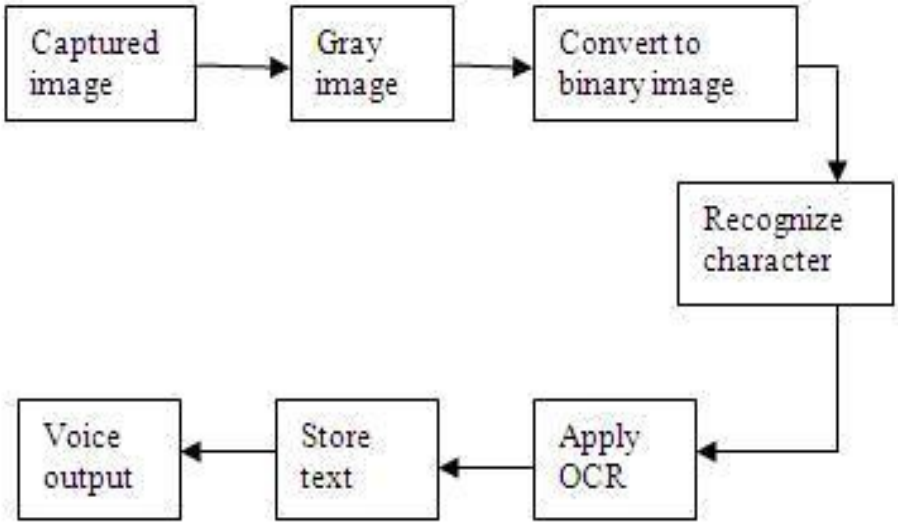

Figure 3.5 Image Capture and Pre-processing

To extract the hand-held object of interest from other objects in the camera view, ask users to shake the hand-held objects containing the text they wish to identify and then employ a motion-based method to localize objects from cluttered background.

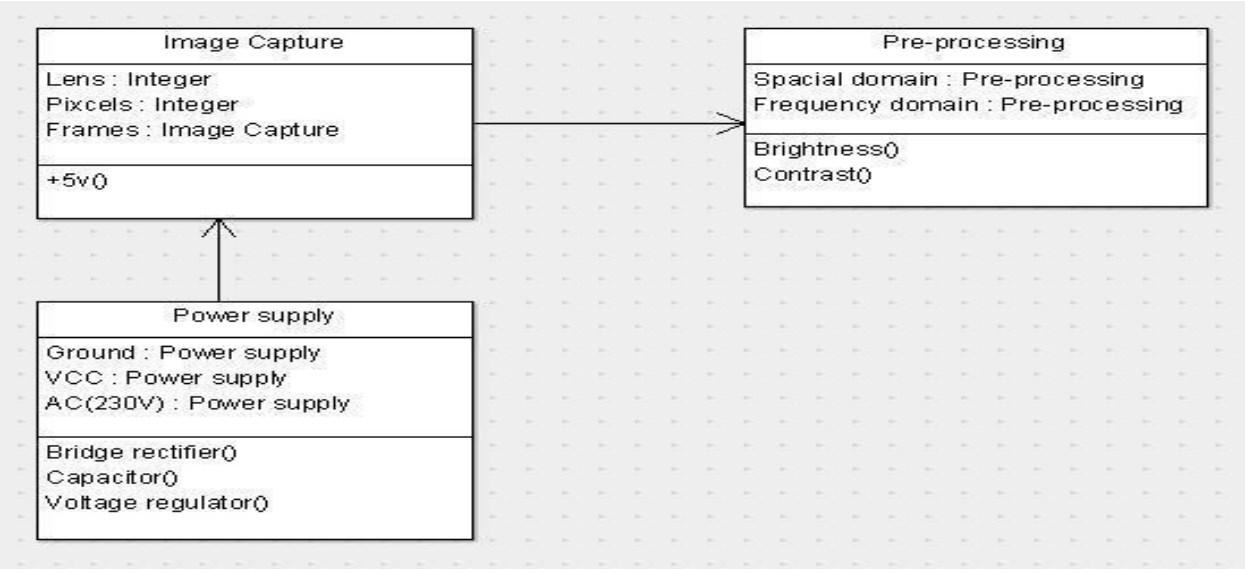

Figure 3.6 Class Diagram for Image Capturing and Pre-Processing Module

In this module the image is capture by the web camera based on using lens, pixels and frames by +5v.The frame from the image is segregated and it undergoes to pre-processor with the spatial domain and frequency domain by brightness(),contrast().Power supply is also used to capture image with the help VCC,AC(230) with function of bridge rectifier(),voltage regulator()**.**

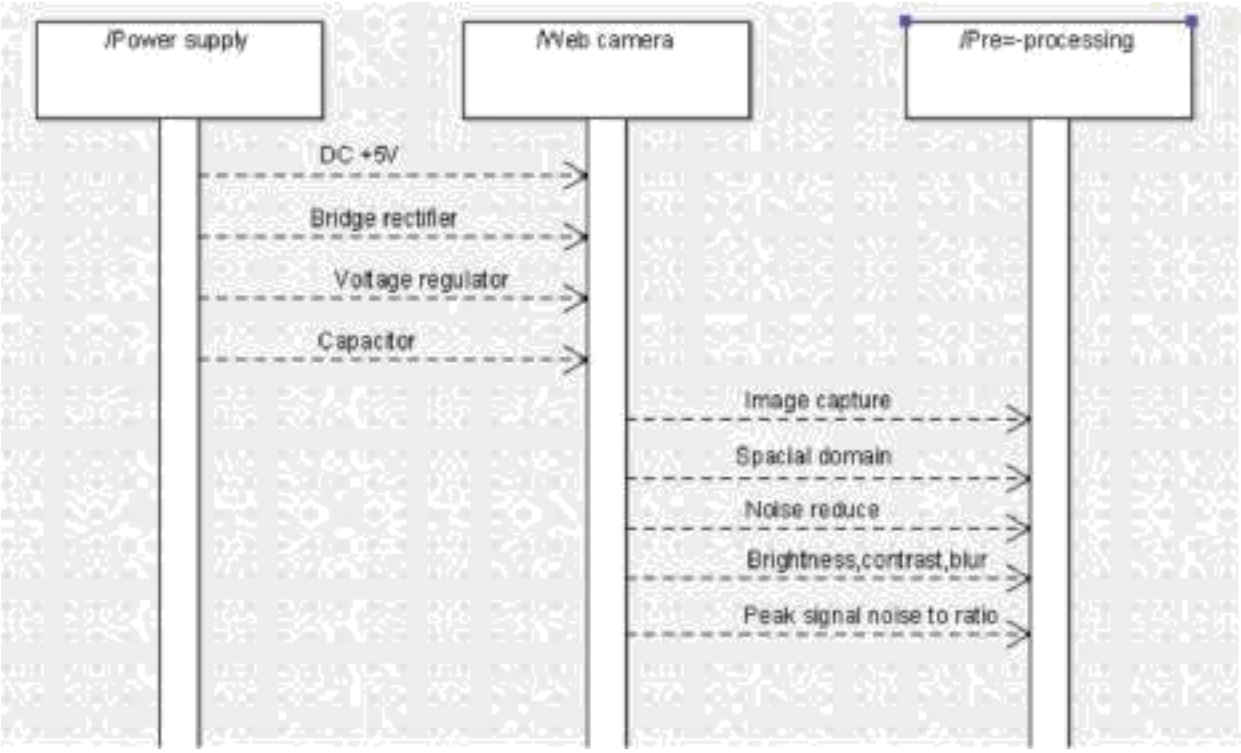

Figure 3.7 Sequence Diagram for Image Capturing and Pre-Processing Module

In this module, the power supply flows direct current to the webcam by +5v, the bridge rectifier are used convert alternating current input to direct current output to the web camera, voltage are used to regulate one or more AC or DC voltages. Webcam are used to capture the segregated image from pre-processing. If the image captured by webcam has brightness, contrast, blur are stored by pre-processor .It also reduce the noise captured by webcam. Therefore it is a PNR. A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. A sequence diagram shows how object interaction arranged in time sequence.

It depicts the object and classes involved in the scenario and the sequence of message exchanged between the object needed to carry out the functionality of scenario.

In this module, the power supply uses +5vdirect current to the web camera and pixels value are segregated to the pre-processing that are pass through the special domain through domain it transform to frequency domain and checks the brightness, contrast, blur and reduce the noise through PSNR and MSC.

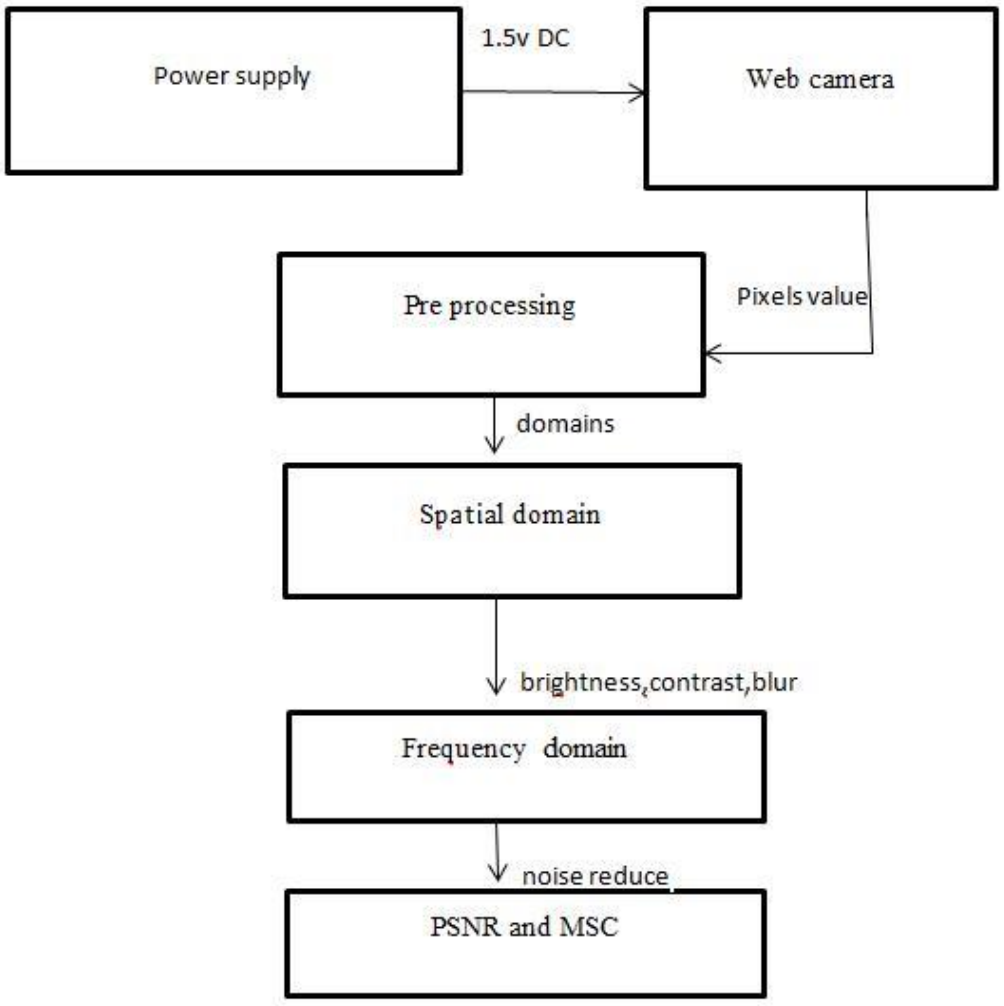

Figure 3.8 Collaboration Diagram for Image Capturing and Pre-Processing Module

3.4 AUTOMATIC TEXT EXTRACTION

In order to handle complex backgrounds, two novel feature maps to extracts text features based on stroke orientations and edge distributions, respectively. Here, stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Adaboost based text classifier. The extraction information from audio and image source restricted to information execution from text. The actual transduction of audio and image data into text is the processing of OCR output.

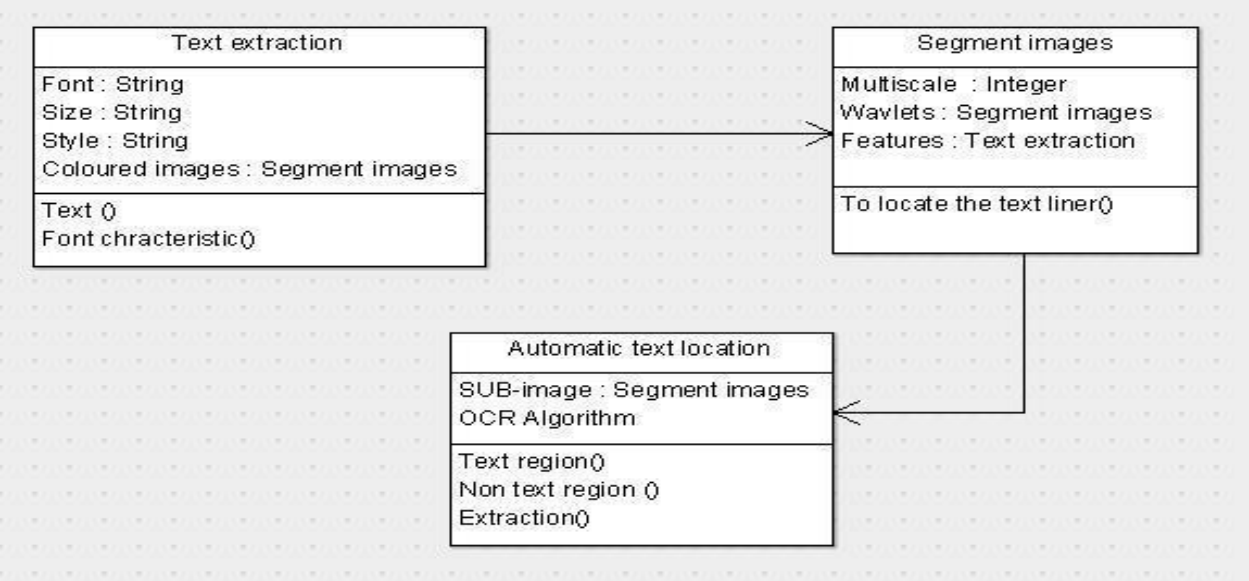

Figure 3.9 Class Diagram for Automatic Text Extraction Module

In this module, the text extraction are extracted based on font, size, style, coloured images by using text(),font characteristic().The extracted text are segmented into multi-scale, wavelets, features are used to locate the text. The automatic text location are 21ecognized by SUB-image ,OCR algorithm based on the text region(),non- text region(),extraction().

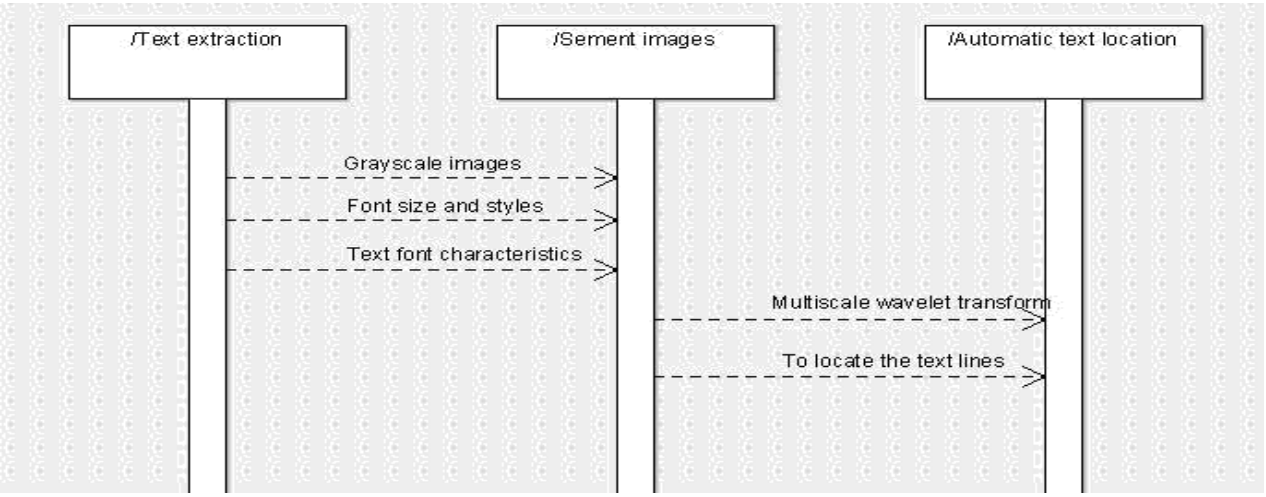

Figure 3.10 Sequence Diagram for Automatic Text Extraction Module

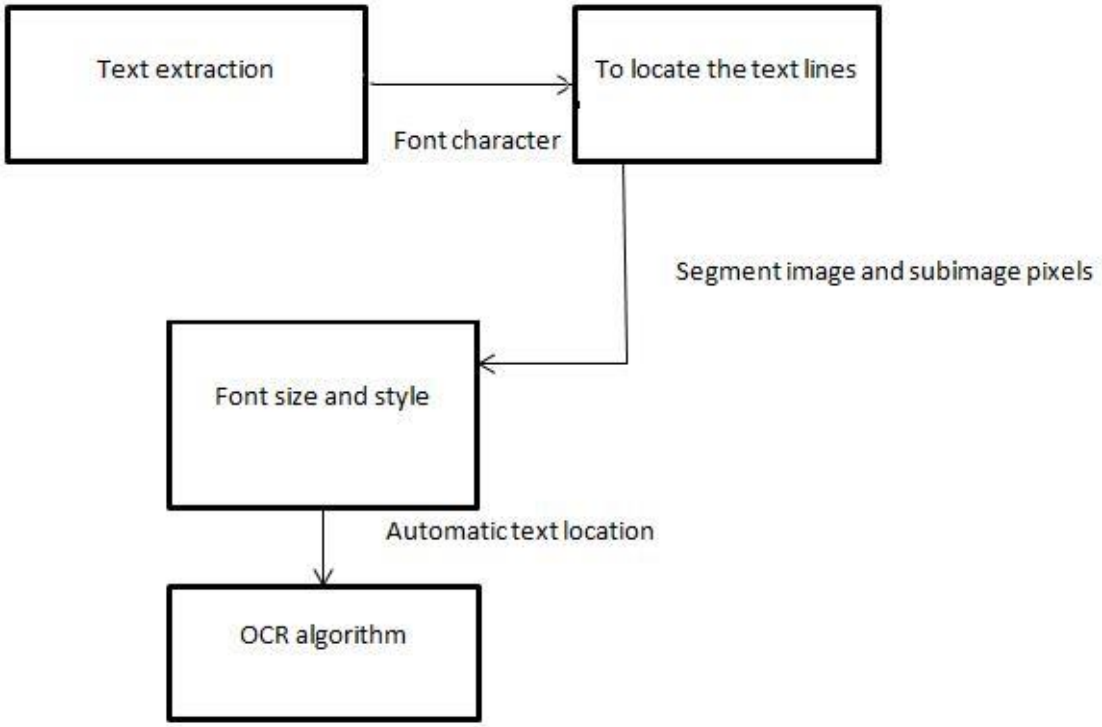

Figure 3.11 Collaboration Diagram for Automatic Text Extraction Module

Text extraction are used to convert grey scale image into segment images by using font size, styles and text font characteristic .The segmented images are used to find the location of text are in the form of multistate wavelet transform. The segmented image are used to locate the text lines through the text region(),non-text region(),extraction().

In this module, the text extraction are used to extract the text from the image by font character, text region and they are used to locate the text lines. The text lines are represented by front size and style through segmented image and sub pixels image. The font size and style are identified by OCR algorithm though automatic text location.

3.5 TEXT RECOGNITION AND AUDIO OUTPUT

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text regions. A text region labels the minimum rectangular area for the accommodation of characters inside it, so the border of the text region contacts the edge boundary of the text characters. However, this experiment show that OCR generates better performance text regions are first assigned proper margin areas and binaries to segments text characters from background. The recognized text codes are recorded in script files. Then, employ the Microsoft Speech Software Development Kit to load these files and display the audio output.

Blind users can adjust speech rate, volume and tone according to their preferences. Static random-access memory (SRAM) is a type of a semiconductor memory that uses bi-stable latching circuitry to store each bit.

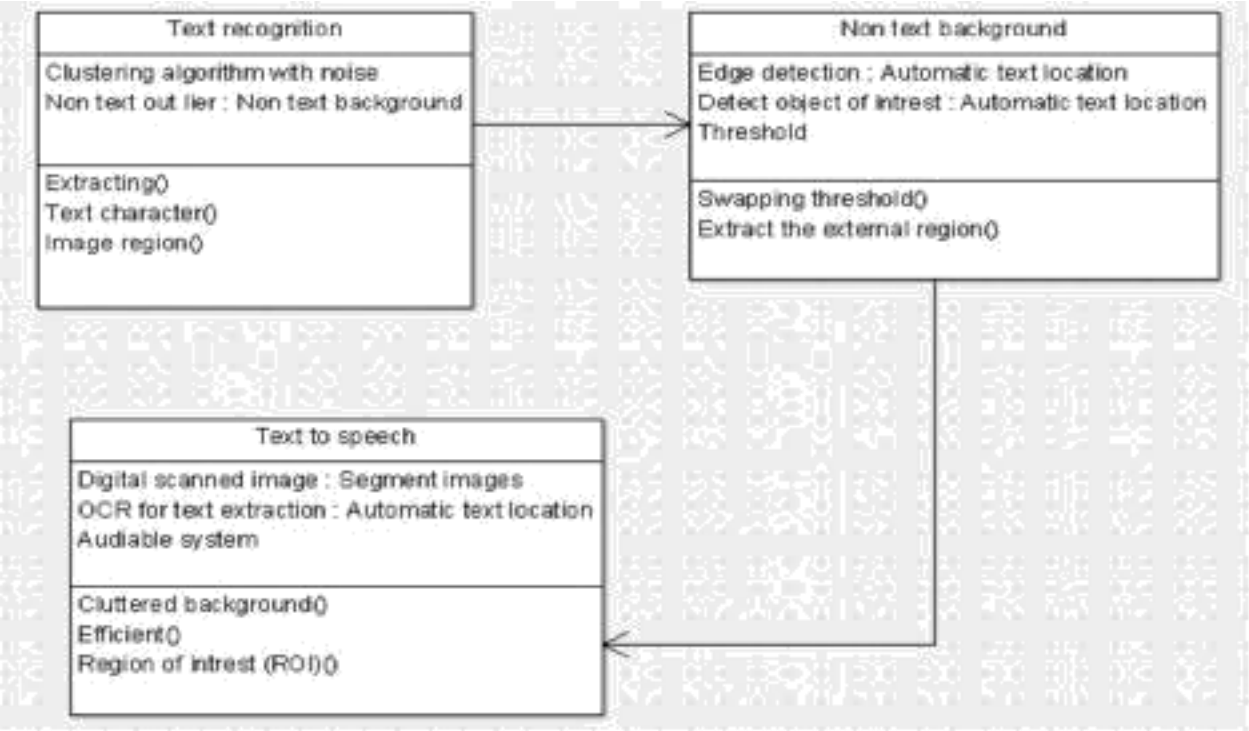

Figure 3.12 Class Diagram for Text Recognition and Audio Output Module

In this module, the text recognition are used to recognize the text with the help of clustering algorithm non noise text outlier by text character().The non-text background are used to detect the object, threshold value and they are used to convert text to speech conversion module by using OCR algorithm to recognize the character of the text. Class diagram is a type of static structure of a system by showing the system classes, their attribute and the relationship between the classes. Private visibility hides information from anything outside the class partition. Public visibility allows all other to view the marked information. It is used for general conceptual modeling of the systematic of the application, and for detailed modeling translating the model into programming code.

Class diagram can also used for data modeling. The classes in the class diagram represent both the main elements, interaction in the application, and the classes to be programmed.

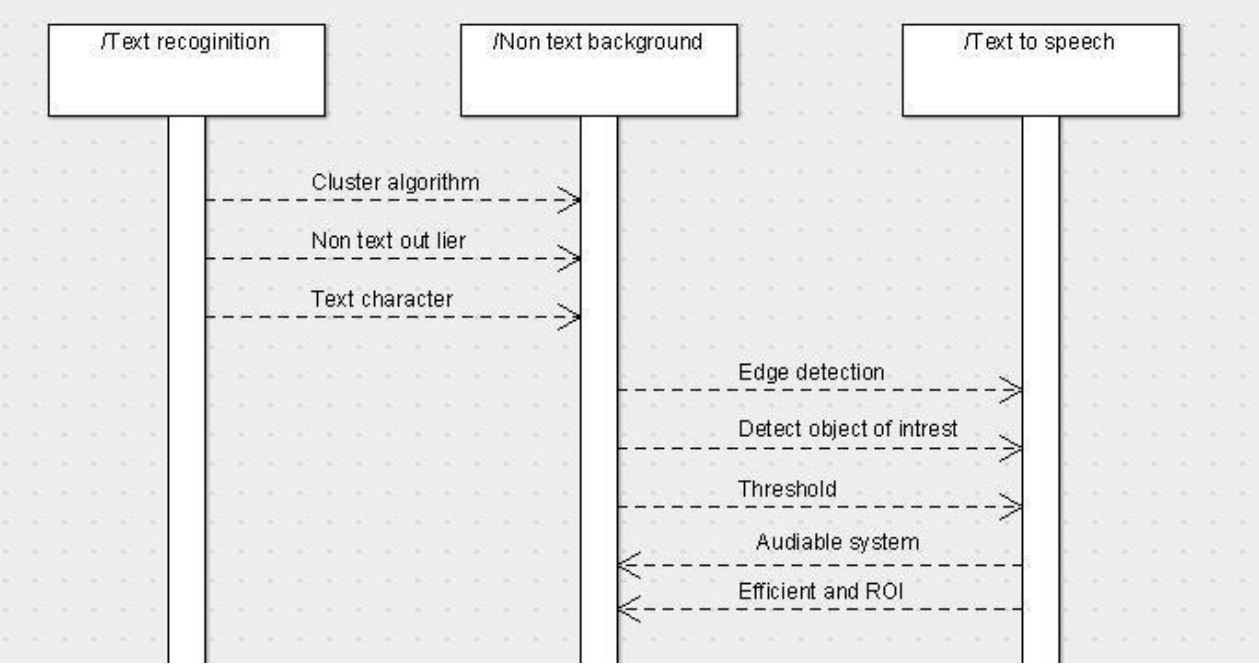

Figure 3.13 Sequence Diagram for Text Recognition and Audio Output Module

In this module, the text recognition are used to recognize the text by cluster algorithm with the non-text outlier to identify the text character for recognition. The non-text background are used to convert text to speech to detect the object, detection of edges, threshold values, the conversion are audible and efficient to the non- text background.

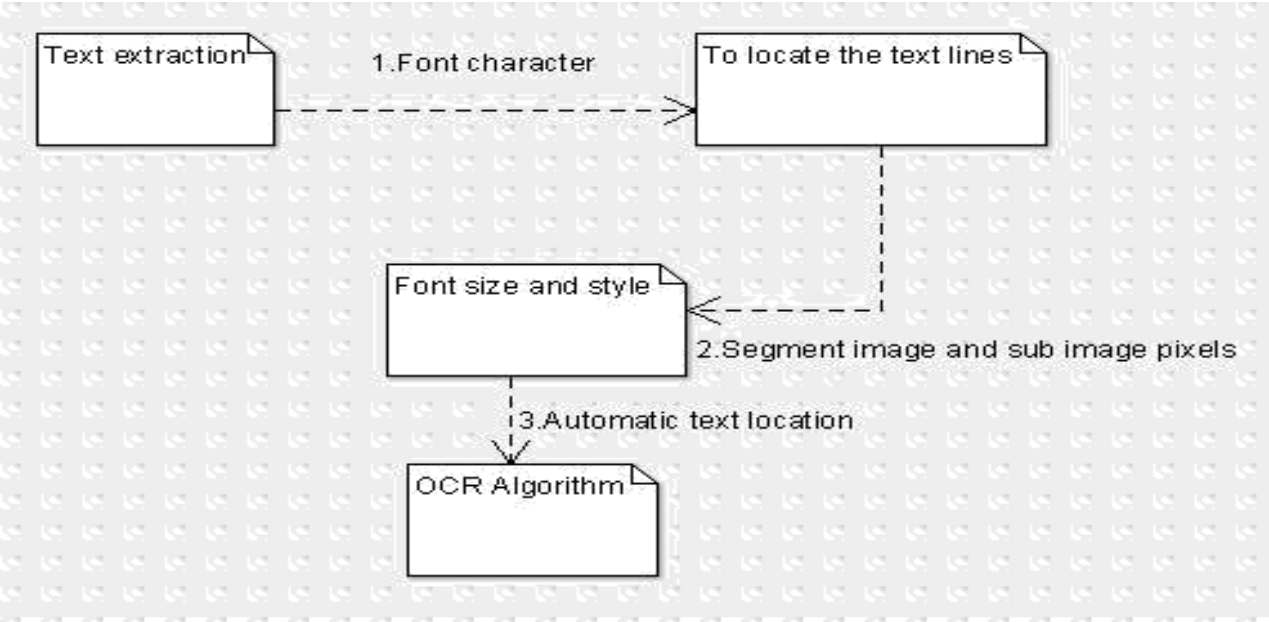

Figure 3.14 Collaboration Diagram for Text Recognition and Audio Output Module

In this module, the text recognition are used to process through non text background and edge detection by non- text outlier. The non-textile are used to convert text to speech by threshold range. The conversion module are very efficient by digital scanned image. The images are passes through of Region Of Interest(ROI).

CHAPTER 4

HARDWARE DESCRIPTION

The hardware components used in the proposed system are listed below

- Raspberry Pi 3
- HDMI to VGA converter
- Ultrasonic sensor
- Webcam
- Speakers / Headphones
- SD Card
- Monitor,Keyboard,Mouse
- Power cable

4.1 RASPBERRY PI

Raspberry Pi seems to be new in the world and many people really don't know what the Raspberry Pi is. Raspberry Pi can be defined as a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It's capable of doing everything you would expect a desktop computer to do, from browsing the internet and playing high definition video, to making spreadsheets, word-processing, and playing games. It is great bonding with Arduino and can do a lot with Arduino.

There are two models of Raspberry Pi, model A and model B. These two are bit similar with few advance features on model B compared to model A. Model B has 512 MB RAM, two USB port where as Model A has 256 MB

RAM and just a USB port. Besides, Model B has Ethernet port while Model A does not. The overview of the Raspberry Pi Model B is shown in Figure 5.1.

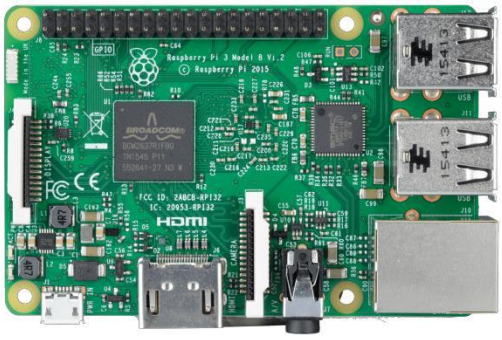

Figure 4.1 Raspberry Pi 3 Model B

Different components of the Pi are named in the Figure 5.1 and brief description on each component is given in following sections.

SD Card Slot:

Raspberry Pi doesn't have the real hard drive as in laptop and computer, SD card is taken as solid state drive (SSD) which is used to install operating system and all others software and store everything. This card is needed to insert into the slot for using the Raspberry Pi. SD card may be 2GB, 4GB or 16GB.

Micro USB Power:

The power port is a 5V micro-USB input and supply should be exactly 5v as it doesn't have onboard power regulator. So, power supply shouldn't exceed than 5V.

HDMI Out:

This output port is used to connect the Raspberry Pi with a monitor via HDMI (High Definition Multimedia Interface). Hence, any screen or TV can be connected to it which consists of HDMI port.

Ethernet and USB port:

Both the Ethernet port and USB port on Model B are supplied via the onboard LAN9512 chip. It is a high-speed USB 2.0 hub with a 10/100 Ethernet controller (Donat, 2014). USB ports are used to connect the inputs (keyboard, mouse). Almost everything that can connect to computer via USB also can connect with Raspberry Pi.

RCA Video Out and Audio Out:

Audio and RCA video jacks are present on the board for audio and video out. The Raspberry Pi does support sound over its HDMI output, but there is a standard 3.5-mm audio jack to plug in headphones but USM mikes may work or not. For video, the RCA jack sends video to any connected RCA video device.

GPIO Headers(Pins):

GPIO pins stands for general purpose of input output pins. These pins are used to connect any number of physical extensions with the Raspberry Pi. Raspberry Pi has pre-installed libraries that allow us to access the pins using programming languages like C, C++ or python.

Chips (Broadcom):

The most important component in a Raspberry Pi is chip that is Broadcom which is placed at the middle of the board. The chip consists of ARM11 processor running at 700 MHz and a Videocore4 GPU and can be over clocked to at least 800 MHz without a problem.

4.2 RASPBERRY PI 3 MODEL B

Recently, Raspberry Pi 3 Model B has been lunched recently which Broadcom BCM2836 ARM Cortex-A7 Quad Core Processor has powered Single Board

Computer running at 900MHz, 1GB RAM and 4 Quad USB ports. It is the advanced version of Model B and is 6 times faster than Model B Raspberry Pi. In addition, it has combined 4-pole jack for connecting your stereo audio out and composite video out and advanced power management. Figure 5.1 shows the top view of the board with labels of some important components (raspberrypi.org, 2015a).

4.3 HARDWARE REQUIRED FOR RASPBERRY PI

Raspberry Pi can't start alone, it needs many others peripherals (hardware). There is brief description of the hardware requirements in the following section (Bates, 2014).

Power Supply:

As mentioned already in above theory portion, Raspberry Pi needs 5V power supply. If supply exceeds 5V then it can't guaranteed to work properly. And the power supply also need to supply at least 500 milliamps (Ma), and preferably more like 1 amp (A). If the supply is 500 Ma or less, it is likely to have the mal-function of keyboard and mouse. It is not good idea to power the Raspberry Pi from USB port of computer and hub as they mostly provide current less than required. Hence, the Raspberry Pi requires a Micro-USB connection which is capable of supplying at least 700 Ma (or 0.7 A) at 5V.

Storage:

A separate hardware is required for the storage purpose in Raspberry Pi. For this, SD card is used, mostly 4 GB and 8 GB if needed. The operating system and all files are stored in the card. We can buy blank SD card and install operating system or buy a pre-installed one.

Input:

External keyboard and mouse are required to provide input to the Raspberry Pi. No any additional software is needed to use the keyboard and mouse.

Monitor:

We can use monitor or TV with HDMI port or DVI inputs as the screen for the Raspberry Pi. For DVI inputs, HDMI-to-DVI converters are required which can be finding easily in a market. Monitor is most important for the Raspberry Pi as it is the only way to see what we have done on it.

Network:

As in laptop or computer, we can access to internet and network in Raspberry Pi as well. For that, we can use wired Ethernet connection which is easier option or Wi-Fi module to access Wi-Fi in the Raspberry Pi.

4.4 POWER SUPPLY UNIT

A power supply (sometimes known as a power supply unit or PSU) is a device or system that supplies electrical or other types of energy to an output load or group of loads. The term is most commonly applied to electrical energy supplies, less often to mechanical ones, and rarely to others.

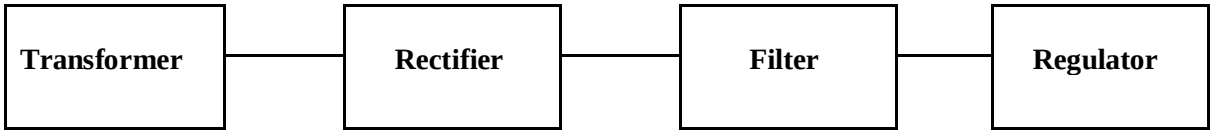

Figure 4.2 Block diagram of power supply

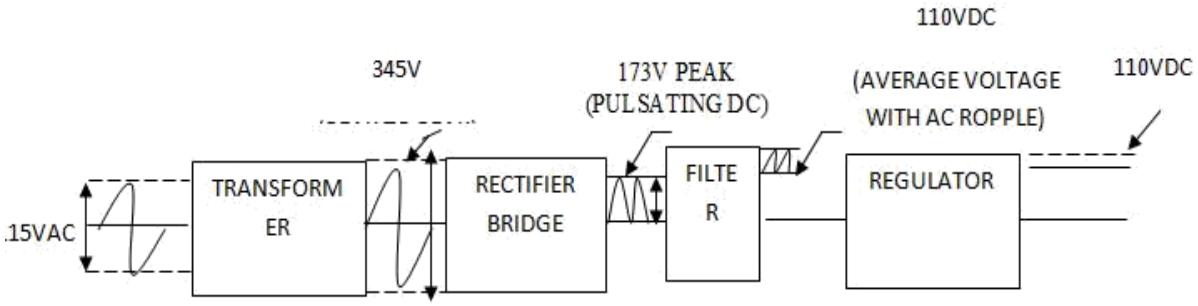

Figure 4.3 Processing of power supply

The transformer steps up or steps down the input line voltage and isolates the power supply from the power line. The RECTIFIER section converts the alternating current input signal to a pulsating direct current. However, as you proceed in this chapter you will learn that pulsating dc is not desirable. For this reason a FILTER section is used to convert pulsating dc to a purer, more desirable form of dc voltage.

The final section, the REGULATOR, does just what the name implies. It maintains the output of the power supply at a constant level in spite of large changes in load current or input line voltages. Now that you know what each section does, let's trace an ac signal through the power supply. At this point you need to see how this signal is altered within each section of the power supply. Later on in the chapter you will see how these changes take place. An input signal of 115 volts ac is applied to the primary of the transformer. The transformer is a step-up transformer with a turns ratio of 1:3. You can calculate the output for this transformer by multiplying the input voltage by the ratio of turns in the primary to the ratio of turns in the secondary; therefore, 115 volts ac ´ 3 = 345 volts ac (peak-to- peak) at the output. Because each diode in the rectifier section conducts for 180 degrees of the 360-degree input, the output of the rectifier will be one-half, or approximately 173 volts of pulsating dc. The filter section, a network of resistors, capacitors, or inductors, controls the rise and fall time of the varying signal; consequently, the signal remains at a more

constant dc level. You will see the filter process more clearly in the discussion of the actual filter circuits. The output of the filter is a signal of 110 volts dc, with ac ripple riding on the dc. The reason for the lower voltage (average voltage) will be explained later in this chapter. The regulator maintains its output at a constant 110-volt dc level, which is used by the electronic equipment (more commonly called the load).

Simple 5V power supply for digital circuits:

- Brief description of operation: Gives out well-regulated +5V output, output current capability of 100 Ma
- Circuit protection: Built-in overheating protection shuts down output when regulator IC gets too hot
- Circuit complexity: Very simple and easy to build
- Circuit performance: Very stable +5V output voltage, reliable operation
- Availability of components: Easy to get, uses only very common basic components
- Design testing: Based on datasheet example circuit, I have used this circuit successfully as part of many electronics projects
- Applications: Part of electronics devices, small laboratory power supply
- Power supply voltage: Unregulated DC 8-18V power supply
- Power supply current: Needed output current + 5 Ma
- Component costs: Few dollars for the electronics components + the input transformer cost

4.5 WEBCAM

A webcam is a video camera which feeds its images in real time to a computer or computer network, often via USB, Ethernet or Wi-Fi. Their most

popular use is the establishment of video links, permitting computers to act as videophones or videoconference stations. This common use as a video camera for the World Wide Web gave the webcam its name. Other popular uses include security surveillance and computer vision. Webcams are known for their low manufacturing cost and flexibility, making them the lowest cost form of video telephony. They have also become a source of security and privacy issues, as some built-in webcams can be remotely activated via spyware.

Lens:

Webcams typically include a lens, an image sensor, support electronics, and may also include a microphone for sound. Various lenses are available, the most common in consumer-grade webcams being a plastic lens that can be screwed in and out to focus the camera. Fixed focus lenses, which have no provision for adjustment, are also available. As a camera system's depth of field is greater for small image formats and is greater for lenses with a large f-number (small aperture), the systems used in webcams have a sufficiently large depth of field that the use of a fixed focus lens does not impact image sharpness to a great extent.

Image sensors can be CMOS or CCD, the former being dominant for low-cost cameras, but CCD cameras do not necessarily outperform CMOS-based cameras in the low cost price range. Most consumer webcams are capable of providing VGA resolution video at a frame rate of 30 frames per second. Many newer devices can produce video in multi-megapixel resolutions, and a few can run at high frame rates such as the PlayStation Eye, which can produce 320×240 video at 120 frames per second.

Support electronics read the image from the sensor and transmit it to the host computer. The camera pictured to the right, for example, uses a Sonix SN9C101 to transmit its image over USB. Typically, each frame is transmitted

uncompressed in RGB or YUV or compressed as JPEG. Some cameras, such as mobile phone cameras, use a CMOS sensor with supporting electronics “on die”, i.e. the sensor and the support electronics are built on a single silicochip to save space and manufacturing costs.

Portability:

Most webcams feature built-in microphones to make video calling and videoconferencing more convenient. The USB video device class (UVC) specification allows for interconnectivity of webcams to computers without the need for proprietary device drivers. Microsoft Windows XP SP2, Linux and Mac OS X (since October 2005) have UVC support built in and do not require extra device drivers, although they are often installed to add additional features.

Features(LOGITECH WEBCAM C100):

- Plug-and-play setup (UVC)
- Video capture: Up to 640 x 480 pixels
- Photos: Up to 1.3 megapixels (software enhanced)
- Frame rate: Up to 30 frames per second (with recommended system)
- Universal clip fits notebooks, LCD or CRT monitor

Figure 4.4 Webcam

CHAPTER 5

SOFTWARE DESCRIPTION

5.1 START UP WITH RASPBERRY PI

Raspberry Pi is a small computer; hence operating system (OS) should be installed. As the Raspberry doesn't have hard drive, OS is installed in the external memory. For that, memory card (SD card) is used for the installation of operating system and all the required software and supporting files are stored in the same SD card.

There are different types of operating system but we preferred to talk about NOOBS (New Out of the Box Software) as it is suitable for the beginners. We can either buy a preinstalled SD card or empty SD card. In pre-installed SD card, NOOBS is already copied and ready to boot.

If you don't have suitable screen for the Raspberry Pi, we can use the monitor of a laptop. Besides, we can use keyboard and mouse of laptop as well for Raspberry Pi. For that, two software are needed to install in a laptop which screen is going to use for Raspberry Pi. Installation and configuration of the software is discussed briefly in the following sections. Xming

This is the first software need to be installed which can be downloaded from the link, Download HYPERLINK "http://sourceforge.net/project/downloading.php?group_id=156984&filename=Xming-6-9-0-31-setup.exe" HYPERLINK "http://sourceforge.net/project/downloading.php?group_id=156984&filename=Xming-6-9-0-31-setup.exe"Xming HYPERLINK "http://sourceforge.net/project/downloading.php?group_id=156984&filename=Xming-6-9-0-31-setup.exe" and install it in the laptop. After completion of

Installation, run the application called ‘Xlaunch’ and verify that the settings are as shown in three following figures; Figure 5.1, Figure 5.2 and Figure 5.3.

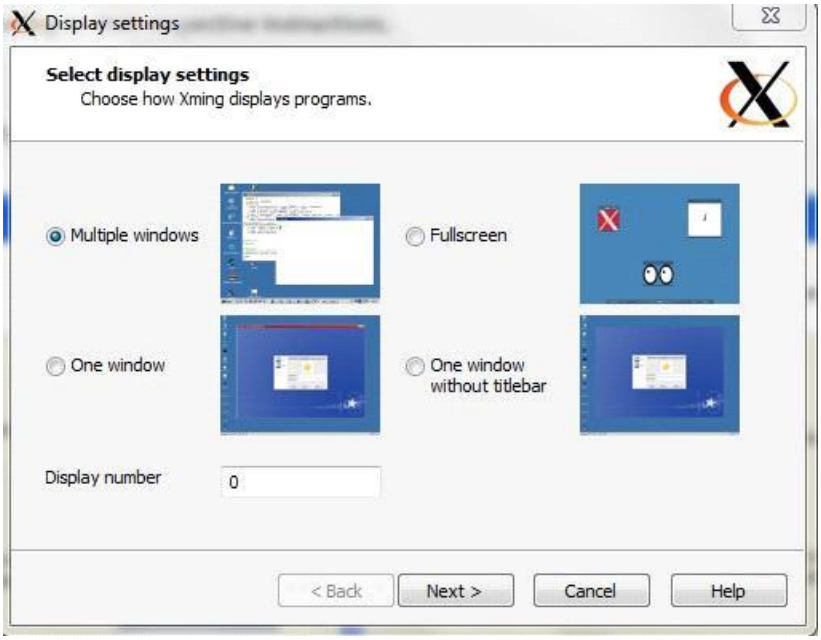

Figure 5.1 Xming Configuration step 1

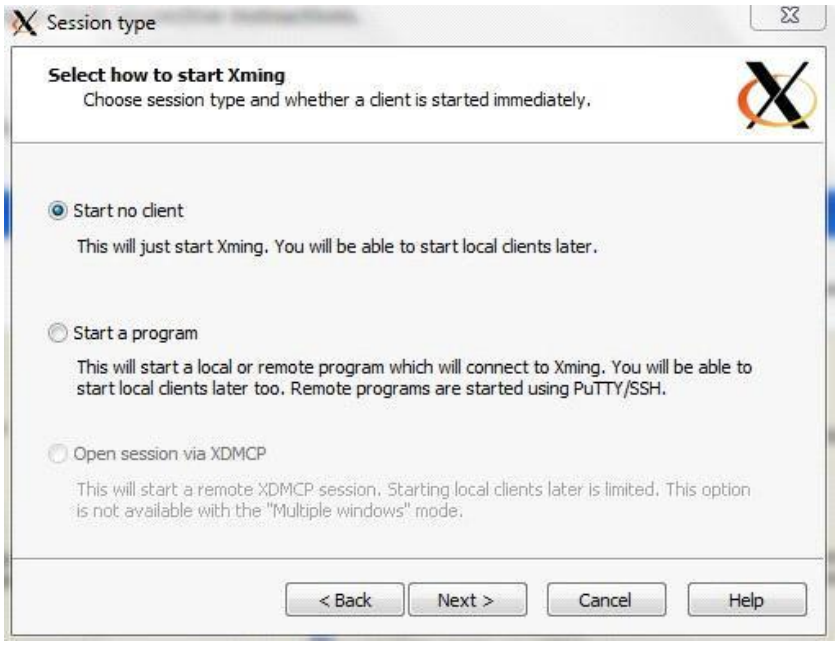

Figure 5.2 Xming Configuration step 2

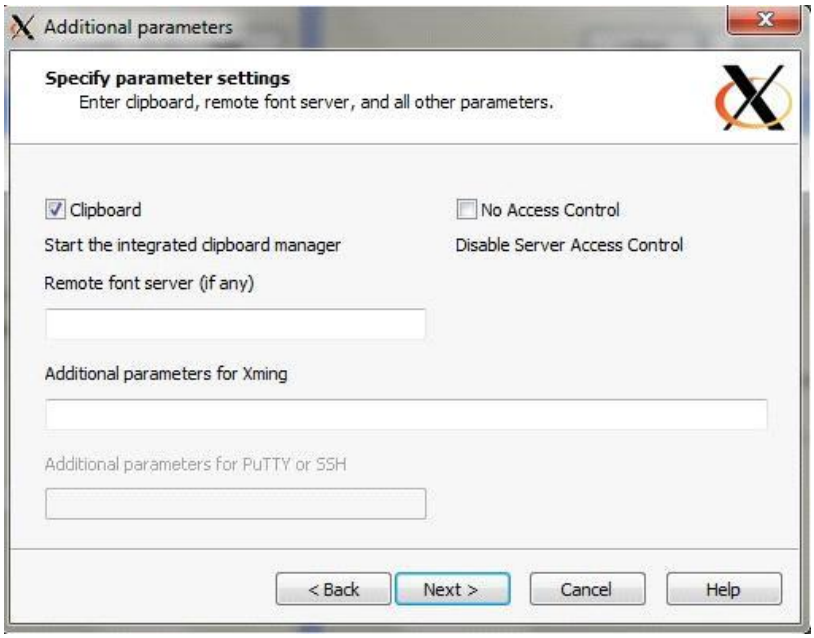

Figure 5.3 Xming Configuration step 3

Finally, click ”Next” button shown in Figure 12-3, it goes to last configuration dialogue box where ”Finish” button is needed to click for the completion of the setting. After Completion of configuration, double click the application named as”Xming”.

5.2 PUTTY

This is the primary software need to be installed. It can be downloaded in the provided link as follow, Download HYPERLINK “http://www.chiark.greenend.org.uk/~sgtatham/putty/” HYPERLINK “http://www.chiark.greenend.org.uk/~sgtatham/putty/”Putty HYPERLINK “http://www.chiark.greenend.org.uk/~sgtatham/putty/”. As, it is downloaded, it needs to be installed following some few normal steps of installation. For Configuration, double click the icon of Putty after the completion of installation and enter the IP address of Raspberry Pi as shown in Figure 5.4.

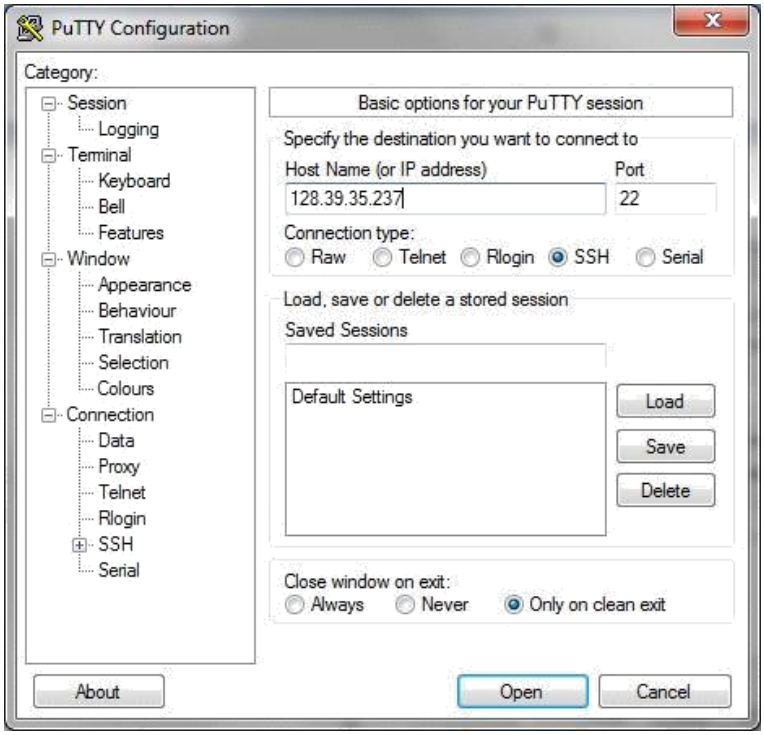

Figure 5.4: PuTTy Configuration Step 1

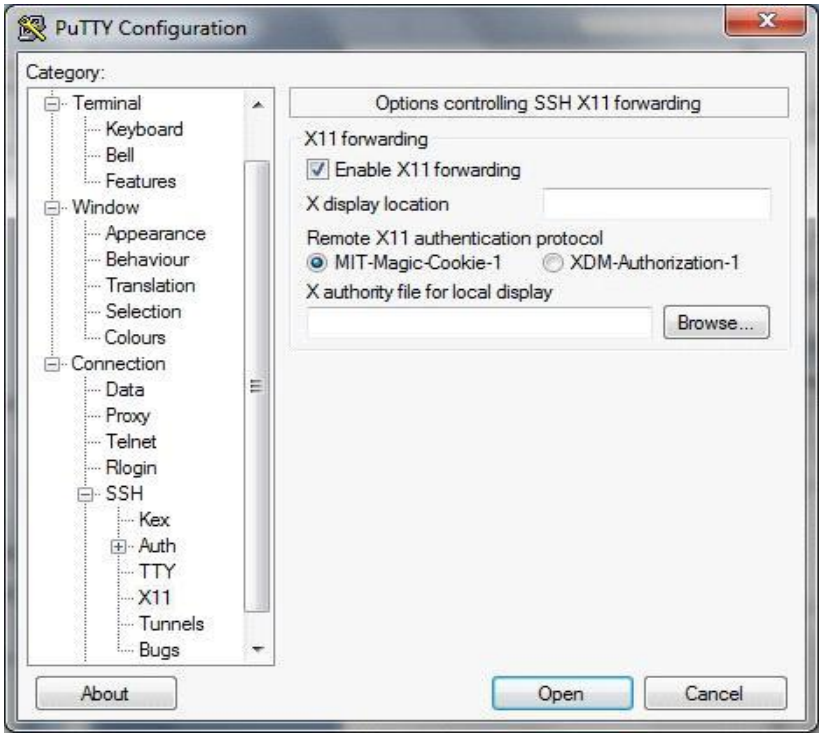

Figure 5.5 Putty Configuration step 2

Finally, click on “open” button and as a result; window will be displayed as shown in Figure 5.5.

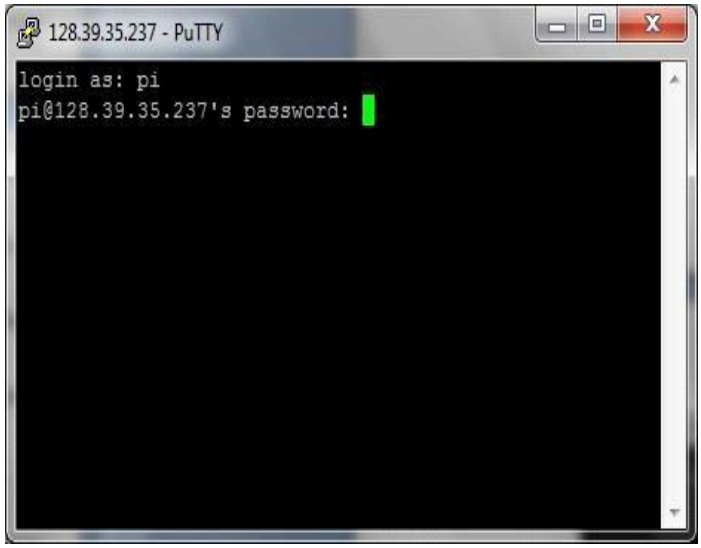

Figure 5.6 PuTTy login windows

In this window, username and password of the Raspberry Pi should be entered. The default username for a Raspberry Pi is pi. Press enter after entering correct username and password and you will see the windows where you should enter the text ”lesson” as shown in Figure 5.6.

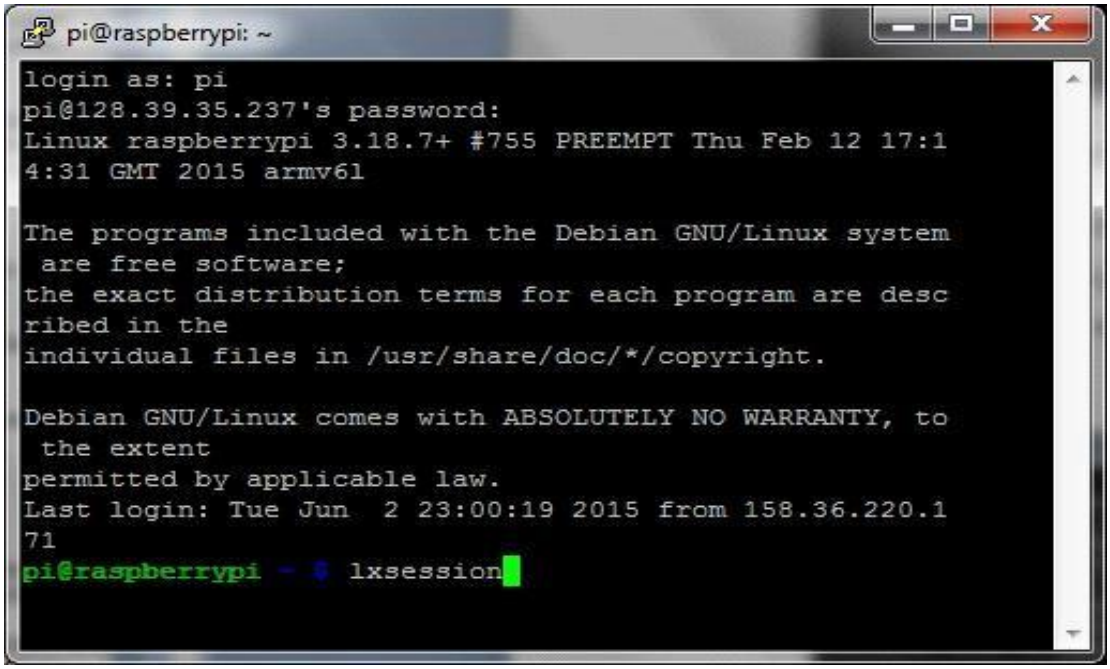

Figure 5.7 PuTTy success windows

Then, you will be able to see the desktop of Raspberry Pi in the laptop as shown in Figure 5.7 and mouse and keyboard can be used for the Raspberry Pi.

5.3 INSTALLATION OF APPLICATIONS ON RASPBERRY PI

There are many applications that are needed to install in the Raspberry Pi for the completion of the thesis. For data logging, MySQL apache5 and php my admin are needed to install whereas for the web-page development, PHP is needed to install. Web page is used for the monitoring and managing purpose

5.4 OCR

Optical character recognition is the translation of optically scanned bitmaps of printed or written text into digitally editable data files. OCRs developed for many world languages are already under efficient use.

Computer Vision:

Human intelligence is described as the capability to make decisions based on information which is incomplete and noisy. This is the ability which makes human being the most superior amongst all living creatures in the world.

There are five senses that provide information to humans for making everyday decisions and out of these the senses of hearing and vision are the sharpest of all. The auditory sense helps us recognize sounds and classify them. It is this sense which tells us that the person on the phone is a friend because his voice is *recognizable*. We can differentiate between an endless variety of sounds, voices, utterances and put them in exactly the slots they belong to, animal sounds, musical notes, wind swishing, the footsteps of a family member, all are within recognition range of a person with a normal sense of hearing.

The other one and the one more profound is the human vision which allows us to identify a known person in a crowd of unknowns merely by casting

a cursory glance at him. It allows us to pick an object that belongs to us from a number of those looking exactly as ours, and by being able to recognize a miss-spelt word in a sentence and unconsciously correct it. The fact is that the human mind is capable of identifying an image on features spontaneously determined and not predefined or predetermined.

With the development of technology these human processes are imitated to create intelligent machines, so much for the immense growth of robotics and intelligent decision making systems and yet the work done so far is not comparable to any natural involuntary human action or process. The hindrance however is that it is not practically possible to imitate all the functions of the human mind and make computer vision as efficient and accurate as the human eye but even though such a possibility may be remote, efforts are consistently being made to bring them as close to it as possible.

Artificial Intelligence is a broad field of computer science taking into account a number of other disciplines that form the bulk of its study. One of the most popular definitions of Artificial Intelligence (AI) was given by Elaine Rich and Kevin Knight as, “Artificial Intelligence (AI) is the study of how to make computers do things which, at the moment, people do better” [20]. One important branch of Artificial intelligence is Computer vision, shows few sub-branches of Computer Vision, the area of research which aims to imitate human vision and forms the basis of all image acquisition, its processing, document understanding and recognition. Computer vision relies on a solid understanding of the physical process of image formation to obtain simple inference from individual pixel values like shape of the object and to recognize objects using geometric information or probabilistic techniques

In its own turn document understanding is a vast and difficult area for the focus of research today lies in being able to make content based searches which

hope to allow machines to look beyond the key words, headings or merely topics to find a piece of information. A far more streamlined field of Document Recognition and understanding is Optical Character Recognition which attempts to identify a single character from an optically read text image as a part of a word that can be then used to process further information on. The area gains rising significance as more and more information each day needs to stored processed and retrieved rather than being keyed in from an already present printed or handwritten source.

Character Recognition:

Character recognition is a sub-field of pattern recognition in which images of characters from a text image are recognized and as a result of recognition respective character codes are returned, these when rendered give the text in the image.

The problem of character recognition is the problem of automatic recognition of raster images as being letters, digits or some other symbol and it is like any other problem in computer vision

Character recognition is further classified into two types according to the manner in which input is provided to the recognition engine. Considering figure 5.8 which shows the classification hierarchy of character recognition, the two types of character recognition are:

- On-line character recognition
- Off-line character Recognition

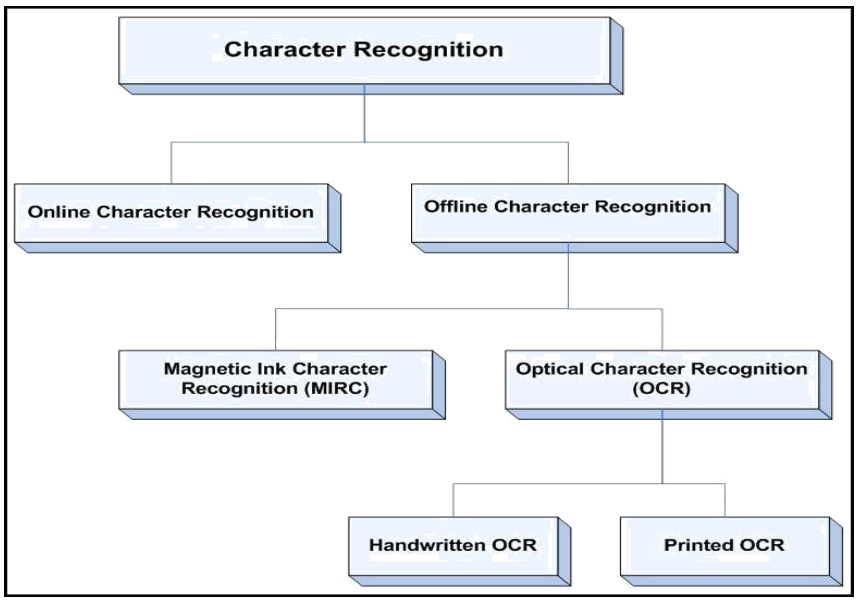

Figure 5.8 Classification of Character Recognition

Online Character Recognition:

Online character recognition systems deal with character recognition in real time. The process involves a dynamic procedure using special sensor based equipment that can capture input from a transducer while text is being written on a pressure sensitive, electrostatic or electromagnetic digitizing tablet. The input text is automatically converted with the help of a recognition algorithm to a series of electronic signals which can be stored for further processing in the form of letter codes. The recognition system functions on the basis of the x and y coordinates generated in a temporal sequence by the pen tip movements as they create recognizable patterns on a special digitizer as the text is written.

Offline Character Recognition:

There is a major difference in the input system of off-line and on-line character recognition which influences the design, architecture and methodologies employed to develop recognition systems for the two. In online recognition the input data is available in the form of a temporal sequence or real time text generated on a sensory device thus providing time sequence contextual

information. On the contrary in an off line system the actual recognition begins after the data has been written down as it does not require real time contextual information.

Offline character recognition is further classified in two types according to the input provided to the system for recognition of characters. These are,

- Magnetic ink Character Recognition (MICR)
- Optical Character Recognition (OCR)

Magnetic ink Character Recognition (MICR):

MICR is a unique technology that relies on recognizing text which has been printed in special fonts with magnetic ink usually containing iron oxide. As the machine prepares to read the code the printed characters become magnetized on the paper with the North Pole on the right of each MICR character creating recognizable waveforms and patterns that are captured and used for further processing. The reading device is comparable to a tape recorder head that recognizes the wave patterns of sound recorded on the magnetic tape. The system has been in efficient use for a long time in banks around the world to process checks as results give high accuracy rates with relatively low chances of error. There are special fonts for MICR, the most common fonts being E-13B and CMC-7.

Optical Character Recognition (OCR) :

Optical Character Recognition or OCR is the text recognition system that allows hard copies of written or printed text to be rendered into editable, soft copy versions. It is the translation of optically scanned bitmaps of printed or written text into digitally editable data files. An OCR facilitates the conversion

of geometric source object into a digitally representable character in ASCII or Unicode scheme of digital character representation .

Many a times we want to have an editable copy of the text which we have in the form of a hard copy like a fax or pages from a book or a magazine. The system employs the use of an optical input device usually a digital camera or a scanner which pass the captured images to a recognition system that after passing it through a number of processes convert it to a soft copy like an MS Word document.

When we scan a sheet of paper we reformat it from hard copy to a soft copy, which we save as an image. The image can be handled as a whole but its text cannot be manipulated separately. In order to be able to do so, we need to ask the computer to recognize the text as such and to let us manipulate it as if it was a text in a word document. The OCR application does that; it recognizes the characters and makes the text editable and searchable, which is what we need. The technology has also enabled such materials to be stored using much less storage space than the hard copy materials. OCR technology has made a huge impact on the way information is stored, shared and communicated.

OCRs are of two types,

- OCRs for recognizing printed characters
- OCRs for recognizing hand-written text.

OCRs meant for printed text recognition are generally more accurate and reliable because the characters belong to standard font files and it is relatively easier to match images with the ones present in the existing library. As far as hand writing recognition is concerned the vast variety of human writing styles and customs make the recognition task more challenging. Today we have OCRs

for printed text in Latin script as an everyday tool in offices while an OCR for hand writing is still in the research and development stage to have more result accuracy.

Optical Character Recognition (OCR) is one of the most common and useful applications of machine vision, which is a sub-class of artificial intelligence, and has long been a topic of research, recently gaining even more popularity with the development of prototype digital libraries which imply the electronic rendering of paper or film based documents through an imaging process.

History of OCR:

The history of OCR dates back to the early 1950s with the invention of *Gismo* a machine that could translate printed messages into machine codes for computer processing. The product had been a combined effort of David Shepard, a cryptanalyst at (Armed Forces Security Agency) AFSA and Harvey Cook. The successful achievement was then followed by the construction of the world's first OCR system, also by David Shepard under his Intelligent Machines Research Corporation. Shepard's customers included The Readers' Digest, Standard Oil Company of California for making credit card imprints, Ohio Bell Telephone Company for a bill stub reader and The U.S. Air force for reading and transmitting teletype messages.

Another success came in the area of postal number recognition, which although crude in its inception stage became more refined with advancement in technology. By 1965 Jacob Rainbow an American inventor had patented an OCR for sorting mails which was then used by the U.S. Postal Service. The OCR has since become an interesting field of research posing numerous complexities and offering unique possibilities in the areas of Artificial

Intelligence and Computer Vision. As advancements were made and new challenges were undertaken it became more and more clear that the scope of such an undertaking, though attractive would be daunting and adventurous.

OCR Processes:

The OCR process begins with the scanning and subsequent digital reproduction of the text in the image. It involves the following discrete sub-processes, as shown in figure 5.3.

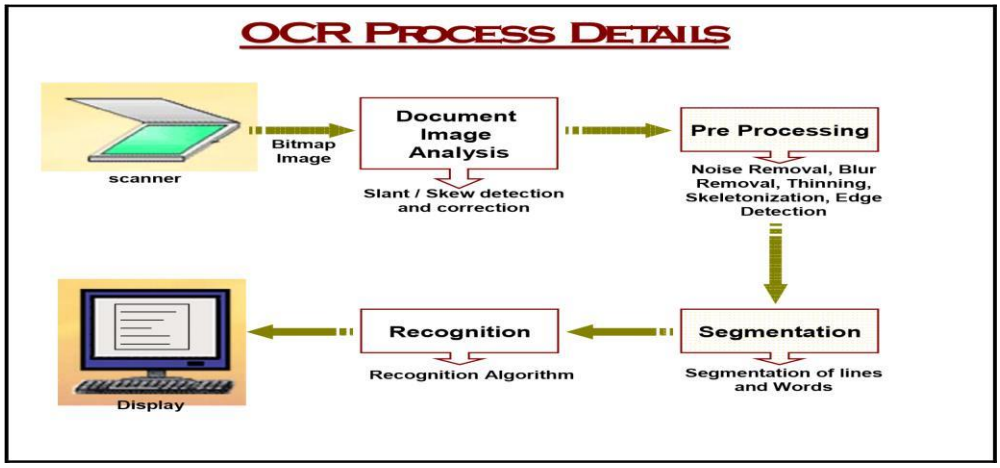

Figure 5.9 OCR Processes

Scanning:

A flat-bed scanner is usually used at 300dpi which converts the printed material on the page being scanned into a bitmap image.

Document Image Analysis:

The bitmap image of the text is analyzed for the presence of skew or slant and consequently these are removed. Quite a lot of printed literature has combinations of text and tables, graphs and other forms of illustrations. It is

therefore important that the text area is identified separately from the other images and could be localized and extracted.

Pre-processing:

In this phase several processes are applied to the text image like noise and blur removal, banalization, thinning, skeletonization, edge detection and some morphological processes, so as to get an OCR ready image of the text region which is free from noise and blur.

Segmentation:

If the whole image consists of text only, the image is first segmented into separate lines of text. These lines are then segmented into words and finally words into individual letters. Once the individual letters are identified, localized and segmented out in a text image it becomes a matter of choice of recognition algorithm to get the text in the image into a text processor.

Recognition:

This is the most vital phase in which recognition algorithm is applied to the images present in the text image segmented at the character level. As a result of recognition character code corresponding to its image is returned by the system which is then passed to a word processor to be displayed on the screen where it can be edited, modified and saved in a new file format.

CHAPTER 6

RESULTS AND CONCLUSION

A prototype system to read printed text on hand-held objects for assisting blind persons has been described. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. This method can effectively distinguish the object of interest from background or other objects in the camera view. The figure 6.1 and 6.2 describes the running of the program used in the described prototype.

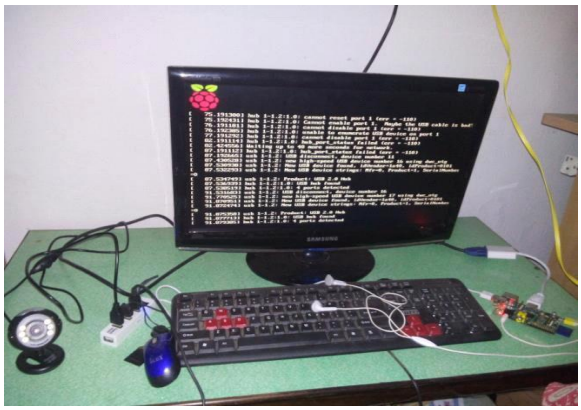

Figure 6.1 Running process of the code

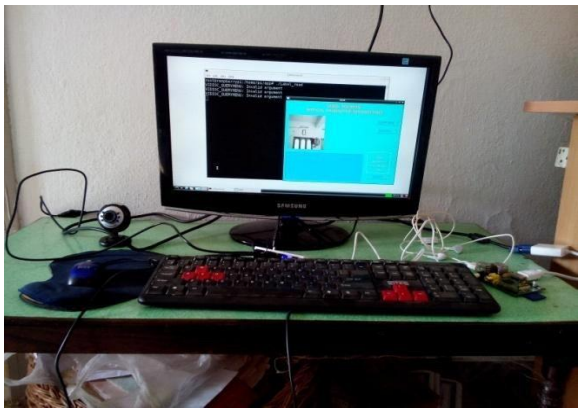

Figure 6.2 Image Capturing Result

To extract text regions from complex backgrounds, we have proposed a novel text localization algorithm based on models of stroke orientation and edge distributions. The corresponding feature maps estimate the global structural feature of text at every pixel. Block patterns project the proposed feature maps of an image patch into a feature vector. Adjacent character grouping is performed to calculate candidates of text patches prepared for text classification. An Adaboost learning model is employed to localize text in camera-based images. Off-the-shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users. Our future work will extend our localization algorithm to process text strings with characters fewer than three and to design more robust block patterns for text feature extraction. We will also extend our algorithm to handle no horizontal text strings. Furthermore, we will address the significant human interface issues associated with reading text by blind users It has been developed by integrating features of all the hardware components and software used. In this paper, the camera acts as input for the paper. As the Raspberry Pi board is powered the camera starts streaming. The streaming data will be displayed on the screen using GUI application. When the object for label reading is placed in front of the camera then the capture button is clicked to provide image to the board. Using Tesseract library the image will be converted into data and the data detected from the image will be shown on the status bar. The obtained data will be pronounced through the ear phones using Flite library. Presence of every module has been reasoned out and placed carefully thus contributing to the best working of the unit. In this by using highly advanced ARM11 board this paper has been implemented.

REFERENCES

1. E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri (2017)“An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking”, Institute of Electrical and Electronics Engineering,vol-18,pp-2568-2576.

2. Mallapa D.Gurav, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I.Byakod, Shivaleela Kanade (2017) “A Reading aid for the Blind People using OCR and OpenCV”, International Journal of Scientific Research Engineering and Technology, Vol.-6,pp-546-548.

3. Miss. Kirti P. Bhure, Mrs. J. D. Dhande (2017) “Object Detection Methodologies for Blind People”, International Journal of Research in Information Technology, vol-5,pp-194-198.

4. Mrs. Shilpa Reddy K, Mounika S.k, Pooja K, Sahana N (2017), “Text to Speech for the Visually Impaired”,International Research Journal of Computer Science,vol.-4,issue 05,pp-220-225.

5. S. M. Lucas(2005), “ICDAR 2005 text locating competi-tion results,” in Proc.Int. Conf. Document Anal. Recognit, vol. 1, pp. 80–84.

6. Shahab, F. Shafait, and A. Dengel (2011), “ICDAR 2011 robust reading competition: ICDAR Robust Reading Competition Challenge 2: Reading text in scene imag-es,” in Proc. Int. Conf. Document Anal. Recognit,vol-7, pp. 1491–1496.

7. Suchita Wankhade, Mrunali Bichukale, Shruti Desai, Shraddha Kamthe, Archana Borate (2017) “Smart Stick for Blind People with Live Video Feed”, International Research Journal of Engineering and Technology,vol.-4,issue-03,pp-1774-1778.

8. X. Chen, J. Yang, J. Zhang, and A. Waibel (Jan. 2004), “Auto-matic detection and recognition of signs from natural scenes,” Institute of Electrical and Electronics Engineering Trans. Image Process.,vol. 13, no. 1, pp. 87–99.

APPENDIX- I

PUBLICATIONS

Batchu Aditya, G.Dhanushkodi, J.A.Anest Bala Rubasingh, V.Ajit Babu and T.Helan Vidhya, "SMART ASSISTANCE FOR BLIND PEOPLE USING RASPBERRY PI 3" in 2017 International Conference on "Smart Structures and Systems", Chennai, pp-106-109.

Smart Assistance for Blind People using Raspberry Pi3

Batchu Aditya[1], G.Dhanushkodi[2], J. Anest Bala Rubasingh[3], V.Ajit Babu[4], T.Helan Vidhya[5]
1,2,3,4 UG Students and 5Assistant Professor
Department of Electronics and Communication Engineering,
Rajalakshmi Engineering College, Chennai, Tamil Nadu, India.
1batchu.aditya9@gmail.com, 2dhanushkodi1107@gmail.com, 3anestjarputharaj@gmail.com,
4ajitbabu96@gmail.com, 5helanvidhya.t@rajalakshmi.edu.in

Abstract--Object detection is technology of computer vision and image processing which deals with detection of instances of semantic objects of particular classes in digitalized images and videos. Outline of this paper is a system for detecting the type of obstacle; converting text into audio for reading purpose and detecting the kind of obstacle for blind people using Raspberry pi3.This technique helps the blind people to see the world virtually with an audio output. Optical Character Recognition (OCR) algorithm can be used to translate print version or hand-written text into audio output using Raspberry Pi. Ultrasonic sensors are attached with motor which detects the presence of obstacle in 180°. In this proposed model Raspberry pi3 is the medium which provides interface between Ultrasonic sensor and camera, which is implemented to provide the required output to the blind user that can be listened through headphones.

Keywords—OCR algorithm, ultrasonic sensor, Raspberry Pi3, motor

I. INTRODUCTION

By the survey of World Health Organisation (WHO), 254 million people are visually impaired,out of which 37 million are blind and 217 million people have low vision. Out of 37 million people 15 million people were from India and 26% of them were children [1]. The statistical analysis of percentage of blind people with various abnormalities is shown in fig.1. These people struggle every day from moving to other places without the help of others. As technologies has dominated the people's task in various fields, this research is to make the life of blind people in much easier way to walk from one place to another and to read like normal people with smart assistance. In the process of assisting the blind people, recent trends in development of computer vision, camera and Raspberry pi has made it easy to guide the visually challenged people with camera based products and existing products such as OCR [4]. With these, smart assistance can be built to assist the blind people in both guiding them from one place to another as well as to read. OCR system is sub-branch of computer vision and in turn a class of AI. OCR can be translated by scan bitmaps of print version or hand-written text into audio output using Raspberry Pi. OCR could be developed for global languages which are already under significant use. This OCR method is used in extracting region of moving objects by a mixture-of- Gaussian-based subtraction method. Text localization, recognition are always conducted to get text information [7].

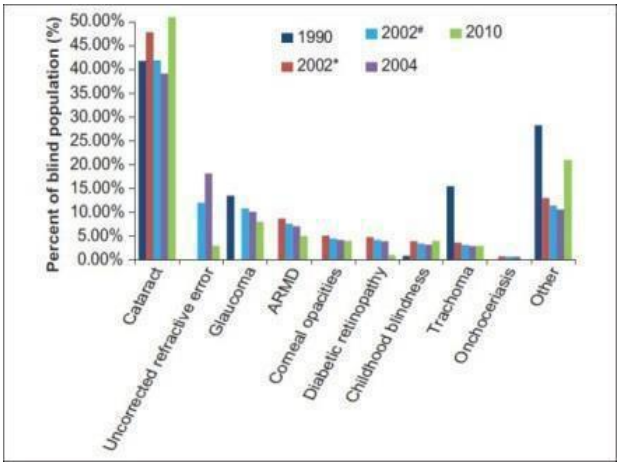

Fig 1. Percentage of blind population with various abnormalities.

For automatically localization of text from the objects, a text localization [7] and Tesseract algorithm [2] with learning of required features of stroke orientations and distribution of edge pixels in Ada boost model. The recognized text codes are the output to the blind people in speech recognization.The Performance of proposed text localization algorithm is, when recognition process is completed, the character code of the text file are processed by Raspberry pi on which characters are recognized using Tesseract algorithm [2],[7] and the audio output is listened through the headphones.

This paper is arranged that the next section describes the proposed methodology followed by Hardware and Software used in this model and final section contains the conclusion and future work of the project.

II. PROPOSED SYSTEM

In proposed system, a prototype system shown in fig. 2 reads printed text on hand-held objects for assisting blind persons has been described. OCR is a vast field of

professional research in recognizing patterns and in artificial intelligence and computer vision [4]. It is widely used as a form of information entry from printed paper. It is a method of digitizing printed text so they can be electronically edited, searched, stored, more compact, displayed on-line and used in machine process such as cognitive computing, machine translation ,text-to-text, key data and text mining. In order to solve the common aiming problem for blind users, a motion-based method to detect the object of interest has been proposed, while the blind user simply shakes the object for a couple of seconds. The automatic detection of region of interest, text localized algorithm was evaluated individually as unit tests to ensure its effective and robust features of the whole system. This prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person has been evaluated. Two calibrations were applied to prepare for the system test. First, instruction is given to the blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, a camera with a reasonably wide angle range has been employed.

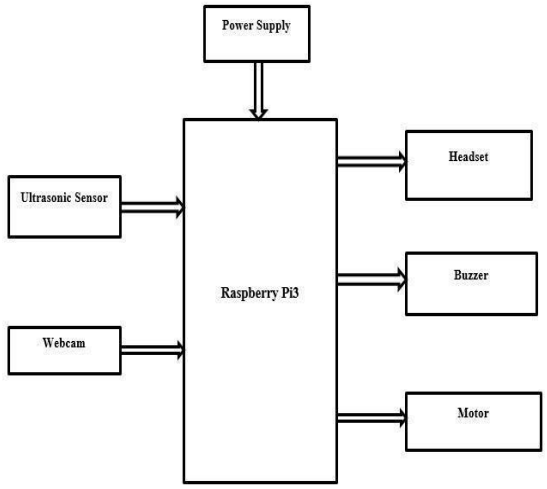

Fig. 2 Block diagram of the proposed work

Second, in an applicable system, a text localized algorithm may prefer higher recall by sacrificing some precision and accuracy. By using ultrasonic sensor, distance between the blind people and the obstacle can be measured and then the distance is heard through ear phones[3]. Also as the motor has been attached to the ultrasonic sensor, the object at an angle of 180° in front of the blind user can be detected and the output recognized by OCR algorithm [4]. This output is the electronic conversion of images of typed, handwritten or printed text, to form a scanned document; a scene-photo text is superimposed on an image. In existing system, images of each character are to be trained, and worked on one font at a time. Advanced systems are worthy enough to produce high degree recognition of accuracy for most of the fonts which are usual now and with support for a variety of digital image file format inputs. Some systems are capable of reproducing formatted output that closely approximates the original page including images, columns, and other non-textual component.

III. HARDWARE IMPLEMENTATION

Raspberry pi3 is a single board computer which is used to detect the object and read the content. The program to accomplish the required task is installed in Raspberry pi in python language.

Ultrasonic sensors are used which evaluate attributes of a target by interpreting the echoes from radio or sound waves respectively[3]. It generates high frequency sound waves and evaluates the echo which is received by the sensor, sensor calculates the time interval between sending the signal and receiving the echo to determine the distance to an object and the resultant output is given out in the form of audio. It is mainly used to detect the presence of obstacle in front of the user and the output is indicated with the help of the headphones. It provides precise, non-contact distance measurements within a 3cm to 3m range.

Camera is as main device in detecting image of the product or board then image is processed internally and separate labels from images by Open CV library and finally identifies the product and identified product name is pronounced through voice.

IV. SOFTWARE IMPLEMENTATION

Raspbian OS is Debian based computer OS for raspberry pi3. It is a free operating system. Raspbian was created by MIKE THOMPSON and PETER GREEN as an independent project. It was built over 35000 Raspbian packages which are based on best performance on Raspberry pi. It is still under gradual development with an importance on improving the performance and stability. This software is the platform in which all the process is been carried out.

Python IDE coding tool is platform where the code detects an obstacle and the obtained code is used to convert the captured image into audio output. Python language has been chosen as coding tool because the language is very compatible and easy for coding and user friendly in nature. IDE stand for Integrated Development Environment which is a coding tool that is used to write text and also for testing and debugging the code in most easier way, as they generally offer code completion or code insight by highlighting, resource management, debugging tools, and even though the IDE is a strictly defined concept, it started to redefine as other tools such as notebooks started gaining more and vast features that traditionally belongs to IDEs. For example, debugging code could also be possible in Jupiter Notebook. Because of all the features that IDEs offer, they are extremely useful for development: they

make the coding more efficient and comfortable and this would be no different for data science. However, given the fact that there aren't only the traditional IDEs to be considered, but also new tools, such as notebooks, you might be wondering which development environment to use when you're just starting out with data science.

V. SYSTEM ARCHITECTURE

A system is proposed to read printed text and to detect the obstacle. As shown in fig.3, the components are connected in such a way that buzzer, webcam and ultrasonic sensor are connected in GPIO and monitor and keyboard are connected in USB port and headphone is connected in of Raspberry pi [3]. Raspbian OS is stored in 16GB micro SDcard and connected in micro SDcard slot. A power supply of 230V AC supply is given to the Raspberry pi to start the process. Rectifier is used to convert alternate current to DC current in order to support the step down transformer produces 12 AC constantly. Raspberry pi board should be placed on any non-metal surface in order to avoid short circuit.

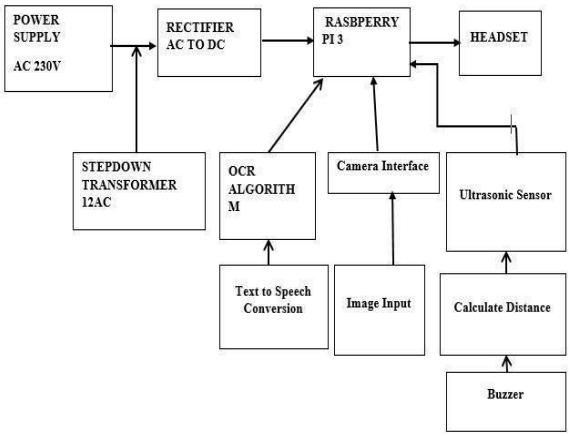

Fig 3. Block diagram of system architecture

In order to make the ultrasonic sensor work, a python coding is fed into the Raspberry pi and the program is executed. The distance of the minimum ultrasonic range, which is set to a range of 200 cm in the program. To run the python code, Python IDE is used. The output of the code is results in alarming the buzzer.

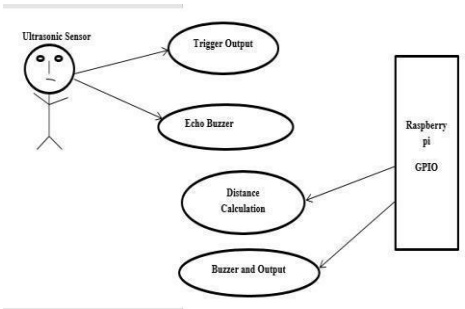

Fig.4 User case diagram for object detection[5]

From fig.4, the working of ultrasonic sensor is explained very clearly[3]. The trigger output and echo input are connected to the Raspberry pi GPIO. Once the echo input is transmitted by the ultrasonic sensor, gets reflected by the obstacle and is received by the trigger output [3].The distance is measured by the formula, Pulse duration = Pulse end – Pulse Start ----- (1) Distance = Pulse duration * 17150 ---------- (2)

The output of the program is shown in Table I

TABLE I. OUTPUT OF BUZZER WITH MEASURED DISTANCE

Distance(cm)	Buzzer output
2000	OFF
200	ON
145	ON

Usage of webcam is very useful as it captures the image as input which is to be converted into audio output. Webcam used here is an 360° camera which covers the entire view of the user. To process the input image, OCR algorithm is fed. OCR algorithm converts the captured image into the machine readable format, which is converted and stored. Text to Speech Conversion coding is executed simultaneously so that once the OCR algorithm coding converts image into machine readable form , this coding performs the task of converting the stored machine readable format into audio output[6]. Then the resulted output is obtained with the use of headphones.

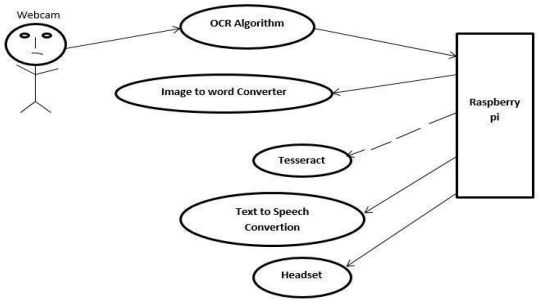

Fig.5 User case diagram for text to speech conversion

From fig.5, the working of the webcam with OCR algorithm and text to speech conversion can be viewed. First, when the image is captured by the webcam, the image is processed by the OCR algorithm by Tesseract [2],[4],[6]. Then it is processed by text to speech conversion process, which results in audio output through headphone[6].

VI. CONCLUSION AND FUTURE WORK

In this proposed model the Text to Speech conversion technique, Image and obstacle detection technique has been successfully implemented using Raspberry pi3and ultrasonic sensor[3],[6] .The OCR algorithm is used for processing the image and audio output that is obtained is audible and clear [4]. It is an economical device which anyone can afford and can be used very efficiently and accurately. This system is very

compatible and really helpful device for the blind people.In future systems, the finger point detection and tracking can be added to adaptively instruct blind people to aim for the object. Also image processing is used to identify the type of object which can be added as a useful feature to help the blind user to detect the kind of object in front of them. The concept of YOLO is a real time neural network which is added to detect the type of object which is in front of them and address the blind user in the form of audio output, which makes the user to see the real world

VII. REFERENCES

[1] Advance Data Reports from the National Health In-terview Survey (2008).

[2] An overview of the Tesseract OCR (optical charac-ter recognition) engine, and its possible enhancement for use in Wales in a pre-competitive research stage Prepared by the Language Technologies Unit (Canol-fan Bedwyr), Bangor University April 2008.

[3] E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri "An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking", IEEE Journal,2017.

[4] MallapaD.Gurav, Shruti S. Salimath, Shruti B. Hatti,VijayalaxmiI.Byakod,ShivaleelaKanade"A Reading aid for the Blind People using OCR and OpenCV", IJSRET Journal 2017.

[5] Miss. Kirti P. Bhure, Mrs. J. D. Dhande "Object Detection Methodologies for Blind People", IJRIT Journal 2017.

[6] Mrs.Shilpa Reddy K, MounikaS.k, Pooja K, SahanaN, "Text to Speech for the Visually Impaired", IRJCSJournal, 2017.

[7] S. M. Lucas, "ICDAR 2005 text locating competi-tion results," in Proc.Int. Conf. Document Anal. Recog-nit., 2005, vol. 1, pp. 80–84.

[8] Shahab, F. Shafait, and A. Dengel, "ICDAR 2011 robust reading competition:ICDAR Robust Reading Competition Challenge 2: Reading text in scene imag-es," in Proc. Int. Conf. Document Anal. Recognit., 2011, pp. 1491–1496.

[9] Suchita Wankhade, MrunaliBichukale, Shruti Desai, ShraddhaKamthe, Archana Borate "Smart Stick for Blind People with Live Video Feed",IRJETJournal, 2017.

[10] . Chen, J. Yang, J. Zhang, and A. Waibel, "Auto-matic detection and recognition of signs from natural scenes," IEEE Trans. Image Process.,vol. 13, no. 1, pp. 87–99, Jan. 2004

RAJALAKSHMI ENGINEERING COLLEGE

DEPARTMENT OF ECE

PROGRAM OUTCOMES (POs)

Engineering Graduates will be able to:

PO1 Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO2 Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3 Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4 Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5 Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6 The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7 Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8 Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9 Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10 Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11 Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12 Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM SPECIFIC OUTCOMES (PSOs)

PSO1: An ability to carry out research in different areas of Electronics and Communication Engineering fields resulting in journal publications and product development.

PSO2: To design and formulate solutions for industrial requirements using Electronics and Communication engineering.

PSO3: To understand and develop solutions required in multidisciplinary engineering fields.

COURSE OUTCOMES (CO)

CO1	To conceive an idea and develop confidence in designing, analyzing and executing the project in the emerging fields of Electronics and Communication and multidisciplinary research areas.
CO2	Identification of modern tools for the implementation of project through simulation and prototype.
CO3	Develop products that meet the specified needs in industrial applications with appropriate consideration for the public health and safety, societal, environmental and ethical considerations.

EC6811 –PROJECT WORK

Project Title: SMART ASSISTANCE FOR BLIND PEOPLE USING RASPBERRY PI 3

Batch Members: Ajit Babu V (211614106015)

Anest Bala Rubasingh J A (211614106022)

Batchu Aditya (211614106038)

Dhanushkodi G (211614106049)

Name of the Supervisor: T.Helan Vidhya

CO - PO – PSO matrices of course

PO/PSO / CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	2	3	3	3	3	3	3	3	3	3	2	3	1	3
CO2	3	3	2	3	3	-	-	-	2	1	1	1	2	2	2
CO3	3	3	3	3	3	2	2	3	3	3	3	2	3	2	2
Average	3	2.66	2.66	3	3	2.5	2.5	3	2.66	2.33	2.33	1.66	2.66	1.66	2.33

Note: Enter correlation levels 1, 2 or 3 as defined below:

1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High), If there is no correlation, put -“

Signature of the Supervisor