Multi Utility Equipment for Assisting Disable Persons

Bangala Sairam, S.P.V. Subba Rao, T. Ramaswamy

Abstract--- In our society disable people faces lot of problems in terms of communication and literacy. When we deals with blind people they faces difficulty in reading the normal books and newspapers which are not Braille scripted and we face lot of problems while communicating with deaf and dumb people. So, this project focuses to find the unique solution to these problems. The project mainly uses four techniques such as image to text, text to speech, speech to text and sign to speech conversion techniques. Visually impaired people can hear the text on paper using image to text and text to speech techniques this is can achieved by Tessaract OCR (Optical Character Recognition) technique and espeak speech synthesizer. Vocally and hearing impaired people can express their views by using sign to speech and speech to text conversion techniques. This makes use of Raspberry pi, camera, microphone, GSM, GPS and speakers to build the system. This is a method where a computer is made to speak.

Keywords--- Optical Character Recognition (OCR), Raspberry pi(RPI), espeak, GSM, GPS, Speakers, Microphone, camera etc.

I. INTRODUCTION

According to World Health Organization(WHO) it is estimated approximately 1.3 billion people suffers with some form of vision impairment in which 36 million are blind and 466 million people suffers with hearing impairment across the world[9]. These people face a lot of problems in our society especially in terms of literacy and communication. So in this project we developed an image to speech, speech to text and sign to speech based application for the disable people using raspberry pi**.** The project present work is also developed to provide a security mode to blind people using GSM and GPS module. The GPS system finds the current location it stores the data of the current location, so it allows a blind user to send his/her location to their relatives through SMS using GSM whenever they are in panic situation. The blind people text reading requires the use of Braille reading system and digital speech synthesizer [1]. Since the blind people are not able to read the simple warnings in walls that surround us. The development of this application that can perform the image to speech conversion. Image to text and text to speech techniques achieved by Tesseract OCR (Optical Character Recognition) the OCR technology is enables the recognition of texts from image, the OCR has been widely used in scanned or photographed documents [3]. The speech synthesizer is enables text in digital format into human voice and played through an audio system. Text to speech system is automatic conversion of sentences into speech[5]. This project also deals with gesture Recognition and speech to text conversion using OpenCV, espeak and Google speech Recognition API tools.

II. LITERATURE SURVEY

I. Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Persons by Chucai Yi[1] : This paper propose a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives. This paper proposes an efficient and effective motion-based method to define a region of interest (ROI) in the video by asking the user to show the object. This method extracts moving object region by a mixture-of-Gaussians based background subtraction method. In the extracted ROI, text localization and recognition are conducted to acquire text information. Text characters in the localized text regions are then binaries and recognized by off-the-shelf optical character recognition (OCR) software. The recognized text codes are output to blind users in speech

II. Review of text to speech convention by Poonam S. Shetake [5] In this paper text to speech convention transforms information stored as data or text into speech. TTS systems make it possible to access textual information over the telephone. The synthesizer produces speech signals of 16 bits, the sampling rate of which is determined by the sampling rate of the diaphone database used. A text to speech (TTS) synthesizer is a system that can read text aloud automatically, which is extracted from Optical Character Recognition (OCR).

III. Embedded Sign Language Interpreter System For Deaf and Dumb People by Geethu G Nath, Anu V S[8] :Main objective of the project is building a device that assist deaf and dumb people to convey their messages to normal people. The system consists of two main parts. The first part is hand sign recognition and the second one is speech recognition. This ensures a bi directional communication avoiding the assistance of an interpreter.

III. EXISTING SYSTEM

In the existing system blind people using the Braille scripted books and specially made Braille supported computer displays, keyboards and finger mounting sensors to read which are expensive and gesture recognition systems

Manuscript received February 01, 2019

Bangala Sairam, PG Scholar, Digital Systems and Computer Electronics, Sreenidhi Institute of Science and Technology, Yemnapet, Ghatkesar, Rangareddy, Telangana. India. (e-mail: sairam.bangala@gmail.com)

Dr.S.P.V. Subba Rao, Professor, Dept of ECE, Sreenidhi Institute of Science and Technology, Yemnapet, Ghatkesar, Rangareddy, Telangana, India. (e-mail: spvsubbarao@sreenidhi.edu.in)

Dr.T. Ramaswamy, Associate Professor, Dept of ECE, Sreenidhi Institute of Science and Technology, Yemnapet, Ghatkesar, Rangareddy, Telangana, India. (e-mail: ramaswamyt@sreenidhi.edu.in)

uses the sensor gloves to recognize the sign language. The design of gloves is bulky with lots of flip sensors and its usage is un-comfortable and expensive. For speech Recognition uses in existing systems are training modules and recorded samples which do not produce the good performance [5].

IV. PROPOSED SYSTEM

In my proposed project camera module is used for capturing the image which is further converted to text and speech to make blind people to understand the content in the books and for gesture recognition instead of using sensor gloves camera is used so that system design is simple and easy to use, inexpensive. For speech Recognition Google Speech Recognition API (application program interface) is used which producing the accurate results and better performance. In addition we also developed a security mode for blind so that whenever they are in panic situation they can protect themselves by sending their live location to relatives or nearby police station.

V. PROJECT BLOCK DIAGRAM

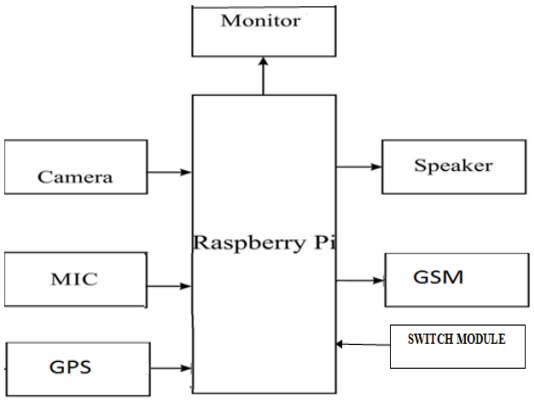

Fig 3.1: System block diagram

Block diagram description

The fig 3.1 shows the system block diagram consists of a Raspberry Pi 3, speaker and MIC, camera, we are giving 5V DC to power the Raspberry Pi. In this system we are using the mobile communication application GSM and GPS. The camera must be physically being pointed towards the text and an image picture is captured. This picture is then processed by the Raspberry Pi and the audio output is heard through the speaker.

Hardware Requirements

Hardware:

1. Raspberry pi 3
2. Camera
3. Microphone
4. Speaker
5. GSM module
6. GPS module

Functional Description

Raspberry PI

The fig 3.2 shows Raspberry Pi 3 Model B features a quad-core 64-bit ARM Cortex A53 clocked at 1.2 GHz. This puts the Pi 3 roughly 50% faster than the Pi 2. Compared to the Pi 2, the RAM remains the same – 1GB of LPDDR2-900 SDRAM, and the graphics capabilities, provided by the Video Core IV GPU, are the same as they ever were. The Pi 3 now includes on-board 802.11n Wi-Fi and Bluetooth 4.0. Wi-Fi, wireless keyboards, and wireless microphone now work out of the box.

Fig 3.2: Raspberry pi

GSM Module

GSM has become the world's fastest growing communications technology of all time and the leading global mobile standard, spanning 218 countries. The GSM is connected to raspberry pi controller. GSM is used for transmitting mobile voice and data services. Using GSM location information sent through SMS to their relatives mobile number.

GPS Module

The global positioning system and global mobile communication are interfaced to the raspberry pi controller to detect the blind people current location or area. The GPS will be sending location information to the controller.

Software

1. Raspbian OS (Linux)
2. SSH (secured shell client) PUTTY
3. VNC (Virtual Network Computing) viewer
4. Open cv (computer vision)
5. Tesseract OCR (Optical Character Recognization)
6. Espeak

Programming using python scripting language

Software Tool Description

- **Tesseract OCR (optical character recognition):**

According to the National Library of Canada, digitization usually refers to the process of converting a paper- or film-based document into electronic form bit by bit this can be achieved by OCR (Optical Character Recognition). Optical character recognition is performed using various techniques, five basic steps typically do not change. These include pre-processing, image segmentation, pattern classification, correction and post-processing. Tesseract OCR (optical character recognition) is one of the most accurate open source software run through a command line interface system. The Tesseract OCR engine was the HP Research Prototype in the UNLV Fourth Annual Test of OCR Accuracy [10].

- **Espeaks** :

Espeak is a compact open source software speech synthesizer for Linux, Windows, and other platforms. It uses a formant synthesis method, providing many languages in a small size and it has also been used by Google Translate. When espeaks are using then the speech is clear, and can be used at high speeds, but is not as natural or smooth as larger synthesizers which are based on human speech recordings [2].

- **OpenCV(computer vision)**:

OpenCV is a cross platform library using which we can develop the real time computer vision. It mainly used in image processing and video capturing the data. It includes the some feature of a face detection and object detection. We have used it for gesture recognition [8].

Working of Project

- project working can be explained in module wise such as implementation of OCR and TTS too.

Image to speech for blind

- First whatever text blind people want to read can be captured using camera. A camera is an optical instrument for recording or capturing images, which may be stored locally, transmitted to another location, or both.
- The captured image is converted in to grayscale image; this grayscale image is further converted in to threshold image.
- This threshold image is given to tessaract to extract text using optical character recognition technique (OCR).
- The extracted text string is given to speak speech synthesizer which produces the voice output.

Speech to text for deaf

- The voice input can be taken from microphone by setting the threshold value
- The time scale is set to capture and record the voice sample
- The voice sample is given to the Google speech recognition API which is used to convert the speech input in to text output

Sign to speech for dumb

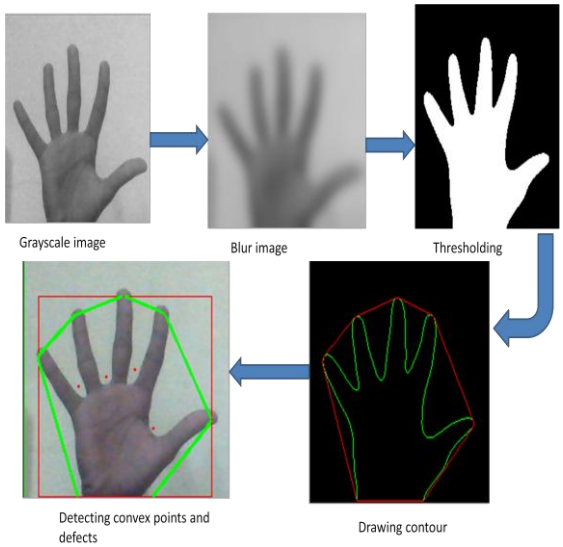

Fig 3.3: Gesture detection process

- The fig 3.3 shows the steps involved gesture detection, input can be taken using camera module
- Captured image and convert it in to gray scale image
- Define the Region of interest (ROI) here our ROI is hand
- The grayscale image is further smoothened to reduce the noise and details in the image this can be achieved by applying Gaussian blurring technique.
- Gaussian blurring is applied to track the shape of object here is hand
- Grayscale is converted to threshold image by applying Otsu's binarisation method. In this method Open CV tool is used to calculate/approximate the threshold value of a bimodal image from its image histogram.
- finally contours is drawn to image to find the convex and defect points
- The convex points are generally, the tip of the fingers. But there is other convert point too. So, we find convexity defects, which is the deepest point of deviation on the contour by this the number fingers raised can be found
- Depending upon the fingers raised the appropriate strings are programmed
- The voice output can be produced using speak speech synthesizer.

The project also focussed on security issues faced by impaired people. Whenever they struck in unknown location or attacked by strangers, they can send their live location to relatives to track them for this a panic switch is provided so that by pressing panic switch they can send location information using GPS and GSM modules.

Project Flowchart

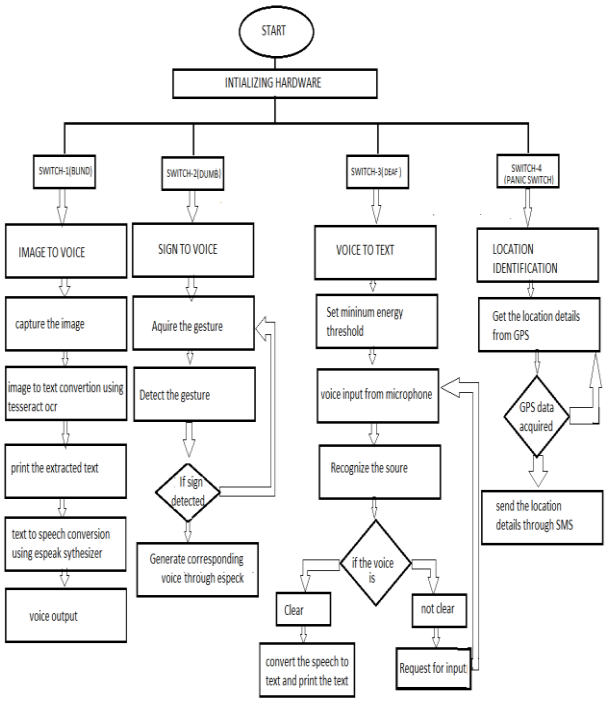

Fig 3.4: System Flowchart

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

ows the project flow chart which explains the indivual functioning of the modules. Once the system is initialized the user provided with four switches depends on required operation corresponding switch can be pressed. The detailed functioning of indivual modules has explained in previous section.

VI. RESULTS

Image to Text

Fig 4.1: Image to text speech shown on monitor

The fig 4.1 shows the image to text conversion process corresponding outputs can be seen in monitor. The text is further converted into speech.

Sign to Speech

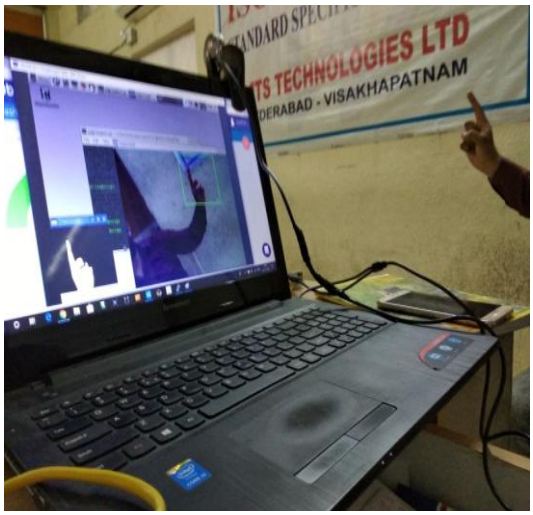

Fig 4.2: finger1 raised image capturing

The fig 4.2 shows the finger 1 sign which will be converted into speech output as “HI”.

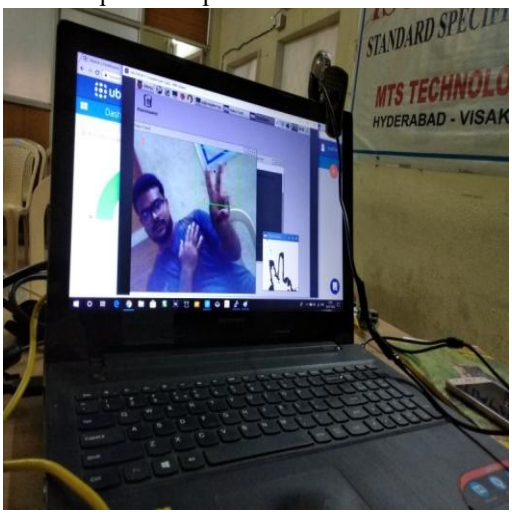

Fig 4.3: Finger2 raised image capturing

The fig 4.3 shows the finger 2 sign which will be converted into speech output as “HELLO”.

Fig 4.4: Finger3 raised image capturing

The fig 4.4 shows the finger 3 sign which will be converted into speech output as “HOW ARE YOU”

Fig 4.5: Finger4 raised image capturing

The fig 4.5 shows the finger 4 sign which will be converted into speech output as “IAM FINE”

Fig 4.6: Finger5 raised image capturing

The fig 4.6 shows the finger 5 sign which will be converted into speech output as "WHAT ABOUT YOU"

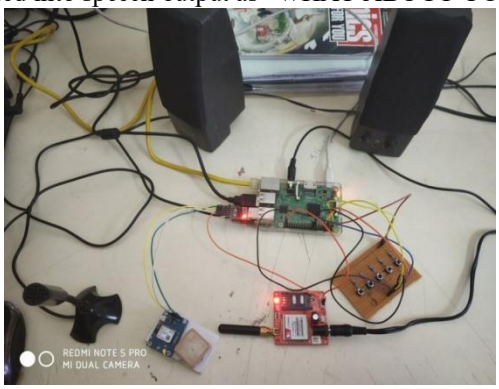

Fig 4.7: Hardware arrangement

The fig 4.7 shows the hardware arrangement of project which consists of modules like mic, camera, Raspberry pi, GSM, GPS and switch modules.

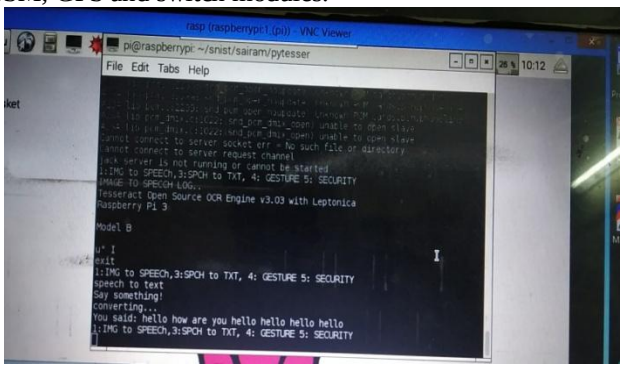

Fig 4.8: Speech to text conversion

The fig 4.8 shows the speech to text conversion output.

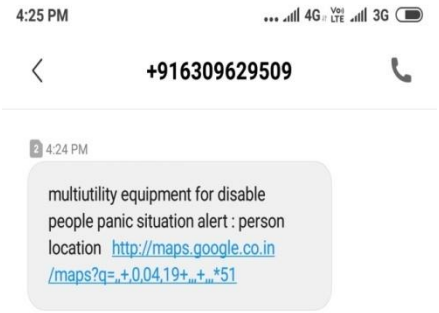

Fig 4.9: Security mode: GPS location is sent through SMS

The fig 4.9 shows the screenshot of alert message sent to relatives when they are in panic situation, the person location can be tracked by clicking on the link.

VII. CONCLUSION& FEATURE SCOPES

In this paper we have proposed a system that runs on the latest technology and makes use of image to speech, speech to text and sign to speech image capturing for disable people using OpenCV libraries, Tesseract OCR and Google speech Recognition API. Our algorithm successfully processes the image and detects the text part then reads it out clearly with set volume and also converts the speech and gesture input into text and voice respectively. In this project we have implemented security mode for assisting disable peoples using raspberry pi controller. We are using mobile communications GSM and GPS they can be send live location through SMS. Results are accurate and system design simple and with less hardware.

The feature scope of project can be extending to process handwritten text and handle non horizontal text strings. In sign to speech more signs can be incorporated so as to bring accurate results for by deaf and dumb people.

REFFERENCES

1. Chucai Yi, Student Member, IEEE , Yingli Tian, Senior Member, IEEE , and Aries Arditi "Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons" 2013 IEEE/ASME transactions on mechatronics
2. S. Venkateswarlu, D. B. K. Kamesh, J. K. R. Sastry, Radhika Rani"Text to Speech Conversion". Indian Journal of Science and Technology, Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/102967, October 2016.
3. K Nirmala Kumari, Meghana Reddy J. "Image Text to Speech Conversion Using OCR Technique in Raspberry Pi". International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 5, Issue 5, May 2016.
4. Nagaraja L, Nagarjun R S, Nishanth M Anand, Nithin D, Veena S Murthy. "Vision based Text Recognition using Raspberry Pi". International Journal of Computer Applications (0975 – 8887) National Conference on Power Systems & Industrial Automation 2015.
5. Poonam S. Shetake, S. A. Patil, P. M. Jadhav "Review of text to speech conversion methods" 2014.
6. THOMAS, S. Natural Sounding "Text-To-Speech Synthesis Based On Syllable-Like Units", Department Of Computer Science And Engineering Indian Institute Of Technology Madras, 2007.
7. DUTOIT, T. "A Short Introduction to Text-to-Speech", Kluwer Academic Publishers, Dordrecht, Boston, London, 1997.
8. Geethu G Nath, Anu V S "Embedded Sign Language Interpreter System For Deaf and Dumb People" Department of Electronics and Communication Sree Buddha College of Engineering, Pattoor Kerala, India International Conference on Innovations in information Embedded and Communication Systems (ICIIECS) 2017.
9. Vision impairment and hearing impairment @http://who.int/en/news-room/fact-sheet/deatials
10. Ray Smith "An Overview of the Tesseract OCR Engine" Google Inc, Ninth international conference on Document Analysis and Recognition (ICDAR), IEEE 2007.