

# Design and Implementation of Communication Assistance System for Deaf, Dumb and Blind People

Author: C L Nidhi<sup>1</sup>; M C Chandrashekhar<sup>2</sup>; M Z Kurian<sup>3</sup>

Affiliation: SSAHE<sup>1</sup>; SSAHE<sup>2</sup>; SSAHE<sup>3</sup>

Email: [nidhi.cl99@gmail.com](mailto:nidhi.cl99@gmail.com)<sup>1</sup>; [mccshekhar@gmail.com](mailto:mccshekhar@gmail.com)<sup>2</sup>; [mzkurian@yahoo.com](mailto:mzkurian@yahoo.com)<sup>3</sup>

**Abstract:** Human beings are able to speak, see, listen and respond according to the situations. But some people are underprivileged of responding and acknowledging. Making a solitary minimal framework for such individuals with hearing, visual and vocal hindrance is an intense activity. The proposed system portrays a new unorthodox communication system framework for specially abled people in a single flexible device. The framework contributes a method for a visually impaired individual to peruse a book and it very well may be given by capturing a picture through Logitech camera which gets the content from the picture and processes it to speech via speaker, that is Text-To-Speech (TTS) conversion using OpenCV tool. For hearing impaired people, it guides a way to read the content from the screen, speech-to-text conversion (STP) is incorporated. To assist dumb people communicate with normal people, text-to-voice conversion is done and is read out by speak. Blind people are able to read the words using Tesseract OCR (online character recognition).

**Keywords:** e-Speak, OpenCV, Speech-to-text (STP), Tesseract OCR, Text-to-speech (TTS)

## 1. INTRODUCTION

In the digital era with the advancement in information and communication technology, symbol language is the toughest way for the communication between deaf dumb people and the external world to provide the improved and easy lifestyle of dumb, deaf and blind people the proposed system is designed and developed. The proposed system helps normal and deaf dumb

people to communicate with each other effectively and easily. To resolve these barriers with visually and vocally impaired people, the system is designed by making use of a tiny creditcard-sized computer named Raspberry-Pi 3model B with OpenCV software. The proposed system provides the technique for the deaf, dumb and blind people to communicate each other as well as with the normal people. A large number of approaches have been presented using web cameras and OpenCV algorithms to detect and identify the symbols. The gesture recognition is the most crucial area for the development of the smart human communication system. The gesture identification area has huge approaches to design and research the smart interaction system. When vocally impaired or hearing impaired person tries to interact with the external world they find very difficulties in exchanging the information because the normal people don't have the knowledge about sign language. Sign language is mostly used by the vocally impaired ones and hearing impaired ones, which is a communication skill they are used instead of voice or speech to convey their messages or information to another person.

## 2. LITERARTURE REVIEW

In [1], Kanika Rastogi proposes the SHAROJAN Framework which makes utilization of the Wearable Development, Texas Instrumentation Equipment, and Arduino Circuit Sheets to give a system for correspondence to all of a sudden abled individual having one or the majority of the above say handicaps. There's another structure called the

SHAROJAN Extension with a genuine goal to cross any limit in the midst of the time spent correspondence between the outwardly weakened, in need of a hearing aid and unfit to talk individuals [2]. Rohith Rastogi et al, states that in general day by day presence correspondence, the principle issue for about hard of hearing, doltish individuals and for Stun solitary it is hard to take notes of specific things [3].

Rehman et al, proposes a structure that is utilized to beat any obstruction of correspondence among outwardly weakened, deaf and nitwit people [4]. In [5], a gadget has been proposed which can change over hand developments into a voice, the contraption includes a Remote Glove, contains accelerometer and flex sensors.

### 3. RELATED WORK

The image-based technique doesn't require any external devices like hand gloves and helmet etc. due to this advantage lot of research is going on the image-based technique approach, but sensor-based technique disabled persons need to wear devices. Symbolic sign recognition system gaining importance in many application areas such as multimedia, human interface communication and security. Disabled persons face problems in communicating with normal peoples because normal people cannot understand the symbolic language.

The conclusion is, if a wearable device like glove is used, it would increase the cost of the device and the complexity too; it is also used for one person at a time which reduces the efficiency and capability of the system proposed. The Sharojan framework is helpful as it acts as the bridge between the specially abled people. The device helps in overcoming the Braille system, like blind people were feeling it by their hands whereas this system just conveys the message through a speaker. Gesture control prototypes also increases the cost; practically, they are difficult to use in real life situations. Hand signal conversion also takes quite more time to process. Hence, the proposed system helps in overcoming all the cons of the existing system.

When the dumb people communicating with normal people, the text written by dumb is spelled out by the speaker which can be understandable by normal person. For deaf disabilities, the voice/speech is translated into text by using a website called [speechtexter.com](http://speechtexter.com). The proposed system in this paper is a portable device where the user sends and receives a message as a text. The device acts as a mobile phone. When people try to interact with others using this device, it takes input from the microphone. Then the device converts speech to text (STT) and displays it on the device screen based on the conversation. The user replies to the text message and the device converts text to speech (TTS) conversion and output are obtained from the speaker.

### 4. IMPLEMENTATION

Hardware components used for implantation are Logitech camera, Raspberry Pi module, microphone and speaker. Raspberry Pi supports external Micro SD card in the range from 8 GB to 128 GB. GPIO pins are digital pins on an IC (integrated circuit), is to connect or control the input and output devices. The DSI (Display-Serial-Interface) is developed by MIPI (Mobile Industry Processor Interface) alliance which is designed to decrease the designing cost of display controllers in wireless devices. Only Raspberry-Pi Model-B and B+ contain the Ethernet port for networking concept. By using this port raspberry pi can able to connect to a network or internet using a standard Rj45 LAN connector on the Ethernet port. Figure 1 shows the translation process from text-to-speech.

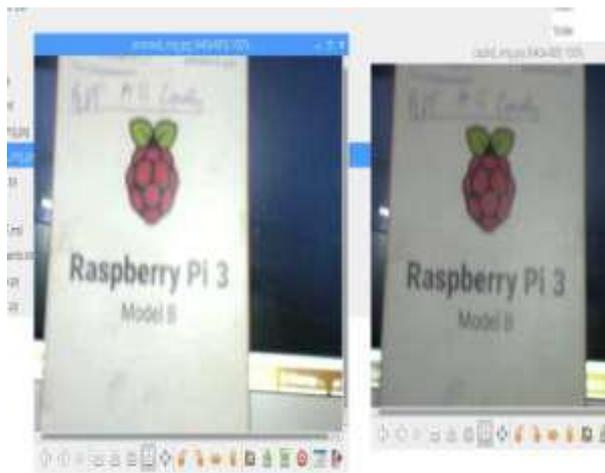
#### 4.1 Tesseract OCR

Tesseract OCR (Online Character Recognition) is the algorithm used to extract text from the picture. Python Tesseract OCR is an online character recognition tool can be used on the various OS platform. Tesseract OCR is the process of electronically extracting text from pictures and reusing the converted text in a variety of ways such as document editing, text-to-voice conversion. OCR is a technology that is capable of converting documents such as scanned papers, PDF files and captured image into editable data. It is used to convert the scanned pictures into text format. It can

be used by programmers to extract typed, printed text from images using an API.

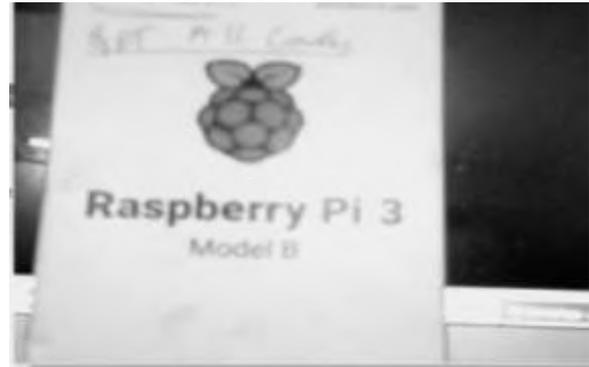
#### 4.2 Camera Initialization

Logitech C270 camera is initialized. If an image from the camera is scaled by a factor, then all the other parameters should be scaled (multiplied/divided, respectively) by the same factor. The matrix of intrinsic parameters does not depend on the scene viewed. So, once estimated, it can be re-used as long as the focal length is fixed (in case of zoom lens). The image is captured from the camera and then processed further. If the captured image is blur, the brightness is optimized for better grasping. Figure 1 depicts the image of enhanced brightness. The optimization of parameter selection is significantly balancing the imperfections of the scale parameters.



**Fig 1: Brightness Enhanced Image**

The brightness enhanced image is converted to grayscale image for better understanding. If image has low brightness, it must be passed through a high-pass filter or else Tesseract OCR's binarization stage will ignore the page and the dark border must be cropped manually to obtain the efficient output. Thresholding method can be useful to read dark text that is overlaid upon gray shapes. After this, it helps reducing the noise, again making it easier for Tesseract to correctly OCR the processed image. Figure 2 shows the grayscale image.



**Fig 2: Grayscale Image**

#### 4.3 Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. The process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image.

#### 4.4 Text Recognition

Detecting text from images is a prototypical modern puzzle that incorporates image processing, computer vision, and machine learning. This is done using OpenCV tool, a highly accurate deep learning text detector used to detect text in natural scene images. And then, it is passed to Tesseract which processes further using python library files. While processing in Tesseract, it calls a number of flags which controls input text and OCR engine mode.

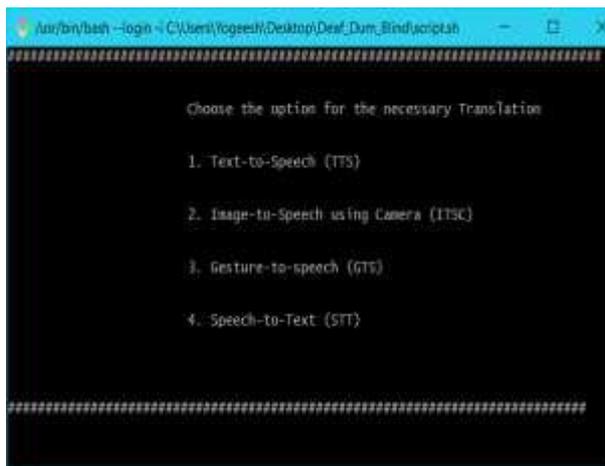
#### 4.5 Converting to mp3 format

The text that is recognized is converted to mp3 format which speaks out the text via a microphone or a speaker. This is done using Google text-to-speech. And the text is finally read out which helps assisting blind people to communicate with the world. gTTS (Google Text-to-Speech) installation is a Python library and CLI instrument to interface with Google Translates content to-discourse API. It composes spoken mp3 information to a document,

a record like item (bytestring) for further sound control, or stdout.

## 5. RESULTS

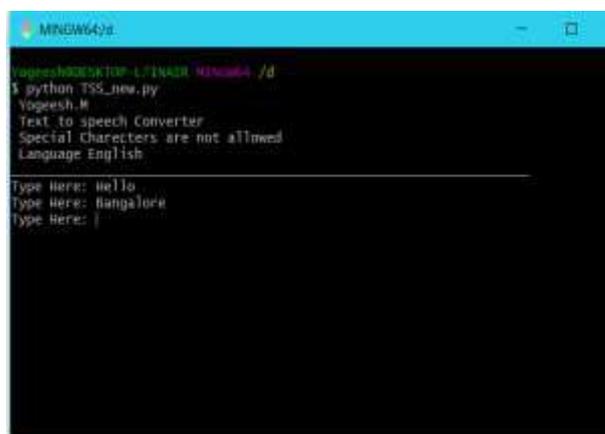
The proposed communication system framework provides a single compact device for a deaf, dumb and blind people to interact with each other easily. Figure 3 shows the system interface for choosing the necessary operation.



**Fig 3: Choices window**

In this proposed system, the VNC viewer is used to display the result remotely (Wireless access/Remote access). VNC is a simple, secure, and ready to use remote access sharing protocol which uses the Remote Frame Buffer Protocol (RFB) communication protocol to remotely control another laptop / a personal computer.

The code was successfully simulated on Windows 10 OS using SpyDer IDE of python 3 Version as shown in fig 4 and Raspberry Pi 3 Model B on python 3.



**Fig 4: Simulation Window**

## 6. ADVANTAGES

Some of the advantages of the proposed system are-

- Eliminates the communication gap
- For dumb, it converts into text and speech
- For deaf, converts speech to text
- For blind, converts text to speech

## 7. APPLICATIONS

- Used in hospitals, in shopping malls, and in public places like bus stations, railway stations, govt. offices, etc. to reduce communication gap
- The device can be used by the vocally-visually-hearing impaired people to communicate with each other as well as with the normal people easily

## 8. CONCLUSION

The proposed system has been designed. The working flow has been analyzed for the process. The TTS involves converting text to speech; STP involves converting speech to text using a microphone and monitor. Gesture control is used for controlling and communication between specially abled people. The implementation is done using Tesseract OCR and eSpeak. The simulation result has been verified.

## 9. REFERENCES

- [1] Kanika Rastogi and Pankaj Bhardwaj, "A Review Paper on smart glove converts gestures into speech and text", International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 4, No. 5, pp. 92-94, 2016.
- [2] K.V Fale, Akshay Phalke, Pratik Chaudhari and Pradeep Jadhav, "Smart Glove: Gesture Vocalizer for Deaf and Dump People", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, No. 4, pp. 801806, 2016.
- [3] Rohit Rastogi, Shashank Mittal and Sajan Agarwal, "A Novel Approach for Communication among Blind, Deaf and Dumb People",

International Conference on Computing for Sustainable Global Development, pp. 605-610, 2015.

[4] Rehman, Salman Afghani, Muhammad Akmal and Rahee Yousaf, "Microcontroller and Sensors Based Gesture Vocalizer", International Conference on signal processing, robotics and automation, pp. 82-87, 2008.

[5] Vikram Sharma M, Vinay Kumar N and D R Ambika, "Virtual Talk for Deaf, Mute, Blind and Normal Humans", International Conference on Texas Instruments India Educators, pp. 316-320, 2013.

[6] Nikhita Praveen, Naveen Karanthand and Megha MS, "Sign Language Interpreter Using a Smart Glove", International conference on advances in Electronics, Computers, and communications, 2014.

[7] Nakul Nagpal, Dr. Arun Mitra and Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 3, No. 5, pp. 147-149, 2015.

*i*Journals