Smart Communication for Deaf and Dumb

Chiranjeevi C R

Dept. of CSE, JSSATEB, Bengaluru, India.

ABSTRACT

With regard to Deaf & Dumb individuals, communication with others is a way longer struggle for them. They are unable to speak with traditional individuals properly. They face difficulties in finding jobs and living a traditional life like others. In this paper, we are introducing a two-way smart communication system for Deaf & Dumb and also for normal people. The system consists of two main parts: The first part is for Deaf & Dumb person to convey their messages to a normal person by using our hardware system and the second one is for a normal person who can also respond them easily without learning a sign language by using our system. This ensures a two-way smart communication system and will make life less demanding for them. The overall accuracy of the system is 92.5%, with both the hands involved. [1]

Key Words: *STT, TTS, e-speak, OCR.*

INTRODUCTION

One of the most precious gifts to a human being is an ability to see, listen, speak and respond according to the situations. But there are some unfortunate ones who are deprived of this. Making a single compact device for people with hearing and vocal impairment is a tough job. Communication between deaf-dumb and normal person have been always a challenging task. This paper proposes an innovative communication system framework for deaf, dumb and people in a single compact device. We provide a technique for a person to read a text and it can be achieved by capturing an image through a camera which converts a text to speech (TTS). It provides a way for the deaf people to read a text by speech to text (STT) conversion technology. Also, it provides a technique for dumb people using text to voice conversion. The system is provided with four switches and each switch has a different function. The dumb people can communicate their message through text which will be read out by e-speak, the deaf people can be able to hear others speech from text. All these functions are implemented by the use of Laptop.[2]
The number of Deaf & Dumb people is over five- hitter of the population. Linguistic communication is principally used by Deaf & Dumb to speak with each other. The most downside today moon-faced by Deaf & Dumb folks is to talk with those that don't understand linguistic communication. In contrast, writing is associate degree possibility; it's thought as a slow and inefficient manner of communication. Thus a viable possibility would be to rent an expert linguistic communication translator. In this paper, we are introducing two- way smart communication system for Deaf & Dumb and Normal people; the project is building a system that assists Deaf & Dumb people to convey their messages to Normal people. The system consists of two main parts: The first part is for Deaf & Dumb person to convey their messages to a normal person and the second one is for a normal person who can also respond them easily without learning a sign language with the help of GUI.

Key Words: *GUI.*

SYSTEM REQUIREMENTS

Camera to capture image:

An RGB image can be viewed as three images (a red scale image, a green scale image and a blue scale image) stacked on top of each other. In MATLAB, an RGB image is basically a M*N*3 array of color pixel, where each color pixel is a triplet which corresponds to red, blue and green color component of RGB image at a specified spatial location. Similarly, A Grayscale image can be viewed as a single layered image.

Key Words: *MATLAB, RGB.*

E-speak synthesizer:

E-Speak is a compact multi-platform multi-language open-source speech synthesizer using a (format synthesis method. Informant synthesis, voice speech (vowels and sonorants consonants) is created by using formants. Unvoiced consonants are created by using prerecorded sounds. Voiced consonants are created as a mixture of a formant- based voiced sound in combination with a pre- recorded unvoiced sound. The E-speak Editor allows to generate formant files for individual vowels and voiced consonants, based on a sequence of key frames which define how the formant peaks (peaks in the frequency spectrum) vary during the sound. A sequence of formant frames can be created with a modified version of Praat, a free scientific computer software package for the analysis of speech in phonetics. The Praat formant frames, saved in a spectrum.dat file, can be converted to formant key frames with E-speak Edit.

Tesseract OCR:
It is an optical character recognition engine for various operating systems. Tesseract up to and including version 2 could only accept TIFF images of simple one-column text as inputs. These early versions did not include layout analysis, and so inputting multi-columned text, images, or equations produced garbled output. Since version 3.00 Tesseract has supported output text formatting, hOCR positional information and page layout analysis. Support for a number of new image formats was added using the Leptonica library. Tesseract can detect whether text is mono spaced orproportionally spaced.

Key Words: *hOCR, Leptonica.*

Speechtexter:
Speechtexter is an online multi-language speech recognizer that can help you type long documents, books, reports, blog posts with your voice. If you need help, please visit our help page at https://www.speechtexter.com/help .This app supports over 60 different languages. For better results use a high-quality microphone, remove any background noise, and speak loudly and clearly. It can create text notes/sms/emails/tweets from users' voice.

Microphone:
Microphone is used to give speech input that wouldbe later converted into text using speech texter so that deaf could read it easily as they cannot hear.[3]

OpenCV:
OpenCV is a library, a cross-platform and is an opensource tool of various programming functions that focus mainly on real-time computer vision. It can be used for various purposes such as face recognition, object identification, mobile robotics, segmentation, gesture-recognition, etc.

METHODOLOGY AND IMPLEMENTATION
The Project is divided into 4 different modules:

1. Text-to-Speech (TTS)
2. Image-to-Speech using camera (ITSE)
3. Gesture-to-Speech (GTS)
4. 4.Speech-to-Text (STT)

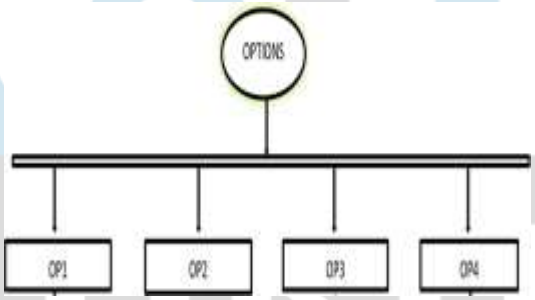

Fig-1: Data flow of system

1. Text-to-speech (TTS)

The first process text to speech conversion is done for the dumb masses who cannot speak. The Dumb people convert their thoughts to text which could be transferred to a voice signal. The converted voice signal is spoken out by E-speak synthesizer. After selecting the option OP1 the OS and sub process imported. Call text to speech function and enter the text as input. After entering the text from keyboard, the E-speak synthesizer converts text to speech. The process also provided with the keyboard interrupt ctrl+C. [4]

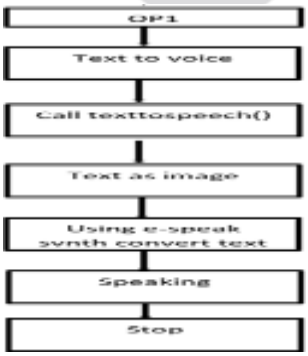

Fig-2: Text-to-Speech

2. Image-to-speech using camera (ITSC)

The second process is developed for dumb people who cannot speak. In order to help them, we have interfaced the Logitech camera to capture the imageby using OPENCV tool. The captured image isconverted to text using Tesseract OCR and save the text to file out.txt. Open the text file and split the paragraph into sentences and save it. In OCR, the adaptive thresholding techniques are used to change the image into binary images and they are transferred to character outlines. The converted textis read out by the E-speak.[5]

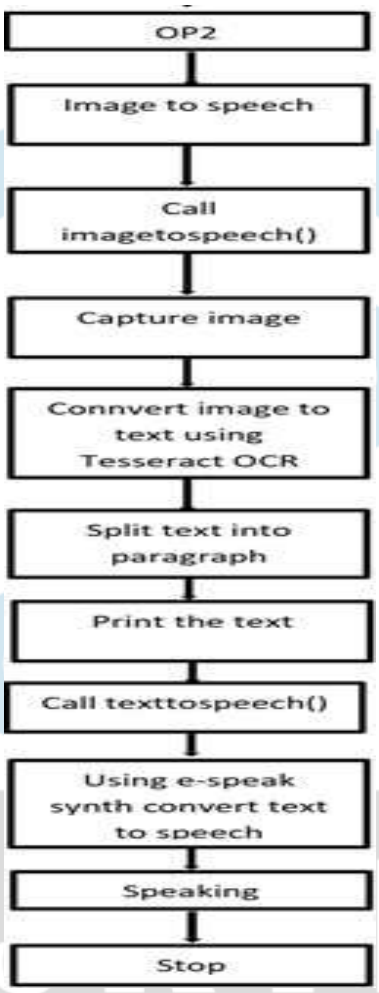

Fig-3: Image-to-Speech

3. Gesture-to-speech (GTS)

The third process is developed for the vocally impaired people who cannot exchange the thoughts to the normal people. Dumb people use gesture to communicate with normal people which are majorly not understandable by normal people. The process starts with the capturing of image and crops the useful portion. Convert the RGB image into gray scale image for better functioning, Blur the cropped image through Gaussian blur function and pass it to the threshold function to get the highlighted part of the image. Find the contoursand an angle between two fingers. By using convexhull function, we can implement the finger point. Count the number of angles which is less than
90 degree which gives the number of defects. According to the number of defects, the text is printed on display and read out by the Speaker.[6]

Key Words: *OpenCV, Gaussian blur, contours.*

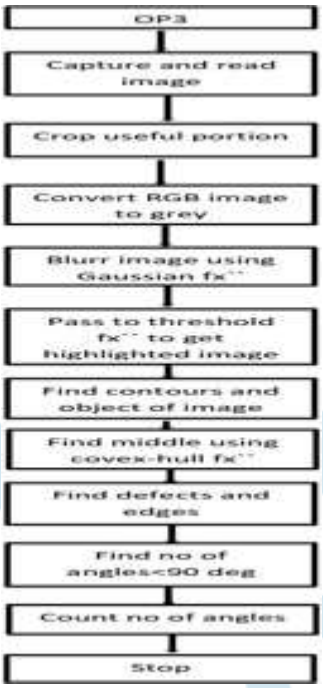

Fig-4: Gesture-to-Speech

4. Speech-to-Text (STT)

The fourth process is developed for the hearingimpairment, people who cannot understand the words of normal people. In order to help them, our project is provided with a switch which is used to convert the voice of the normal people text. We have used a chromium browser which is automatically connected to URL speechtexter.com. The process is performed by assigning a minimum threshold voltage to recognize the voice signal. The input is given through a microphone which is converted into a text format. The URL supports a variety of languages. If the voice signals recognizable it will print the text else it gives the error signal.[7]

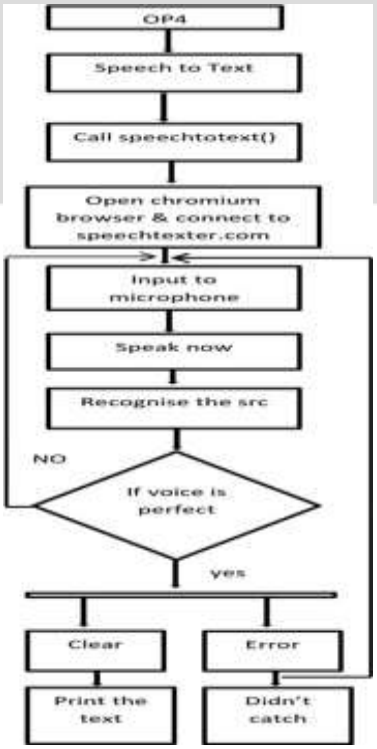

Fig-5: Speech-to-Text

CONCLUSION

This project aims to lower the communication gap between the deaf or mute community and the normal world, help them to lead standard lifestyle. The project is used to convert text/image to voice for dumb, speech to text conversion for deaf and conversion of hand gestures to text for dumb people. We have designed the prototype model for deaf and dumb people into a single compact device. This project can be used as smart assistant for differently abled people to communicate with others and it is a language independent system. It can be further improved by implementing gesture recognition for numbers and alphabets. It can be advanced to take videos as input and segment them into frames from which text readings can be scanned and translated to text or speech format.

REFERENCES

1. Mrs. Neela Harish, Dr. S. Poonguzhali, "Design and development of hand gesture recognition system for speech impaired people", International Conference on Industrial Instrumentation and Control (ICIC) College ofEngineering Pune, India. May 28-30, 2015.

2. Suganya R, Dr.T. Meeradevi, "Design ofA Communication Aid for PhysicallyChallenged", in IEEE Sponsored 2nd International Conference on Electronics and Communication System (Icecs 2015),2015.

3. Meenakshi Panwar and Pawan SinghMehra, "Hand Gesture Recognition for Human Computer Interaction", InternationalConference on Computing, Communication andApplications (ICCCA), Din Digul, Tamilnadu, India, February 2012.

4. https://ieeexplore.ieee.org/document/9080028

5. https://www.researchgate.net/publication/341074410_Twoway_Smart_Communication_System_for_Deaf_Dumb_and_Normal_People

6. https://www.researchgate.net/publication/271484460_Design_of_Smart_eTongue_for_the_physically_challenged_people

7. https://www.researchgate.net/publication/26903 2729_Machine_learning_model_for_sign_language interpretation_using_webcam_images