Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 165 (2019) 259–269

Procedia Computer Science

www.elsevier.com/locate/procedia

INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING 2019, ICRTAC 2019

IoT based Assistive Device for Deaf, Dumb and Blind People

Karmel A*,Anushka Sharma, Muktak pandya, Diksha Garg

Vellore Institute of Technology, Chennai - 600 127, Tamil Nadu, India

Abstract

Focusing and addressing the problems faced by the differently abled people such as visually, audibly and vocally challenged, through a single device is a tough job. A lot of research has been done on each problem and solutions have been proposed separately. But not all of them are addressed together. The aim of the project is to create a single device solution in such a way that is simple, fast, accurate and cost-effective. The main purpose of the device is to make the differently abled people, feel independent confident by seeing, hearing and talking for them. The paper provides a Google API and Raspberry Pi based aid for the blind deaf and dumb people. The proposed device enables visually challenged people to read by taking an image. Further, Image to text conversion and speech synthesis is done, converting it into an audio format that reads out the extracted text translating documents, books and other available materials in daily life. For the audibly challenged, the input is in form of speech taken in by the microphone and recorded audio is then converted into text which is displayed in the form of a pop-up window for the user in the screen of the device. The vocally impaired are aided by taking the input by the user as text through the built-in customized on-screen keyboard where the text is identified, text into speech conversion is done and the speaker gives the speech output. This way the device speaks for the user.

Keywords: Raspberry Pi, Google APIs, Assistive Device, Text to speech, Deaf-Dumb-Blind people

1. Introduction

Approximately 1.3 billion people live with some sort of vision impairment out of which 188.5 million people have a mild vision impairment, 217 million have moderate to severe vision impairment, 36 million people are blind and the majority of people with vision impairment are over the age of 50 years. India is considered to home the largest number of blind people. Around 9.1 billion people are deaf and mute [1]. According to WHO, around 5% of the world's population or 466 million people suffer from disabling hearing loss. Technology is advancing day by day and during the last few decades, it has made our lives easier and convenient. But some- how the physically

impaired part of our society has not been paid enough attention to. They are deprived of the advancements of science and still face plenty of problems in their day to day lives.

*Corresponding Author. Tel.: +91-9840507512
Email Address: karmel.a@vit.ac.in

1877-0509
10.1016/j.procs.2020.01.080

Communication is a major aspect of human lives. But there still exists a gap. Braille and sign language are the means of their communication but it is out of their comfort zone. They always have to learn these traditional modes of communication or they bank on support such as another person. This paper majorly focuses filling this gap by trying to make them feel independent and that they too can walk hand in hand with the other normal people. Raspberry Pi and Google API being the two pillars of this device, make it accurate, efficient and robust. The device consists of three major modules, each dedicated to the visually, audibly and vocally challenged. It uses Raspberry Pi supported by Google API as the main unit, and also consists of a camera, microphone, speaker and a screen. For the visually challenged, the inbuilt camera takes a picture of the writ- ten or printed document and this image is then converted into digital text using the Google Vision API. This text is then converted into audio using the TTS (Text to Speech) library and voice converted output according to the written document or book is obtained. The audibly challenged are aided by recording the speech or audio, converting it into text and displaying it on the screen for the user to read it. The device speaks for the vocally challenged as it provides the user with a customized keyboard on the screen where the user can type the message. This text is converted into speech using the TTS (Text to Speech) library and audio for the input given by the user is obtained in a synthesized voice.

2. Literature

Development of user centered interfaces and technologies have become crucial in the process designing for the differently abled people. Adding an extra element is just not enough to assist the use of technology for the visually disabled [5-7]. Many device-based hardware and software technologies exist to assist the visually disabled. They have functions like reading printed or writ- ten text, expanding characters on braille systems and machines Based on computer vision [3]. Prototypes that function with cell phone, cameras, help in processing images to identify patterns of movement, are applied for musicians who are blind [8,9].

AudioMUD [4] is a multiuser virtual environment exclusively made for the blind people and is associated with spoken cues. The original MUDs (Multi User Dimension) are generally text based and do not contain any sort of graphical interfaces. Users generally use MUD (Multi User Dimension) style games to per- form a set of actions in a virtual environment with a navigable space in the presence of direction, orientation and restrictions. There is high potential for the description of spaces and interactions due to its possible types of interaction and text based interface between players and virtual environment in AudioMUD with collaborative aspects. Their project focuses on the development of a server and client from scratch where the state of the world is stored in the server in such a way that when the server connects to the client, the state of virtual game is received and players can enter or exit anytime. The game starts when the blind user enters the IP name and server in the client, the player comes inside the kingdom of the human body like the respiratory system in a random location with attributes and can explore the system. L.Gonzalez et al [2] suggest a system for the visually disabled to enhance the quality of their life. The wearable system consists of facial recognition to recognize people's faces and can identify a person through prior system training using fisher faces algorithm, obstacles detection where the user wears the device which uses ultrasonic sensors to generate vibration signals that indicate an obstacle, email reader which accesses user's email using POP3 protocol and enables the user to listen to the email using headphones, medication reminder to remind the user about the medication prescribed, MP3 player as a source of entertainment enabling the user to listen to music. Anusha Bhargava et al [3] suggest a system using raspberry pi that uses image acquisition using interfacing a webcam, preprocessing of image to obtain the region of interest, template identification to detect characters and objects, converting image to text using OCR algorithm which scans image and gives a corresponding text output, and save the text data in a text file, and convert text to speech using E-speak for the blind user to hear the text.

Sign language which principally uses manual communication including hand movements, facial expressions to express, connects with people and convey their messages. Lorenzo Monti et al [10] have come up with a wearable device for the deaf-blind users called GlovePi to identify the person, number and position of people, and their facial expressions in front of the user. It mainly comprises of a gardener glove which is attached to capacitive touch sensor with raspberry Pi using a I2C interface. Using many to many architecture in order to include maximum amount of users into an account, the Glove enables the user to register on the server usually by sending a HTTP request and eventually the user is added on the server after which the server sends a updated list of all the connected users and thus uses peer to peer communication to send or receive messages. Amro Mukhtar Hussain et al [12] has designed a mouth gesture recognition system using the help of an infrared sensor that collects the data from the audibly impaired person's mouth and detects the state of the mouth. They have designed three states: OSCS (Open Slow Close Slow), OSCF (Open Slow Close Fast) and OFCS (Open Fast Close Slow). When the sensor reaches its threshold, the sensor indicates and records the signal. Using different combinations, 27 patterns have been achieved which generated 26 alphabetic letters. The output of this proposed system depends on the light reflected from the object that the sensor subjected on, where the intensity supposedly gets affected by the surface color, shape and distance, after which the circuit gets the appropriate output analog voltage range.

Systems that suffice all solutions for the blind, deaf and dumb users in one compact device are rare to find. Kumar.K et al [1] have introduced an arrangement for the visually impaired can understand words using Tessaract which is an OCR (Optical Character Recognition) algorithm by python, vocally impaired can ex- press and communicate by text which is read through E-speak, and audibly impaired can hear by speech to text conversion using OpenCV. Rohit Rastogi et al [11] have put up an ideology that consists of a Sharon bridge which is a wearable technology that makes communication between differently abled on the extent of their capabilities. The Sharon Bridge comprises of small units to form a complete circuit to enable them to convey messages among the differently abled and their different combinations. It comprises of a sensor glove that is made up of arduino circuit board, tactile and flex sensors, and accelerometer which is used to convert the American sign language to audio that is further changed to text which is displayed on the LCD(Liquid Crystal Display) for the user, Arduino GSM(Global System for Mobile communications) shield to communicate over long distances using the internet and GPRS(General Packet Radio Service) wire- less network, Beagle bone that converts analog to digital and vice versa. It works in a way where the message to be sent is the input as text, audio or braille language which is converted to the respective forms for the disabled to hear, speak or see. For long distances, the input in converted and sent through wireless GSM network to the receiver but the user is supposed to possess a phone number. Sharon Bridge works for all combinations of the blind, deaf and dumb.

3. Design

The figure 1 shows the outline of the device. The raspberry Pi is the support system of the device which connects the camera, microphone, speaker and LCD display. The device works for the visually impaired as the camera clicks a picture of the document and the output is in audio format through the speaker, audibly impaired as the microphone takes the spoken words as input and displays it as text on the LCD display, and for the vocally impaired as the user types the message in the LCD and the speakers gives the output as an audio.

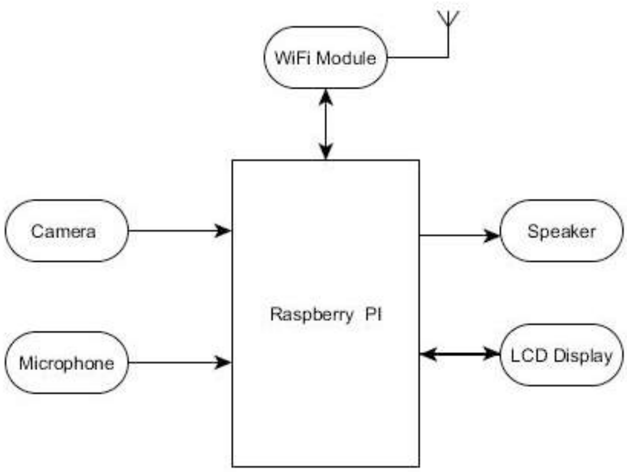

Figure 1 The system architecture with all main modules

3.1 Software Requirements

3.1.1 Google Cloud Vision API

The Google Cloud Vision API (Application Programming Interfacing) encapsulates powerful machine learning models in an easy to use REST API and enables developers and users to apprehend the content of an image. It is used for classification of images into thousands of categories, detecting individual objects and faces within images, and reading printed words contained within images. Optical Character Recognition (OCR) is used to enable the user to detect text within images, along with automatic language identification. Vision API supports a huge and broad set of languages. Initially Conventional neural network (CNN) based model is used to detect localized lines of text and generates a set of bounding boxes. Script identification is done by identifying script per bounding box and there is one script per box. Text recognition is the core part of the OCR which recognizes text from image as shown in Figure 2.

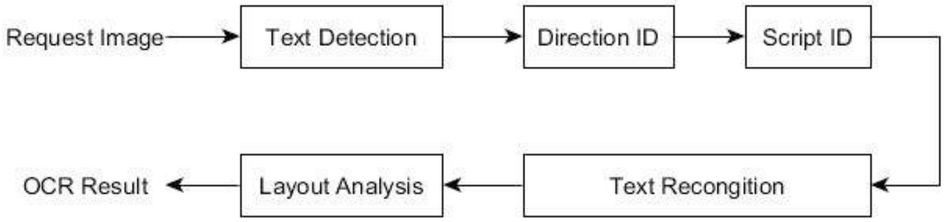

Figure 2 Architecture of Vision API converting Image to text

3.1.1 Tkinter

Various options for the development of graphical user interfaces are provided by python. Tkinter is the standard GUI (graphical user interface) provided as a library for python. GUI applications can be created in a faster and easier way using Tkinter, and it also provides a prevailing object-oriented interface to the Tk GUI toolkit.

3.1.2 Speech to Text:

Google cloud Speech to text aides the developers in the conversion of audio into text as it applies robust neural network models in a convenient API. It enables voice command and control and transcribes audio. It is capable of processing real-time streaming or pre-recorded audio using Google's ML technology. The accuracy is unparalleled as the most advanced deep learning neural network algorithms are applied by Google. It streams text results, returning text as it is recognized from audio stored in a file and is capable of long-form audio as shown in Figure 3.

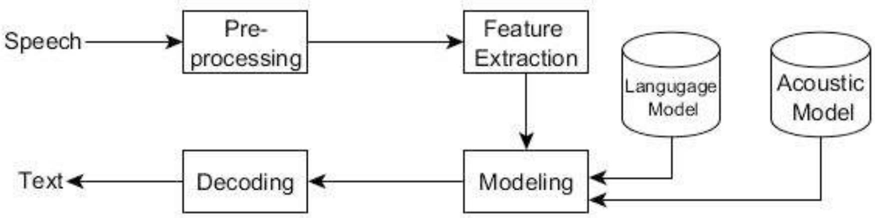

Figure 3 Google speech API converting speech to text

3.1.3 Text to speech

Google Text to Speech API is one of the several APIs available in python to convert text to speech as shown in Figure 4. It is commonly known as the gTTS API. It is an easy and efficient tool which converts entered text, into audio that can be saved as an mp3 file.

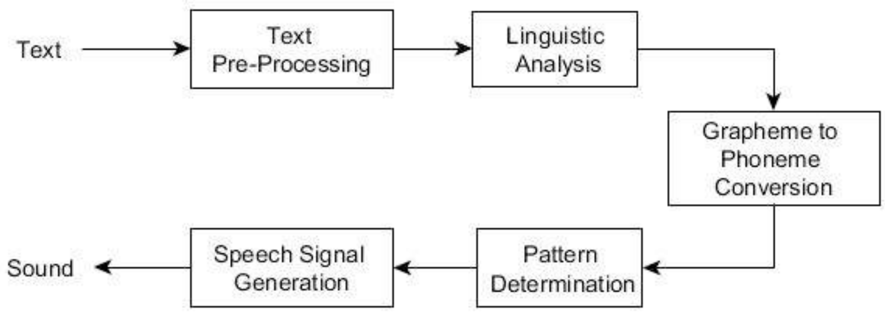

Figure 4 gTTS converting text to speech

3.1.4 Bitwise SSH

Bitwise SSH (Secure Shell) is one of the advanced and flexible SFTP protocol. The bitwise ssh helps us to securely connect with raspberry pi and access all the resources of raspberry pi. In addition, the user can transfer the files from localhost to raspberry pi; compile the programs; and provides a secure link for further connection.

3.2 Hardware Requirements

3.2.1 Raspberry Pi

Raspberry Pi, shown in Figure 5 is a low cost, credit card sized processor, which can easily perform all task we expect from a desktop. It is very easy to connect raspberry pi with computers and TVs. It also provides GPIO (General Purpose Input Output) pins to connect with other components. Because of this efficiency to intercommunicate with the cross-disciplinary domain, it has been used in a variety of projects. Raspberry pi operates in an open source environment such as Raspbian (Linux based operating system).

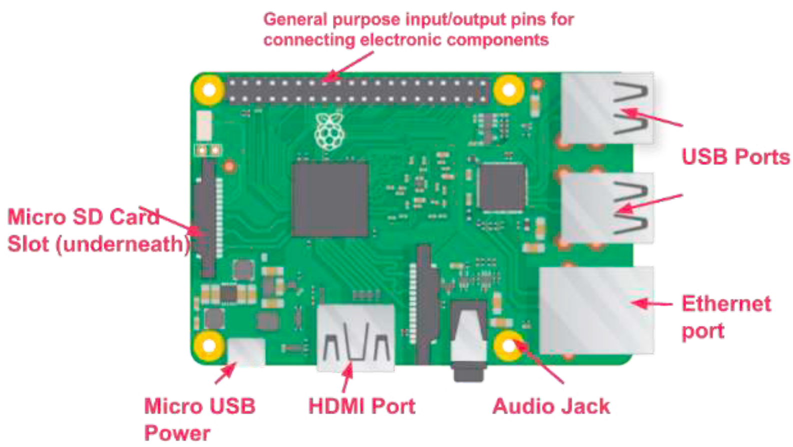

Figure 5 Raspberry Pi

Technical specification

- Broadcom Soc BCM2836 (CPU, GPU, DSP, SDRAM)
- 900 MHz quad-core ARM Cortex A7 CPU (ARMv7 instruction set)
- Broadcom VideoCore IV @ 250 MHz GPU
- 1 GB MEMORY (shared with GPU)
- 4 USB ports
- 17 GPIO(General Purpose Input Output) Peripherals plus specific functions, and HAT ID bus
- 15-pin MIPI camera interface (CSI) Video input connector
- HDMI video outputs, composite video (PAL and NTSC) via 3.5 mm jack
- I^2S audio input
- Analog audio output via 3.5 mm jack; digital via HDMI and I^2S
- MicroSD for storage
- 10/100Mbps Ethernet speed
- 800 mA power rating (4.0 W)
- 5 V power source via MicroUSB or GPIO(General Purpose Input Output) header
- 85.60mm × 56.5mm
- Weighs 45g (1.6 oz)

3.2.2 Camera Module

The camera used by project is a C310HD Logitech webcam, shown in Figure 6 with a resolution of 720p/30fps. The images taken are crisp and contrasted. This camera fits perfectly in the project as it adjusts to the lighting conditions to produce brighter contrasted images. It uses a universal clip to attach itself firmly to the device. It is small, adjustable and agile and is therefore handy in the project.

Technical specifications

- Max Resolution: 720p/30fps
- Lens technology: standard
- Focus type: fixed focus
- Field of View: 60°
- Built-in mic: mono
- Cable length:1.5 m
- Universal clip fits laptops, LCD or monitors

Figure 6 Logitech C310HD Webcam

3.2.3 Screen display

The project consists of a 5inch resistive touch with a high hard- ware resolution and HDMI Interface specially designed for the Raspberry Pi, shown in Figure 7. It has a resistive touch control It is compatible and has a direct connects with any revision of the existing Raspberry Pi. It provides drivers and the backlight can be turned on or off for the lower power consumption. According to the requirements of the project, a keyboard has been hardwired in this 5-inch dis- play for the vocally challenged to type their text in the screen.

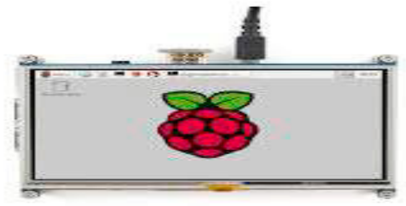

Figure 7 Waveshare 5' inch display

Technical Specification

- Drivers provided (works with your own Raspbian/Ubuntu/Kali/Retropie)
- HDMI interface for displaying, no I/Os required (however, the touch panel still needs I/Os)
- High-quality immersion gold surface plating

3.2.4 Microphone

A mini portable high quality USB microphone, shown in Figure 8 is used in the project. It is a noise cancelling microphone which filters out un- wanted background noise. It comes as a brownie point for the project as it is portable, compact and easy to use. It can be made more efficient according to the user or background by increasing the gain control or capture for better accuracy.

Technical Specification:

- 4.5V Working voltage
- Weight 99.8 g
- 2cm x 2cm x 0.5cm in size

Figure 8 USB Microphone

4. Implementation

The device has been created by formulating a unique design for assisting the differently abled people. It has been divided into three modules for enhancing the experience of the user with the device. The device consists of three modes and a three-way slider to change mode. Each mode is separately dedicated for the blind, deaf and dumb respectively in the device. The device is designed to make the user feel individualistic, self-reliant and self-sufficient. The gist of design of the device is in the figure 1. The main component of the device is the raspberry Pi.

4.4.1 Blind Module

The figure 9 represents the methodology of this module which consists of three steps.

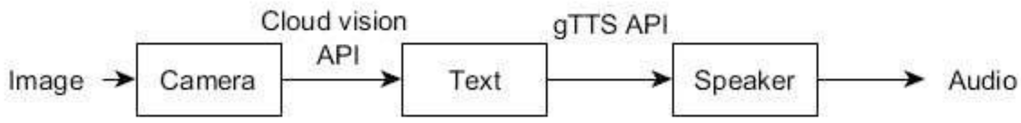

Figure 9 Working of Blind Module

Step 1. For the module to work, the three-way slider is set to the blind mode. The camera connected to the raspberry Pi of the device takes a picture of the written document or book placed on the holder of the device.

Step 2. The picture is saved in JPEG format and is passed to the Google Cloud Vision API to be converted to text where the API extracts the text to be converted.

Step 3. The extracted text then gets converted into speech using the gTTS API and the required text is thus converted to the audio format.

Step 4. This audio is given as an output by the high quality speaker connected to the Raspberry Pi and thus the device enables the visually impaired person to understand the written document or book through the audio.

4.4.2 Deaf module

The audibly impaired can virtually hear using this device as it enables them to read, what is being spoken. The figure 10 describes the respective procedure.

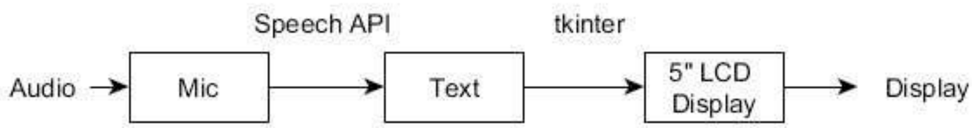

Figure 10 Working of Deaf Module

Step 1.The three-way slider is set to the deaf mode. The audio or the words being spoken to the user, who in this case might be a deaf person, are recorded as input by the USB Microphone connected to the Raspberry Pi of the device and is saved as a file in mp3 format.

Step *2*.This audio file is passed to the Google Speech API which converts the audio into text for the user to understand.

Step 3.The converted text is then displayed on the 5 inch HDMI LCD screen available in the device, as a pop up window exclusively created using python tkinter for this module. This way the user understands everything that is being spoken to him quickly and efficiently. To change modes, the slider can be set accordingly.

4.4.3 Dumb module

This module makes the device handy for the vocally disabled as it enables them to vocalise words by typing it on the screen. The figure 11 explains the methodology.

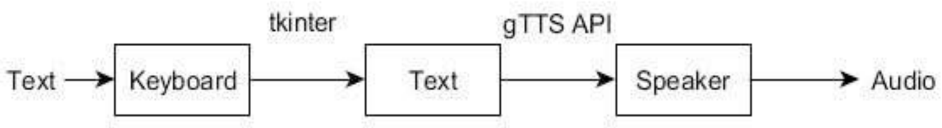

Figure 11 Working of Dumb Module

Step 1.When the three-way slider is used to set the device on the dumb mode, a pop up is displayed along with a customized keyboard which has been created using python tkinter, in it on the HDMI screen connected to the Raspberry Pi.

Step 2.The user who possibly is vocally impaired can type whatever he wants to convey using the keyboard in the screen as text

Step 3.The typed text is converted into audio format using the gTTS API and the audio file of the required text is obtained.

Step 4.The high quality speaker connected to the Raspberry Pi in the device plays this audio file thus vocalising the message given by the impaired person.

Step 5.Modes can be changed in the device according to the convenience of the user.

5. Results

The figure 12 shows the working of the blind module as it reads the text taken as a picture in figure 13 and visually impaired will hear this text through the speakers.

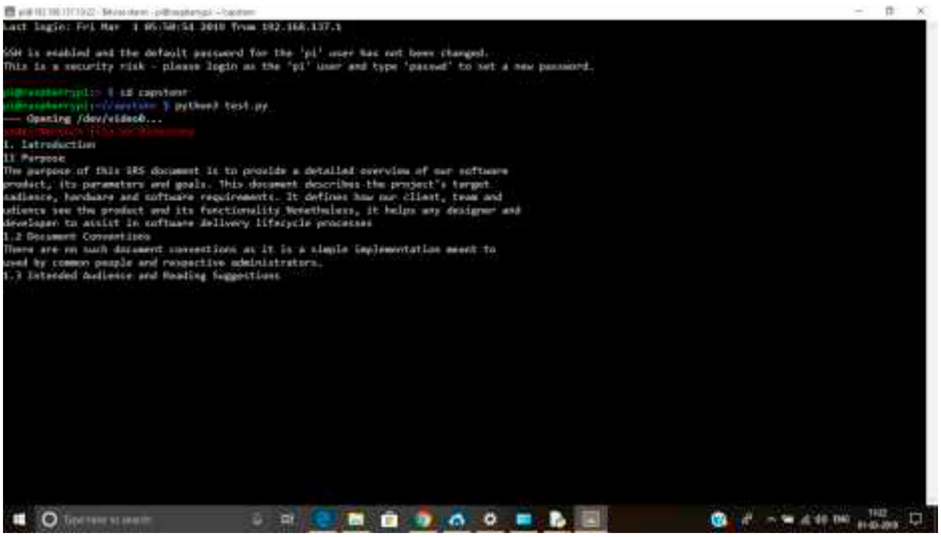

Figure 12. Conversion of image to text using Google Vision API

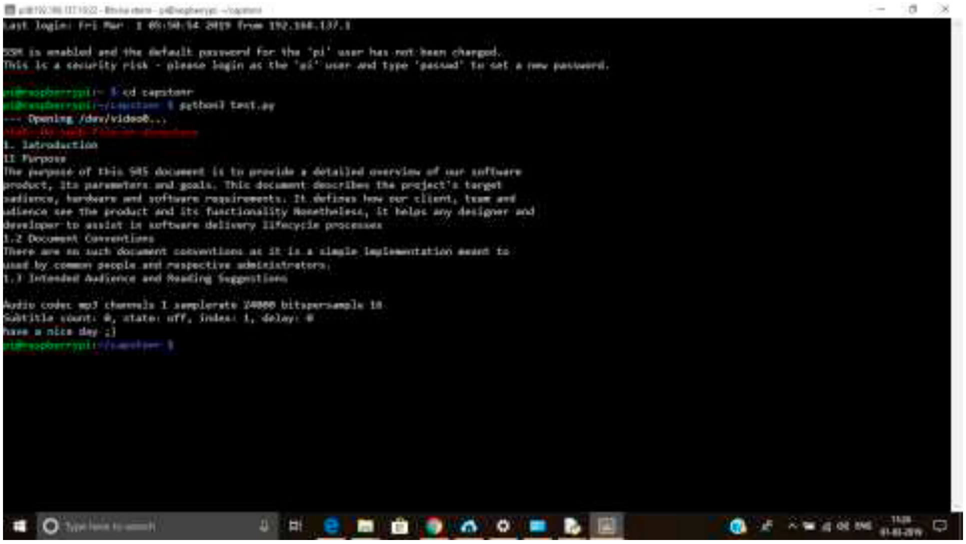

Figure 13. Conversion of gained text from image to Audio

Figure 14 shows how the audibly impaired can read as the audio or spoken words “Muktak doing testing for module 2” are identified as and converted to text.

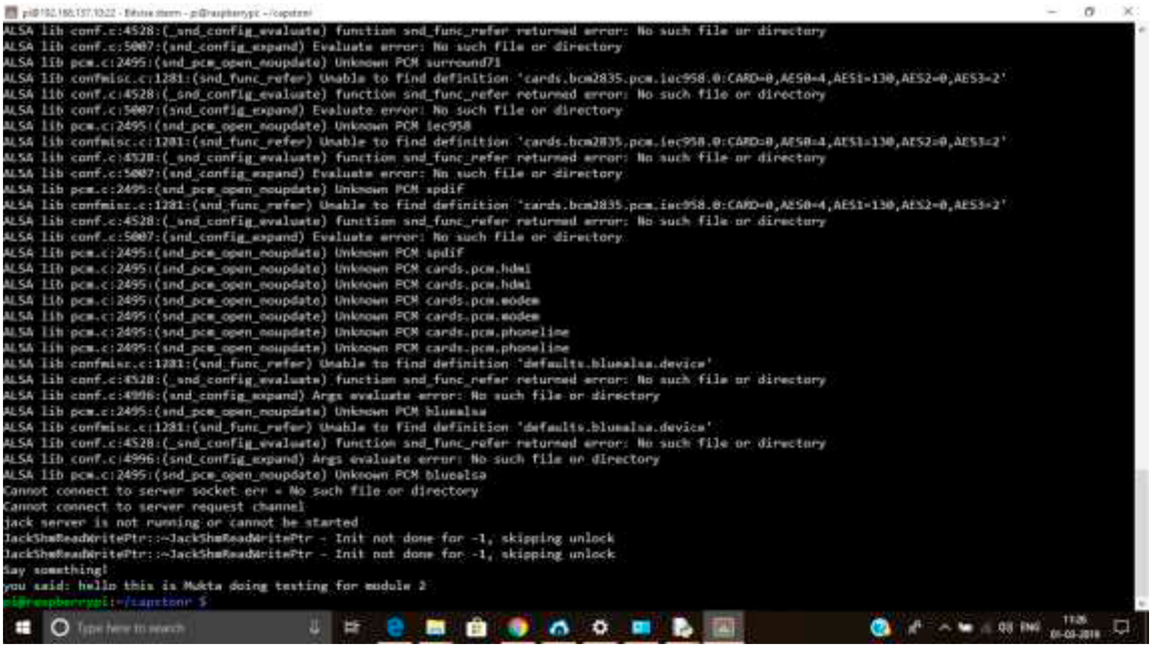

Figure 14. Conversion of Audio to text

Vocally impaired type a message in the keyboard of the screen as shown is figure 15 and this text gets converted to speech which the speaker gives as an output.

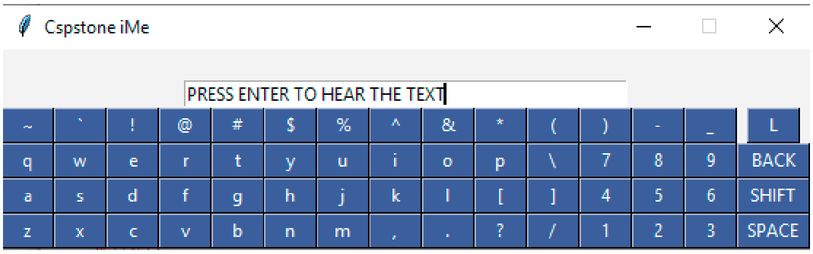

Figure 15 On-screen Keyboard (text to speech)

6. Conclusion

Through this paper, an unprecedented prototype has been created to aid the visually, vocally and audibly disabled. This project not just focuses on empowering and facilitating the differently abled, it is also compact and resource saver. The overall cost has been cut down by eliminating braille books and the energy spent in understanding them. It is a less costly solution, as all the components used in the device are cost effective and efficient. The latest and most trending technology makes this device portable, adaptable and convenient. The device proposed in this paper can be a major help in solving a few of the many challenges faced by the differently abled. To further extend the project, the device can be made more compact and wearable to make it easy for the user to use.

References

[1] N. K., S. P. and S. K., Assistive Device for Blind, Deaf and Dumb People using Raspberry-pi, Imperial Journal of Interdisciplinary Research (IJIR), 3(6), 2017 [Online]. Available: https://www.onlinejournal.in/IJIRV3I6/048.pdf.

[2] L. González-Delgado, L. Serpa-Andrade, K. Calle-Urgiléz, A. Guzhñay-Lucero, V. Robles-Bykbaev and M. Mena-Salcedo, "A low-cost wearable support system for visually disabled people," 2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, 2016, pp. 1-5. doi: 10.1109/ROPEC.2016.7830606

[3] Anusha Bhargava, Karthik V. Nath, Pritish Sachdeva & Monil Samel (2015), International Journal of Current Engineering and Technology, E-ISSN 2277– 4106, P-ISSN 2347– 5161

[4] J. Sanchez and T. Hassler, "AudioMUD: A Multiuser Virtual Environment for Blind People," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, 15(1), pp. 16-22, March 2007. doi: 10.1109/TNSRE.2007.891404

[5] M. Lumbreras and J. Sánchez, "Interactive 3-D sound hyperstories for blind children," in Proc. ACM-CHI '99, Pittsburgh, PA, 1999, pp. 318–325.

[6] R. McCrindle and D. Symons, "Audio space invaders," in Proc. ICDVRAT 2000, Alghero, Sadinia, Italy, Sep. 23–25, 2000, pp. 59–65.

[7] T. Westin, "Game accessibility case study: Terraformers-Real-time 3-D graphic game," in Proc. ICDVRAT 2004, Oxford, UK, 2004, pp. 120–128.

[8] Y. H. Lee and G. Medioni, "Rgb-d camera based wearable navigation system for the visually impaired," Computer Vision and Image Understanding, vol. 149, pp. 3–20, 2016

[9] J. Bajo, M. A. Sanchez, V. Alonso, R. Berj ´ on, J. A. Fraile, and J. M. ´ Corchado, "A distributed architecture for facilitating the integration of blind musicians in symphonic orchestras," Expert Systems with Applications, 37(12), pp. 8508–8515, 2010.

[10] L. Monti and G. Delnevo, "On improving GlovePi: Towards a many-to-many communication among deaf-blind users," 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, 2018, pp. 1-5.doi: 10.1109/CCNC.2018.8319236

[11] R. Rastogi, S. Mittal and S. Agarwal, "A novel approach for communication among Blind, Deaf and Dumb people," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2015, pp. 605-610.

[12] A. M. Hassan, A. H. Bushra, O. A. Hamed and L. M. Ahmed, "Designing a verbal deaf talker system using mouth gestures," 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Khartoum, 2018, pp. 1-4. doi: 10.1109/ICCCEEE.2018.8515838