International Journal of Research in Engineering, Science and Management
Volume-2, Issue-3, March-2019
www.ijresm.com | ISSN (Online): 2581-5792

Raspberry PI based Reader for Blind Person

Manisha S. Thakare[1], Akshay B. Matte[2], Pawan D. Kelzarkar[3], Hemant P. Chavan[4]
[1,2,3]*Student, Department of Electronics, Government College of Engineering, Chandrapur, India*
[4]*Professor, Department of Electronics, Government College of Engineering, Chandrapur, India*

***Abstract*: Character Recognition has become main aspect of Computer vision. Optical Character Recognition is a method where characters are recognized from images digitally. In this project an innovative, efficient and real-time cost beneficial technique that enables user to hear the contents of text images instead of reading through them has been introduced. Text-to-Speech converter is a device that scans and reads English alphabets in the image using OCR technique and changing it to voices. It combines the concept of Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry pi. This kind of system helps visually impaired people to interact with computers effectively through vocal interface. Text Extraction from color images is a challenging task in computer vision. This project describes the design, implementation and experimental results of the device. This device consists of two modules, image processing module and voice processing module.**

***Keywords*: Text Extraction, OCR, Text-To-Speech, Image Processing, Voice Processing**

1. Introduction

Language is the ability to express one's thoughts by means of a set of signs, whether graphical gestural, acoustic, or even musical. It is distinct venture of human beings, who are the only creatures to use such a structured system. Speech is one of its main components. It is by far the oldest means of communication between human being and also the most widely used. No wonder, then, that people have extensively studied it and often tried to build machines to handle it in acoustic way. Most of the Information in digital world is accessible to a few who can read or understand a particular language. Language technologies can provide solutions in the form of natural interfaces so the digital content can reach to the masses and facilitate the exchange of information across different people speaking different languages.

In this project we are taking images as an input. To extract text from the image we need to do image processing. We are implementing some concepts from curricular subject Digital Image Processing. We are doing Python programming for Raspberry Pi under Embedded Platform. Here in this project we are working on Digital Image Processing, this subject is from our syllabus and we have learned about this subject very much. And after learning we are able to perform some basic operations on image. So in this project we have used some of the concepts of digital image processing like segmentation, grey scale conversion etc.

A text to speech converter convert text into speech. Here question arises that whether machine or simply computer can perform same task of text to speech conversion? Answer is not that much easily as human can. The machine has to follow some procedure which is divided in basic two steps: I: character recognition in which we are using OCR that is optical character recognition method. We are implementing for English alphabets. Next step is TTS that is Text to speech conversion in this we have to convert recognized text from OCR into .wav file or simply in speech file.

Optical character recognition (also optical character reader, OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast). [1] Testing of device was done on Raspberry Pi platform device. The Raspberry Pi is initially connected to the internet through Wi-Fi. The software is installed using command lines. The first setup is to download the installation script, second command is to convert it to executable form and the last command starts the script which does the rest of the installation work. Device set up is done as in Fig. 1.

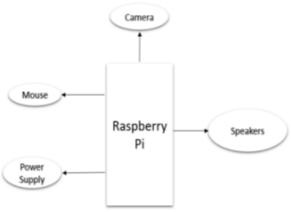

Fig. 1. Raspberry pi setup design model

The webcam is manually focused towards the text. Then, to take a picture, use a Mouse, as an input device. The Mouse set up is done as in fig.1. A delay of around 7 seconds is provided, which helps to focus the webcam, if it is accidently defocused. After delay, picture is taken and processed by Raspberry Pi to hear the spoken words of the text through the earphone or speaker plugged into Raspberry Pi through its audio jack

2. Related work

Visual impairment is one of the biggest limitation for humanity, especially in this day and age when information is

communicated a lot by text messages (electronic and paper based) rather than voice. The device we have proposed aims to help people with visual impairment. In this project, we developed a device that converts an image's text to speech. The basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi's audio jack using speakers or earphones [3].

Optical character Recognition (OCR) is a conversion of scanned or printed text images, handwritten text into editable text for further processing. In this paper, we have presented a robust approach for text extraction and convert it to speech. Testing of device was done on raspberry pi platform. The Raspi is initially connected to the internet through VLAN. The software is installed using command lines. The first setup is to download the installation script, second command is to convert it to executable form and the last command starts the script which does the rest of the installation work. Device set up is done as in Fig.1 The webcam is manually focused towards the text [5]. Then, to take a picture, press pushbutton switch. Pushbutton switch set up is done as in fig A delay of around 7 seconds is provided, which helps to focus the webcam, if it is accidently defocused. After delay, picture is taken and processed by Raspi to hear the spoken words of the text through the earphone or speaker plugged into Raspi through its audio jack.

3. System design

A. *Raspberry Pi camera module*

The camera module is connected with the camera serial interface of the raspberry pi using the 15-pin ribbon cable. Enable the camera support in the configurations. This helps in capturing a 5 MP resolution image by a single command. The command is: sudo raspistill-o image.jpg The Fig. shows a raspberry pi 3 model B with a raspberry pi camera module connected via a 15-pin ribbon cable. Fig. 2.

Fig. 2. Raspberry pi board and camera module

B. *Raspberry Pi 3 Model B*

The Raspberry Pi 3 Model B is the third generation Raspberry Pi. This powerful credit-card sized single board computer can be used for many applications and supersedes the original Raspberry Pi Model B+ and Raspberry Pi 2 Model B. Whilst maintaining the popular board format the Raspberry Pi 3 Model B brings you a more powerful processer, 10x faster than the first generation Raspberry Pi. Additionally, it adds wireless LAN & Bluetooth connectivity making it the ideal solution for powerful connected designs [11].

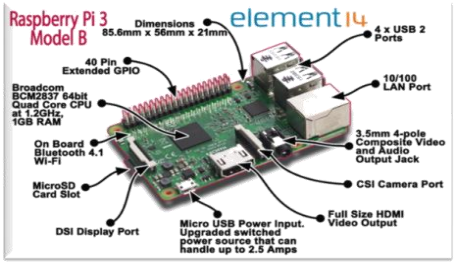

Fig. 3. Raspberry Pi Module

C. *Impedance*

Impedance = Resistance. Measured in ohms, this is another critical specification that you should take a note of. To understand what impedance means, think of a speaker as a hose and water that flows through it as electric current (audio signals). [It is the current (that flows) that is converted into the sound produced by a speaker.] If a hose has a narrow opening (high resistance), less water will flow through it, and if the hose has a bigger opening (low resistance), more water will flow through it [9].

D. *Frequency response*

It is sometimes mentioned as frequency range and is measured in Hertz (Hz). This specification in a speaker tells how low and high a speaker can play. It is mentioned as "XHz-XkHz" - where X is a number. Let's take a hypothetical figure: 65Hz-20kHz. The number at the lower end i.e.,"65Hz" here represents the bass output which means how low a speaker can play. The lower the number, the deeper the bass. And 20kHz (20,000 Hz) represents the highest treble. It is said that the human ear can hear between 20Hz and 20kHz. But, practically, bass frequencies below 30Hz are less heard and more felt [9].

4. Software requirement

A. *Python IDE*

Python is a multi-paradigm programming language: object-oriented programming and structured programming are fully supported, and many language features support functional programming and aspect-oriented programming (including by metaprogramming and metaobjects (magic methods)). Many other paradigms are supported via extensions, including design by contract and logic programming. Python uses dynamic typing and a mix of reference counting and a cycle-detecting

garbage collector for memory management. An important feature of Python is dynamic name resolution (late binding), which binds method and variable names during program execution [12].

5. Conclusion

In this paper, it was converting an image to text conversion in which it was extracting the characters from image. And another one was converting text to speech. No one has ever think about combining these two projects together so it will be useful for blind people. We have combine these two projects so it will be useful for blind people. We have introduced optical character recognition to extract text from images and used festival library to convert it into speech. Text-to-Speech device can change the text in image into sound. This portable device can be used by visually impaired people to read boards, books, computer screens etc. This device does not require internet connection and can be used independently. Through this method, we can make editing process of books or web pages easier.

References

[1] Archana A. Shinde, D. "Text Pre-processing and Text Segmentation for OCR. International Journal of Computer Science Engineering and Technology", pp. 810-812, 2012.
[2] R. Mithe, S. Indalkar and N. Divekar." Optical Character Recognition" International Journal of Recent Technology and Engineering (IJRTE)", ISSN: 2277-3878, Volume-2, Issue-1, March 2013.
[3] R. Smith, "An Overview of the Tesseract OCR Engine", USA: Google Inc."
[4] H. Shah and A. Shah. "Optical Character Recognition of Gujarati Numerical", International Conference on Signals, Systems & Automation, pp. 49-53, 2009.
[5] Simon Monk, "Raspberry Pi Cook Book."
[6] R. Sachdeva, "Text Localization and Extraction in Images Using Mathematical Morphology and OCR Techniques," 2013.