

# Image Text to Speech Conversion Using OCR Technique in Raspberry Pi

Jarlin James, Michael Aldo, Adellina Andrew

**Abstract**— Raspberry Pi is a single board, high performance computer. We are using this system to extract text from a document and convert it to audio for the use of visually impaired persons. Through this device, we want to enable our users to hear the contents of the text image. Concepts of Optical Character Recognition and Text to Speech Synthesis has been combined in this project.

**Index Terms**— Google Speech API, Optical Character Recognition, Raspberry Pi, Raspberry Pi Camera, Tesseract, Visually Impaired,

## 1 INTRODUCTION

Vision impairment refers to a decreased ability to see to a degree it causes problems not fixable by usual means, such as spectacles. The term blindness is used for complete or nearly complete vision loss. Most visually impaired people who are not totally blind are able to read print, either of regular size or enlarged magnification devices. Others use Braille, or some use reading devices which convert printed text to speech or Braille. Access technology such as screen readers, screen magnifiers and refreshable Braille displays enable the blind to use mainstream computer applications and mobile ones. Modern technology has made many useful tools for people who read and write Braille. There are some devices that produce books in Braille and others that let people read information on computers and from the Internet.

## 2 PROJECT OVERVIEW

Optical Character Recognition is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine coded text, whether from a scanned document, a photo of a document, or from subtitle text superimposed on an image. The basic function of OCR Technology is to automatically capture printed text present in scanned images and convert it into a text searchable document.

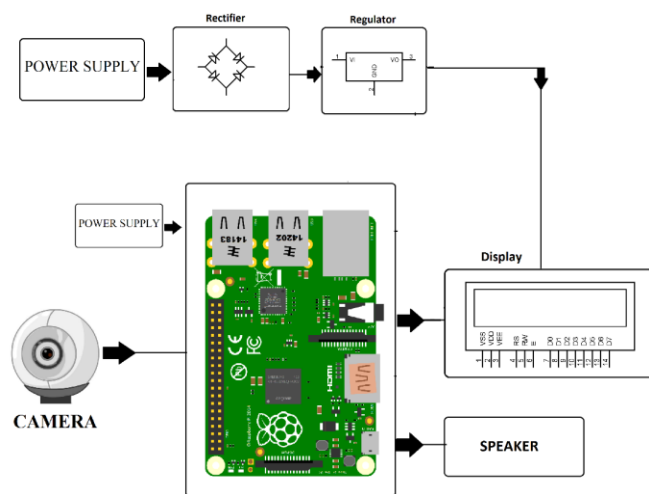
A text to speech (TTS) synthesizer is a computer based system that reads text aloud automatically, regardless of whether the text is introduced to the system that can read text aloud automatically, regardless of whether the text is introduced by a computer input stream or a scanned input submitted to be implemented by both hardware and software. Speech is often based on concatenation of natural speech i.e. Units that are taken from natural speech altogether to form a word or sentence.

## 3 SOFTWARE SPECIFICATIONS

- Operating System: Raspbian
- Language: Python
- Platform: OpenCV
- Library: OCR engine, TTS engine

## 4 HARDWARE SPECIFICATIONS

- Raspberry Pi 3 (Model B)
- Broadcom BCM2837 processor
- LCD Display
- Capacitors
- Resistors
- Battery
- Speaker



## 5 FLOW OF PROCESS

**Image processing:** In the first step the device is moved over the printed page and the inbuilt camera captures the images of the text. The quality of the image captured will be so high so as to have fast and clear recognition due to the high resolution camera. Letters will be extracted and converted into digital form.

**Preprocessing:** It consists of three steps: Skew Correction, Linearization and Noise Removal. The captured image is checked for skewing. There are possibilities of image getting skewed with either left or right orientation. Here the image is first brightened and binarized. The function for skew detection checks for an angle of orientation and if detected then a simple image rotation is carried out till the lines match with the horizontal axis, which produced a skew corrected image. The noise introduced during capturing or due to poor quality of the page has to be cleared for further processing.

**Segmentation:** This operation seeks to decompose an image of sequence of characters into sub image of individual symbol (characters). The binarized image is checked for interline spaces. If inter line spaces are detected then the image is segmented into set of paragraphs across the interline gap. The lines in the paragraphs are scanned for horizontal space intersection with respect to the background. Histogram of the image is used to detect the width of the horizontal lines. Then the lines are scanned vertically for vertical space intersection. Here histograms are used to detect the width of the words. Then the words are decomposed into character width computation.

**Feature Extraction:** In this stage we gather the essential features of the image called feature maps. One such method is to detect the edge in the image, as they will contain the required text.

**Tesseract:** This software is used to convert the image file to text file by extracting the texts from the image and storing it in the file with .txt extension.

**Google Text to Speech API:** This is used to convert the output content into voice which can be heard on the speakers.

#### 4 CONCLUSION

Through the implementation of this system, visually impaired persons can easily listen to the text of the document. Through translation tools, one can convert the text to the desired language and then again by using the Google speech recognition tool he/she can convert that changed text into audio. We can also extend the scope for the long distance capturing. The capture work could also be concentrated on developing an efficient portable.

#### REFERENCES

- [1] K. Lakshmi, Mr. T. Chandra Sekhar Rao, "Design and Implementation Of Text To Speech Conversion Using Raspberry Pi", International Journal of Innovative Technology and Research, Volume No. 4, Issue No. 6, October-November, 2016.
- [2] D. Velmurugan, M.S. Sonam, S. Umamaheswari, S. Parthasarathy, K.R. Arun, "A Smart Reader for

Visually Impaired People Using Raspberry Pi", IJSEC, Research Article, Volume 6, Issue No. 3

- [3] Rithika H., B. Nithya Santhoshi, "Image Text To Speech Conversion In The Desired Language By Translating With Raspberry Pi".
- [4] Nagaraja L., Nagarjun RS, Nishanth M Anand, Nithin D, Veena S Murthy, "Vision based Text Recognition using Raspberry Pi", International Journal of Computer Applications, National Conference on Power Systems & Industrial Automation.
- [5] Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik, "Image to Speech Conversion for Visually Impaired", International Journal of Latest Research in Engineering and Technology, Volume 03, Issue 06, June 2017.
- [6] Aaron James S, Sanjana S, Monisha M, "OCR based automatic book reader for the visually impaired using Raspberry Pi", International Journal of Innovative Research in Computer and Communication Engineering, Volume 4, Issue 7, January 2016.
- [7] Christos Liambas, Miltiadis Saratzidis, "Autonomous OCR Dictating System for Blind People"