A Smart Reading Assistance for Visual Impaired People

[1]Prajakta S. Udgave, [2]S. G. Tamhankar
[1,2]Dept. of Electronics Engineering, Walchand College of engineering, Sangli, Maharashtra, India

Abstract
Visual impairment is one of the major limitations to a human. Communication and information access is one of the challenges for this people. Most of information is communicated with the text message. Text is available in the forms of books, newspapers, digital screens, commercial boards, sign boards etc. Now a day, worlds become more digital, with the availability of digital technology. There is a need of digital aid to assist the visual impaired people on the problem of recognizing printed text in the any image such as scene or printed image. In the motive of supporting visual impaired, a system assists such people by reading out the printed or handwritten text to this people. This system is also applicable to normal human beings who read the text as a speech as possible. In this system, we want to design a device that captures an image through camera, extract only part of image that contains only text and converts that text content into speech. This system is constituted by Raspberry pi 3B+, high definition camera, headphone and power supply. The captured image go through a variety of image pre-processing techniques to detect only that region of the image that holds the text and removes the background. To convert the text image to speech two tools are used. They are image pre-processing techniques and TTS (Text-to-Speech) synthesizers. The aural output is heard through the speakers or earphones.

Keywords
Raspberry PI, Camera, Speaker, Image Processing, Tesseract OCR, TTS Engine.

I. Introduction

Visual impairment or vision loss is a decreased ability of visual system that is not repairable by a usual means such as glasses or contact lenses. It is used to describe any vision loss whether it is complete or partial blindness. Every daily activities like walking, driving, reading and recognizing the objects and places are affected by this disability on such people. Due to diabetic, accident, childhood blindness, some eye infection and other causes, the amount of visually impaired people are increasing day by day. There are many excellent document system available in industries, academics and research laboratory due to digitization of books and any printed material, but this is only for normal human. Braille language is normally used by blind and visual impaired people for touch reading the documents and books which are less reliable to read. Braille materials are complicated to produce and less readily available in braille format for use. It is essential to develop a system with smart algorithm to detect the text content from text image, to satisfy the need of visual impaired community. To enable assistive technology to aid increasing population of visual impaired people with the help powerful tools such as image processing tools and text to speech synthesizer. The proposed system uses Py-tesseract to extract the text from texts image. Py-tesseract is an optical character recognition tool for python. And it will recognize the text present in images. And Espeak text into speech synthesizer is used to convert recognized text to speech. Espeak is an open source software text to speech synthesizer for English and other languages.

II. Literature Review

There are number of devices available for the blind and visually impaired person as a reading assistance. And different methods are present to detect and extract the text from captured image and to convert this extracted text into speech using text to speech synthesizer. The basic properties and limitation of these device and methods are discussed.

In this paper [1], EAST text detector is used. EAST is an Efficient and Accurate Scene Text detector that directly predict the words or text line level. It is a neural network model and which is trained to directly predict the existence of text instances and their geometries from images. They experiment two geometry shapes for text regions, rotated box (RBOX) and quadrangles (QUAD). It is performed on different datasets such as ICDAR 2015, COCO-Text and MSRA-TD500. And also they compare all datasets to compute better speed. This algorithm is outperforms in better accuracy and speed. On the ICDAR 2015 dataset, this algorithm achieves an F-score of 0.7820 at 13.2fps at 720p resolution.

In this paper [2], first text is detected by using connected component method and it is easy and fast. And recognize the text from detected and extracted patterns using feature learning method. It will give better result in recognition task. LABVIEW software is used to convert this extracted text into speech. This system helps visually impaired people as travel aid for reading assistance. This system works faster and better.

In this paper [3], they convert the text information to audio signal. They use OTSU algorithm for deciding the threshold (it will convert grey scale image to binary image) and canny edge detector for segmentation. Graphical user interface in MATLAB software is used to convert text into speech. This system is simple and uses an efficient algorithm for conversion of image to text. This system is very accurate as per the result.

In this paper [4], a mobile app is designed for android device that is named as 'Blind reader'. In this mobile app, mobile screen can be touched by a person by moving his finger tip horizontally from left to right on mobile screen just like a reading book. The system gives voice feedback those words which are touched. If user is visually disable, most of information on screen is invisible by user. Moreover since a blind person is unable to interpret the action in the interface, complexity will remain. To minimize the complication, same the touch type and touch duration time is assigned for distinct action for example, if user single taps on word, the word is readout, but if he uses double tap on a word the line number containing the word is read. Again a shorter tap on any option in the setting panel, will result in triggering. For voice feedback about option, one can use longer tap. This system can give the blind people a sense of haptic feeling. But this system has

requirement of an android phone, so it is not economical.

In this paper [5], to capture text image, finger mounted camera is used. This captured image is analyzed using (OCR) optical character recognition technique. And dataset is loaded to match the detected text with captured image. Once it match the text is synthesized for producing speech output using text-speech engine. In this method, they get 83% accurate result of various input set and hardware setup of finger reader. They uses MATLAB simulator.

In this paper [6], there is a device that converts a text image to speech. This system can be used to extract the text data from any printed text and convey this information to the visually impaired in audio form. To convert text image into speech, two tools are used and they are py-tesseract OCR method and text to speech engine. And output is heard through raspberry pi's audio jack using speaker or earphones. This device will give better accuracy.

In this paper [7], the system enables the blind and visual impaired person to hear the content of text image. It merges tesseract optical character recognition for text detection and recognition from captured image and voice output is produced by using text to speech synthesizer. This system helpful for the visually impaired or blind people to interact with computer effectively through audio interface. This system is an efficient and real time cost beneficial.

III. Proposed system

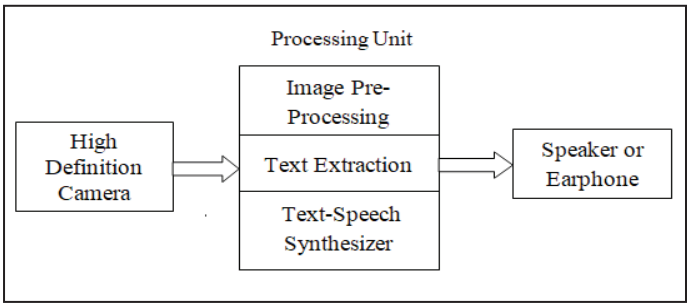

Fig. 1: Basic Block Diagram of Smart Reading Assistance for Visual Impaired People

The system composed of a raspberry Pi 3B+ module, speaker or earphones, HD camera and power supply. The camera must manually alligned to the text and a image is captured. Then this image is processed by the Raspberry Pi and the audio output is heard through the speaker.

First noise is removed from captured or printed image and then it is converted to grey scale image. After performing thresholding we get binary image. Once we get binary image, then segmentation is carried out by using Canny algorithm. Tesseract Ocr will recognise each and every character . Then the recognized text is converted into speech using Espeak text- speech synthesizer. And this speech is heard through speaker or earphone.

References

[1] Xinyu Zhou, Cong Yao, He Wen, Yuzhi Wang, Shuchang Zhou, Weiran He, and Jiajun Liang, "EAST: An Efficient and Accurate Scene Text Detector", 2017.
[2] Akhilesh A. Panchal, Shrugal Varde, M.S. Panse, "Character detection and recognition system for visually impaired people", 2016.
[3] Sandeep Musale, Vikram Ghiye, "Smart reader for visually impaired", 2018.
[4] Shahed Anzarus Sabab, Md. Hamjajul Ashmafee, "Blind reader: An intelligent assistance for blind", 2016.
[5] Sanjana. B, J. Rejina Parvin, "Voice assisted text reading system for visually impaired person using TTS method", IOSR Journal of VLSI and Signal Processing (IOSR-JVSP).
[6] Asha G. Hagargund, Sharsha Vanria Thota, MitadruBera, Eram Fatima Shaik,"Image tospeech conversion for visual impaired", International Journal of Latest Research in Engineering and Technology (IJLRET)
[7] K Nirmala Kumari, Meghana Reddy J,"Image text to speech conversion using OCR technique in Raspberry Pi", IJAREEIE 2016, Vol. 5.

Miss. Prajakta S. Udgave, she is an Mtech student of Walchand College of Engineering, Sangli. She completed her Graduation in Electronics engineering from DKTE, Ichalkaranji. She has a keen interest in the field of digital image processing.

Mr. S. G. Tamhankar, he received his ME degree in Electronics. He is Assistant Professor in department of Electronics Engineering at Walchand College of Engineering, Sangli. He has more than 20 years of teaching experience. Area-Computer Network, IOT, Image Processing.