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

Raspberry Pi based Smart Reader for Visually Impaired People

K. Barathkumar[1], S. Balaji[2], K. N. Desikan[3], S. Padhmanabha Iyappan[4]
[1,2,3]*Student, Dept. of Electrical and Electronics Engineering, Valliammai Engineering College, Chennai, India*
[4]*Assistant Professor, Dept. of Electrical and Electronics Engg., Valliammai Engg. College, Chennai, India*

***Abstract*: According to the World Health Organization, out of 7.4 billion population around 285 million people are estimated to be visually impaired worldwide. It is observed that they are still finding it difficult to roll their day today life and it is important to take necessary measure with the emerging technologies to help them to live the current world irrespective of their impairments. In the motive of supporting them, we have proposed a smart spec for the blind persons which can perform text detection thereby produce a voice output. This can help the visually impaired persons to read any printed text in vocal form. A specs inbuilt camera is used to capture the text image from the printed text and the captured image is analysed using Tesseract-Optical Character recognition (OCR). The detected text is then converted into speech using a compact open source software speech synthesizer, e-Speak. Finally, the synthesized speech is produced by the headphone by TTS method. In this project Raspberry Pi is the main target for the implementation, as it provides an interface between camera, sensors, and image processing results, while also performing functions to manipulate peripheral units (Keyboard, USB etc.,).**

***Keywords*: Raspberry Pi, Tesseract Optical Character recognition, Text to Speech engine, e-Speak.**

1. Introduction

Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. We present a smart device that assists the visually impaired which effectively and efficiently reads paper-printed text. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document. The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board. The design is motivated by preliminary studies with visually impaired people, and it is small-scale and mobile, which enables a more manageable operation with little setup. In this project we have proposed a text read out system for the visually challenged. The proposed fully integrated system has a camera as an input device to feed the printed text document for digitization and the scanned document is processed by a software module the OCR (optical character recognition engine). A methodology is implemented to recognition sequence of characters and the line of reading. As part of the software development the Open CV (Open source Computer Vision) libraries is utilized to do image capture of text, to do the character recognition. Most of the access technology tools built for people with blindness and limited vision are built on the two basic building blocks of OCR software and Text-to-Speech (TTS) engines.

Optical character recognition (OCR) is the translation of captured images of printed text into machine-encoded text. OCR is a process which associates a symbolic meaning with objects (letters, symbols an number) with the image of a character. It is defined as the process of converting scanned images of machine printed into a computer process able format. Optical Character recognition is also useful for visually impaired people who cannot read Text document, but need to access the content of the Text documents. Optical Character recognition is used to digitize and reproduce texts that have been produced with non-computerized system. Digitizing texts also helps reduce storage space. Editing and Reprinting of Text document that were printed on paper are time consuming and labour intensive. It is widely used to convert books and documents into electronic files for use in storage and document analysis. OCR makes it possible to apply techniques such as machine translation, text-to-speech and text mining to the capture / scanned page. The final recognized text document is fed to the output devices depending on the choice of the user. The output device can be a headset connected to the raspberry pi board or a speaker which can spell out the text document aloud.

Table 1
Prevalence of blindness (per thousand) as per the estimate of 2018

Type of Blindness	Urban	Rural	Total
Total Blindness	4.62	6.12	5.57
Economic Blindness	11.34	15.57	13.87
One eye Blindness	8.34	7.15	7.62

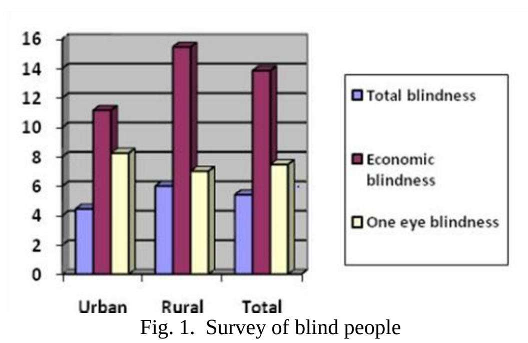

Fig. 1. Survey of blind people

Total Blindness = Visual acuity less than 3/60 in better eye with spectacle correction.

Economic blindness = Visual acuity less than 6/60 in the better eye with spectacle correction.

One eye blindness = Visual acuity less than 3/60 in one eye and better than 6/60 in the other eye with spectacle correction.

2. Block diagram of proposed method

The fig. 2 illustrates the block diagram of proposed method. The framework of the proposed project is the raspberry pi board. The raspberry pi B+ is a single board computer which has 4 USB ports, an Ethernet port for internet connection, 40 GPIO pins for input/ output, CSI camera interface, HDMI port, DSI display interface, SOC (system on a chip), LAN controller, SD card slot, audio jack, and RCA video socket and 5V micro USB connector. The power supply is given to the 5V micro USB connector of raspberry pi through the Switched Mode Power Supply (SMPS). The SMPS converts the 230V AC supply to 5V DC. The web camera is connected to the USB port of raspberry pi. The raspberry pi has an OS named RASPION which process the conversions. The audio output is taken from the audio jack of the raspberry pi. The converted speech output is amplified using an audio amplifier. The Internet is connected through the Ethernet port in raspberry pi. The page to be read is placed on a base and the camera is focused to capture the image.

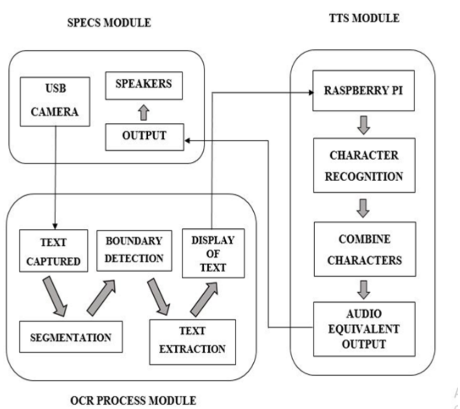

Fig. 2. Block diagram of Proposed Method

The captured image is processed by the OCR software installed in raspberry pi. The captured image is converted to text by the software. The text is converted into speech by the TTS engine. The final output is given to the audio amplifier from which it is connected to the speaker. Speaker can also be replaced by a headphone for convenience. Here it is shown how to style a subsection and sub sub-section also

3. Flow of process

A. *Image capturing*

The first step in which 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 high so as to have fast and clear recognition due to the high resolution camera.

B. *Pre-processing*

Pre-processing stage 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 between ±15 degrees and if detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a skew corrected image. The noise introduced during capturing or due to poor quality of the page has to be cleared before further processing.

C. *Segmentation*

After pre-processing, the noise free image is passed to the segmentation phase. It is an operation that seeks to decompose an image of sequence of characters into sub-image of individual symbol (characters). The binarized image is checked for inter line spaces. If inter line spaces are detected then the image is segmented into sets 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 characters using character width computation.

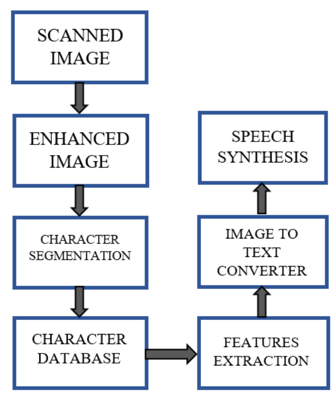

Fig. 3. Flow of process

D. *Feature extraction*

Feature extraction is the individual image glyph is considered and extracted for features. First a character glyph is defined by the following attributes:

- Height of the character;
- Width of the character;
- Numbers of horizontal lines-short and long;
- Numbers of vertical lines-short and long;
- Numbers of circles present;
- Numbers of horizontally oriented arcs;
- Numbers of vertically oriented arcs;
- Position of the various features;
- Pixels in the various regions.

E. Image to text converter

The ASCII values of the recognized characters are processed by Raspberry Pi board. Here each of the characters is matched with its corresponding template and saved as normalized text transcription. This transcription is further delivered to audio output.

F. Text to speech

The scope of this module is initiated with the conclusion of the receding module of Character Recognition. The module performs the task of conversion of the transformed text to audible form. The Raspberry Pi has an on-board audio jack, the on-board audio is generated by a PWM output and is minimally filtered. A USB audio card can greatly improve the sound quality and volume. Two options of attaching a microphone into Raspberry Pi. One is to have USB mic, another to have an external USB sound card.

4. Simulation environment

The image to text and text to speech conversion is done by the OCR software installed in raspberry pi. The conversion which is done in OCR can be simulated in MATLAB. The conversion process in MATLAB includes the following processes.

- Binary image conversion.
- Complementation.
- Segmentation and labelling.
- Isolating the skeleton of character.

A. Sample image

The following image which is captured by the webcam contains the following word. This image is in the jpeg format which has to be converted into text.

Fig. 4. Sample Image

B. Binary conversion

In this section sample image is converted into binary format. The image which was a 3D image initially is converted to 2D image. Binary 0 represents black color of the characters. Binary 1 represents white color of the characters.

C. Boundary marking

Fig. 5. Binary 0 text representation

The area of the text is bordered and the boundary for each character is isolated. The boundary for each character is programmed and it can vary from 0 to 255 bits of characters occupying memory in the database.

Fig. 6. Binary 1 text representation

D. Segmentation and labelling

The isolated blocks of characters are segmented and are automatically labelled for identity. Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. Connected-component labelling is used in computer vision to detect connected regions in binary digital images, although color images and data with higher dimensionality can also be processed. When integrated into an image recognition system or human-computer interaction interface, connected component labelling can operate on a variety of information. Blob extraction is generally performed on the resulting binary image from a thresholding step. Blobs may be counted, filtered, and tracked.

E. Audio output

The programming codes are run in MATLAB and corresponding output is generated. The output is in the form of audio. The audio is heard using headphone or speaker connected to the system. Each character of the word is spelled out first and then the entire word is read out.

5. Hardware implementation

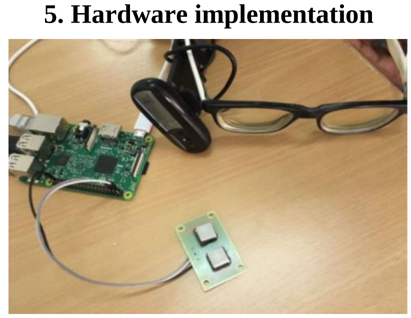

Fig. 7. Hardware Setup

The hardware of the proposed work consists of a raspberry pi board interfaced with a USB camera. WiFi dongle is connected to the system for internet connection which is taken to Pi through LAN cable. A 5MP camera is connected to one of the USB port of raspberry pi. A 5V supply is given to Raspberry pi from the system through a power cable.

6. Experimental outputs

The text document which has to be read out is placed at a considerable distance from the webcam so that the image is clear enough with proper illumination. The proposed system is tested with different input sets as, Text in the form of Black and White, Text with different font styles, Text in the form of colored. In all five sets specs inbuilt camera is tested using OCR process with around 2 samples of 100 letters which yields an accuracy of around 98%.The displayed image is read out by the text to speech engine e-speak.

Table 2
Tested output results

Test Results for Black and White coloured image	
Number of Words	150
Number of Words detected	143
Error Possibility	07

Test Results for Text with Different Font Styles	
Number of Words	100
Number of Words detected	91
Error Possibility	09

Thus the simulated results of various input sets and the hardware setup of spectacles with different input sets and the recognized output is viewed through MATLAB simulator with an audio output.

7. Conclusion

We have implemented text to speech conversion technique using raspberry pi. The simulation results have been successfully verified and the hardware output has been tested using different samples. Our algorithm successfully processes the image and reads it out clearly. This is an economical as well as efficient device for the visually impaired people. We have applied our algorithm on many images and found that it successfully does its conversion. The device is compact and helpful to the society.

Acknowledgement

The authors would like to thank the anonymous reviewers for their constructive comments. We also would like to thank our institution for their support to carry out our project in the laboratory.

References

[1] Bindu Philip and R.D. Sudhaker Samuel`s, “Human Machine Interface – A Smart OCR for the Visually Challenged” International Journal of Recent Trends in Engineering (IJRTE), Volume 3, November 2009.
[2] V. Ajantha Devi and Santhosh Baboo`s “Embedded Optical Character Recognition on Tamil Text Image using Raspberry Pi”, International Journal of Computer Science trends and Technology (IJCST), Volume. 2, Issue 4, July-August 2014.
[3] Jaiprakash Verma, Khushali Desai, and Barkha Gupta`s “Image to Sound Conversion”, International Journal of Advance Research in Computer Science and Management Studies (IJARCSMS), Volume 1, Issue 6, November 2013.
[4] Pooja Deole, and Shruti Kulkarni`s “Finger Reader: A Wearable Device for The Visually Impaired”, International Journal of Infinite Innovations in Technology (IJIIT), Volume 4, Issue 2, October 2015.
[5] Jisha Gopinath, S S Saranya, Pooja Chandran, and S Aravind`s “Text to Speech Conversion System using OCR”, International Journal Of Emerging Technology and Advanced Engineering (IJETAE), Volume 5, Issue 1, January 2015.
[6] M.A.Raja, R Ani, Effy Maria, J Jameema Joyce, and V Sakkaravarthy`s “Smart Specs: Voice Assisted Text Reading system for Visually Impaired Persons Using TTS Method”, International Conference on Innovations in Green Energy and Healthcare Technologies (ICIGEHT), March 2017.
[7] M Abhijit, A Savitri, R Pooja, K Amrutha, and Vikram Shirol`s “DRASHTI – An Android Reading Aid”, International Journal of Computer Science and Information Technologies (IJCSIT), Volume. 6, Issue 4, July 2015.
[8] D Goldreich and I.M. Kanics`s, “Tactile Acuity is Enhanced in Blindness”, International Journal of Research in Science (IJRS), Volume. 23, Issue 8, 2003.
[9] M S Sonam, S Umamaheswari, S Parthasarathy, K R Arun, and D Velmurugan`s “A Smart Reader for Visually Impaired People using Raspberry Pi”, International Journal of Engineering Science and Computing (IJESC), Volume. 6, Issue 3, March 2016.
[10] Catherine A. Todd, Ammara Rounaq, Umm Kulsum Nur and Fadi Boufarhat`s “An Audio Haptic Tool for Visually Impaired Web Users”, Journal of Emerging Trends in Computing and Information Sciences (JETCIS), Volume. 3, Issue 8, August 2012.