International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 03 - Issue 10 || October 2018 || PP. 39-46

A Smart Reader for Visually Impaired People (Standard Image Vs Real Time Image: A Comparative Study)

Ram Nivas Duraisamy[1], Sathya Manoharan[2]
[1]*(Assistant Professor, Department of Electronics and Communication Engineering/ P. A. College of Engineering and Technology, Pollachi, Coimbatore, India)*
[2]*(UG Scholar, Department of Electronics and Communication Engineering / P. A. College of Engineering and Technology, Pollachi, Coimbatore, India)*

Abstract: Optical character recognition (OCR) is the process of identifying the printed characters using photoelectric device and computer software. It converts images of typed, handwritten or printed text into machine encoded text from scanned document or from text superimposed on an image. These images are converted into audio output. OCR is mainly used in the field of research in character recognition artificial intelligence and computer vision and it is also used for pattern recognition and to perform Document Image Analysis. The work focuses on the OCR based automatic book reader for the visually impaired using Raspberry Pi. The aim is to provide assistance to the visually impaired at low cost and to demonstrate the easily designable version. The results of text detection rate, accuracy, text to speech conversion rate and error rate are tabulated. The proposed model's result is compared with MATLAB image processing method.
Keywords: Optical character recognition, Raspberry Pi, Document Image Analysis, Tesseract, Festival

I. INTRODUCTION

The World Health Organization reports that 285 million people are estimated to be visually impaired worldwide. Among the 285 million, 90 Percent live in developing countries. Visually impaired and blind people may be elderly or young. They may be sportsmen and women, gardeners, farmers, chess players, teachers, typists, musicians, lawyers, housewives, computer programmers, physiotherapists and social workers. There are many alternate solutions to the problem of assisting individuals who are blind to read. The problem with the solutions is that they are not cost effective and their implementation is also complex.

People with physical disabilities can use mobility aids, such as wheelchairs, scooters, walkers, canes, crutches, prosthetic devices, and orthotic devices, to enhance their mobility.

Visually impaired people suffer from numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. A Smart device that assists the visually impaired which effectively and efficiently reads paper-printed text is developed. The Smart Reader uses the methodology of a camera based assistive device that can be used by people to read text document. The platform is the combination of image processing technique and embedded systems.

The design is motivated by observation of visually impaired people, and it is small-scale and mobile device, which enables a more manageable operation with a simple setup. The setup has speaker read out for voice output. The Smart Reader has a camera as an input device to feed the printed text document. The Acquired image is digitized and the scanned document is processed by a software module called the Optical Character Recognition engine.

As a part of the software development, the Open source Computer Vision libraries are utilized to capture the text and recognize characters. Most of the access technology tools built for visually impaired are built on the two basic building blocks of OCR software and Text-to-Speech engines. Optical character recognition is the translation of captured image text into machine encoded text. It is defined as the process of converting scanned images into a computer process able format. Optical Character recognition is useful for visually impaired people for accessing the content of the text documents.

Optical Character recognition is also used to digitize and reproduce texts that have been produced with non-computerized system. Digitizing texts also helps to reduce storage space. Reprinting and Editing of text document on still image paper are time consuming and labor intensive. It is widely used to convert books and documents into electronic files for storage and document analysis.

OCR makes it possible to apply techniques such as machine translation, text-to-speech and text mining to the capture or scanned page. The recognized text document is fed to the output devices depending on the choice of the user. The output device can be a headset or speaker connected to the Raspberry PI board or a speaker which can spell out the text document aloud.

II. LITERATURE SURVEY

The implementation of the assist devices for the visually impaired is carried out by different methods. It is found that the classical or traditional methods seemed to be complex, costlier and less accurate. The following are the different methodologies adopted in different text conversions, speech processing and text processing.

Aaron James et al (2016), designed an OCR Based automatic book reader for the visually impaired using Raspberry Pi. It uses Energy Efficient Character Recognition, document image analysis and python programming for implementation. The reader is costlier in implementation [1].

D Dakopoulos et al (2010), presents wearable obstacle avoidance electronic travel aids for blind. Comparative survey among portable or wearable obstacle detection and avoidance systems to inform about the progress in assistive technology for visually impaired people is taken [2].

N Giudice et al (2008), adopted Blind navigation technology and explained about the navigational technologies available to blind individuals to support independent travel on large scale [3].

Kumar S et al (2007), extracted and documented the text image using Matched wavelets and MRF Models. The Method has a novel scheme for the extraction of textual areas of an image using globally matched wavelet filters. A clustering based technique is devised for estimating globally matched wavelet filters using a collection of ground truth images [4].

K Kim et al (2004), proposed a Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm for the textural properties of scripts [5].

Chen X et al (2004), detected and read the texts in natural scenes. He proposed an algorithm for detecting and reading the text from the city scenes as they walk through [6].

Chen J Y et al (2004), automatically detected and recognized the image. He presents an approach to automatically detect and recognize signs its application to sign translation task that further propose a local intensity normalization method to effectively handle lighting variations followed by Gabor transform for obtaining local features [7].

All the above observations seemed to be rugged in construction and costlier in implementation. The complexity is found in technologies, algorithms, hardware constrains and transforms. The desired output is obtained only for the standard images with 100 percent accuracy. The accuracy degrades for real time images.

III. HARDWARE AND SOFTWARE REQUIREMENTS

The raspberry Pi is a small single board computer developed by Raspberry Pi Foundation to promote teaching in educational institutions. It is widely used in majority of the engineering projects. Optical character Recognition is a electronic conversion of hand written or printed text into machine encoded text from the scanned Image. Business cards, handouts of statistical data can be considered as examples. It is a common method of digitizing printed text so that they can be electronically processed. OCR is a widely adopted of research in fields of artificial Intelligence, pattern recognition, and computer vision. Tesseract is an optical character recognition engine for various operating systems, it was considered one of the most accurate open source OCR engine. There are few constrains while using Tesseract as its inputs must be scaled up to height of atleast 20 pixels. The presence of skew or rotation may degrade the performance. Festival is a general multi lingual speech synthesis system developed at center of speech technology research. This includes languages like British, English, Italian, Czech and Spanish.

IV. PROPOSED METHOD

4.1 Introduction

The proposed method requires the input image processor converts in to suitable speech conversion. Since the advanced methodology like Raspberry Pi, Tesseract and Festival were used, the accuracy seems to be good when compared with classical methods of Image processing and OCR. The block diagram of the proposed method is shown in figure1.

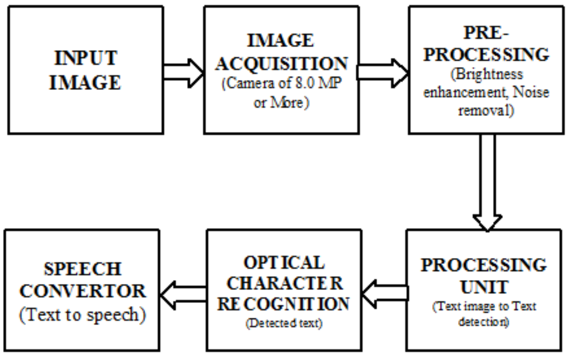

Fig 1 Block diagram of the proposed method

4.2 Image Acquisition

Image acquisition is the capturing of the image. The text to be captured is the test data created by the user which is half the size of A4 size(720*340) resolution.

4.3 Pre-Processesing

Since the captured image is text. The image data or background is the unwanted or irrelevant information. This may affect the accuracy of the system. It is essential to suppress the background information. The image which is converted into binary image may be cropped of suitable dimensions. After the entire pre-processing is over, the detected text image is stored as detected.jpg.

4.4 Optical Character Recognition

The proposed method makes use of OCR inbuilt in Raspberry PI which is difficult in traditional image processing methods which requires more samples to be trained , the accuracy of the application depends on the exact replication of the alphabets in obtaining more accuracy in fonts like Times New Roman, Arial and Calibri could be used. Hand written scripts have less accuracy due to variable strokes by users.

4.5 Text To Speech Conversion

Speech synthesis is the artificial production of speech. A text-to-speech system converts a language text into speech. The quality of the speech synthesizer is judged by its similarity to the human voice and its ability to be understood clearly.

4.6 Pseudo Code

Begin;
Initialize the hardware and software components.
    Camera : Enable/ disable;
    Speaker : Enable/ disable;
    Pi Board : Enable/ disable;
    Start : 0/1;

[The Raspberry pi board is installed with raspbian Operating system. The software part includes tessract, Mplayer, festival.]

Input : Image with text data in majority of the portions with less background information
STEP 1:
Begin (While Start =1)
    Acquisition of image:
    Board : Enable;
    Camera : Enabled;
    Speech Engine : disable;
    Speaker : Disable;
    Acquired image : image .jpg;
    Enhanced image : Binary image (image.jpg);   // for better detection of text data;
    Loop : Disable;
    Goto : step 2;
End
STEP 2:

```
Begin (While End of stroke indication = 0)
        Pi Board : Enable
        Camera : Disable;
        Speech Engine : disable;
        Speaker : Disable;
        Stored folder : Enable;
        Input Image : Enhanced Image''; [Fetched from the destination of stored folder]
        OCR : enabled''; [ Image to text conversion from the enhanced image];
        Notepad Document : Enable''; [ Store the converted text in notepad.txt / speech.txt]
        Loop : Disable;
        Goto : Step3;
End
STEP 3:
Begin (While End of line indication = 0)
        Pi Board : Enable
        Camera : Disable;
        Speaker : enable;
        Stored folder : Enable;
        Input : speech.txt [ Refer Step 2]
        Speech Engine : Enable;
        Process : Text to speech conversion;
        Loop : Enable;
Goto : Step:1;
End
End
```

The hardware and software components are initialized. The key component comprises of camera, speaker and Pi board. The process is started with the start being enabled. The Raspberry pi board is installed with Raspbian Operating system. The software part includes tesseract, Mplayer, festival. The input image is acquired with the help of camera, the image with majority of the portions as text is considered.

The background data is eliminated because it may lead to false strokes. To eliminate background data, the contrast and brightness of the image is adjusted and the natural image is converted in to the binary image.

The process begins with enabling step 1, the peripherals like pi board and camera are enabled. Speech engine and speakers are disabled. The acquired image is 'image.jpg'. The image is enhanced for better detection of text data. The loop is disabled. The completion of step 1 leads to step 2.

The step 2 begins with end of stroke indication. The Pi Board and stored folder are enabled. Camera, speech engine and speakers are disabled. The input image is fetched from the destination folder and it is enhanced. OCR and Notepad are enabled. The converted text is saved as notepad.txt. The completion of step 2 leads to step 3 with loop being disabled.

The step 3 begins with end of stroke indication. The Pi Board, speaker speech engine and stored folder are enabled. The text 'notepad.txt' is the input to the process. The completion of step 3 leads to step 1 with loop being enabled.

V. RESULT AND DISCUSSION

5.1 Steps in Smart Reader

1. The reader operates in 3 modes, Acquisition, text conversion and speech conversion modes.
2. The raspberry Pi is turned on.
3. In the Command prompt, "Python filename.py" is made to run.
4. The image is acquired within 10 seconds from the moment you enable the command.
5. The image is saved in .jpeg format with resolution of 1280 x 720.
6. The image is enhanced and it is converted in to binary image and the background data is completely eliminated. The processed image is saved as process.jpg image.
7. Tesseract tool is enabled. Image to text conversion takes place. Where the text data is fetched from the binary image process.jpg, each combination of line strokes are converted in to appropriate text and they are displayed in command prompt and saved as text.txt file.
8. Festival tool is enabled. Now the text.txt is open and the text is converted in to speech. As soon as the end of the line in reached the device loops back to acquisition mode.

Figure 2. shows the Experimental setup of the Smart Reader. It comprises of camera, display unit, Pi Board and a pair of speakers

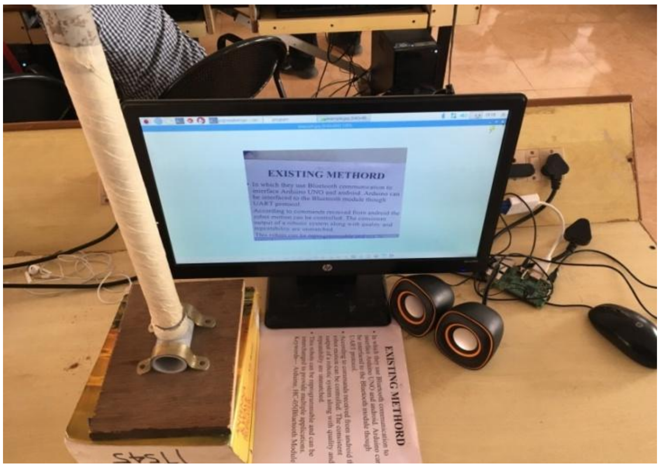

Fig 2 Experimental setup

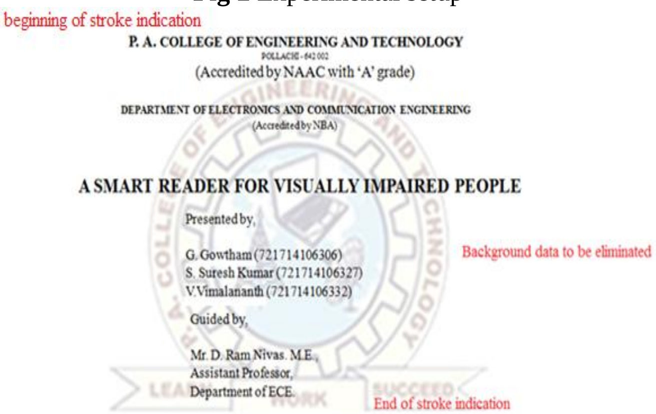

Fig 3 Input image(User Defined)

Figure 3 shows the input image with major area occupied with text data and least background information. The beginning of text is considered to be 'beginning of stroke indication' and the end of text is indicated as 'end of stroke indication'

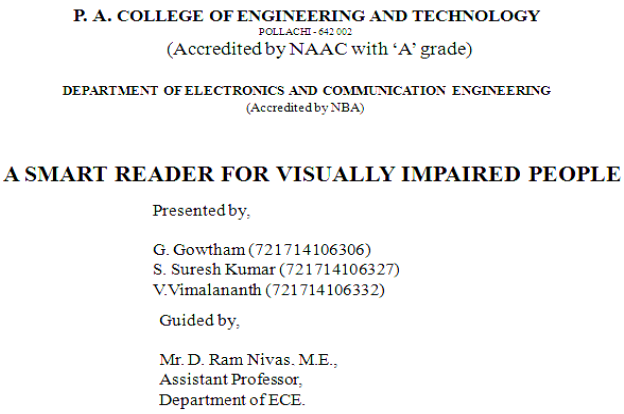

Fig 4 Processed binary Image with less background data

Figure 4 shows the processed binary image with least background information. The image to text conversion is shown in figure 5.

Table 1 : Accuracy					
Sample Image no.	Number of characters to be detected	Detected characters (Method 1) Classical		Detected characters (Method 2) Proposed	
	100 % Accuracy	Count	Accuracy	Count	Accuracy
1(Standard Image)	38	38	100	38	100
2 (Standard Image)	83	83	100	82	98.71
3(User Defined) (Trial 1)	341	280	82.11	330	96.77
4(Trial 2)	341	283	82.99	335	98.24
5(Trial 3)	341	279	81.81	329	96.48
6(Trial 4)	341	300	88.23	338	99.12

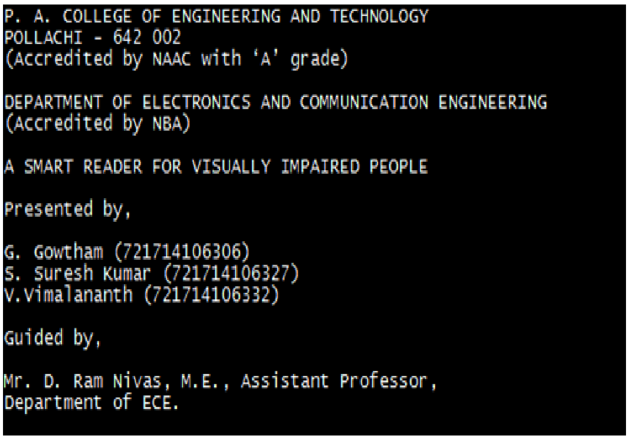

Fig 5 Image to text conversion

Accuracy for 10 different samples 90 percent is achieved. It is determined by the resolution of the camera. When good resolution of more than 12MP is used. We can attain accuracy of more than 98%. Illumination source is also a key factor to be considered. The experiment holds excellent at good day light.

Fig 6 Standard Image (Source: www.in.mathworks.com)

5.2 Comparitive Study

Consider the Table 1: Standard images and user defined images are considered. The Matlab image processing results and the result of the proposed method are tabulated and considered under study. The observed results are as follows. The standard images give 100 percentage accuracy in both the methods where only fewer characters are detected. Considering the second case the first method is 100 percentage accurate but there is a lag in experimental set up result due to degradation phenomenon. Phenomenon's may be poor lighting conditions or tilted acquisition angle. The accuracy percentage is represented graphically in figure 7.

Accuracy = [Detected Characters (Method1 or Method2) / Characters to be detected] X 100 (1)

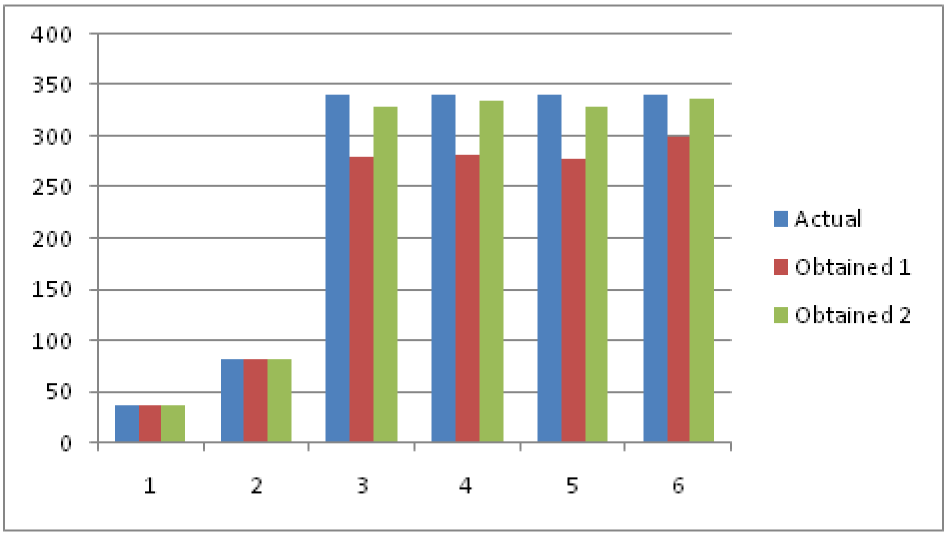

Fig 7 Accuracy Chart

Consider the Table 2: The accuracy in table 2 indicates the text to speech conversion rate. Trails 2 and 4 seemed to be good. The algorithm is accurate that it does not go below 95 percentage of accuracy under good lighting conditions and acquisition rate. The system works well only for prescribed fonts like Times New Roman, Arial, Calibri and Tahoma.

Table 2 : text to speech conversion rate

Trail Numbers	Detected Characters	Accuracy	Description
1	330	96.77	Moderate
2	335	98.24	Good
3	329	96.48	Moderate
4	338	99.12	Good

Consider the Table 3 : It is observed that the prescribed result has less deviation when compared with the original value. The performance of the prescribed method is 80 percentage more accurate than the method 1. The calculations for deviations and improvement percentage are shown below.

Deviation = 100 – Actual Accuracy (2)

Improvement Percentage = [[Deviation 1 – Deviation 2]/ Maximum Deviation] x 100 (3)

Equation (2) and (3) represents Deviation and Improvement Percentage.

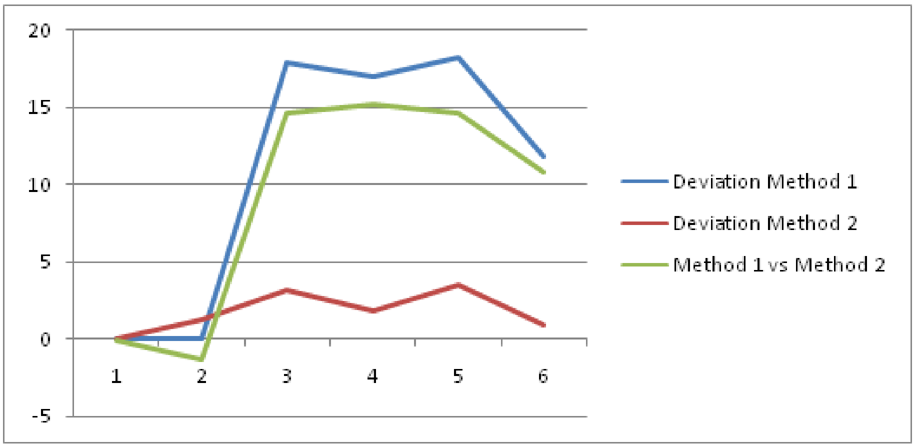

Fig 8 Deviation Chart

VI. CONCLUSION AND FUTURE WORK

A prototype system to read printed text and hand held objects for assisting the blind people is proposed. To extract text regions from complex backgrounds, novel text localization algorithm based on models of stroke orientation and edge distributions is adopted. An image to speech conversion technique using Raspberry Pi was implemented. Output has been tested using different samples. The algorithm successfully processes the image and reads it out clearly. This is an economical as well as efficient device for the visually impaired people. The accuracy for hand written text has to be increased by improving the illumination source and acquisition device. The major limitation of the work is the system lags in case of hand written text. A complex and secured system should be implemented with the help of deep learning algorithms or neural networks for providing better reliability.

REFERENCES

Journal Papers:

[1] Aaron James S , Sanjana S and Monisha M, OCR based automatic book reader for the visually impaired using Raspberry PI, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 7, January 2016, 1111-1118.

[2] Dimitrios Dakopoulos ; Nikolaos G. Bourbakis, *Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) ,Volume:40 , Issue: 1 , Jan. 2010* , 25-35.

[3] Giudice, N. A., & Legge, G. E, *Blind navigation and the role of technology. In A. Helal, M. Mokhtari & B. Abdulrazak (Eds.), Engineering handbook of smart technology for aging, disability, and independence (pp. 479- 500): John Wiley & Sons.*

[4] Sunil Kumar, Rajat Gupta , Nitin Khanna, Santanu Chaudhury and Shiv Dutt Joshi, *Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model, IEEE Transactions on Image Processing (Volume: 16 , Issue: 8 , Aug. 2007)* 2117 – 2128.

[5] Kwang In Kim, Keechul Jung, and Jin Hyung Kim, *Texture-Based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 12, DECEMBER 2003*, 1631-1639.

[6] Xiangrong Chen, Alan L. Yuille, *Detecting and Reading Text in Natural Scenes, Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*

[7] Xilin Chen , Jie Yang Jing Zhang, and Alex Waibel, *Automatic Detection and Recognition of Signs From Natural Scenes, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 13, NO. 1*, JANUARY 2004, 87-99.

Books:

[8] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins , *Digital Image Processing using Matlab* (Pearson Education, 2006)

Website References:

[9] https://in.mathworks.com/help/vision/examples/automatically-detect-and-recognize-text-in-natural-images.html.

[10] https://in.mathworks.com/matlabcentral/answers/110412-extract-text-from-image.