International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 14 Issue 1 –APRIL 2015.

OPTICAL CHARACTER RECOGNITION FOR VISUALLY CHALLENGED PERSONS

C. VIGNESH[#1], MR. M. SENTHIL KUMARAN[#2]

[#1]*PG Scholar, Department of CSE, SCSVMV University, Kanchipuram*
[#2]*Assistant professor, Department of CSE, SCSVMV University, Kanchipuram*

vignesh.jag@gmail.com
senth.kumaran2000@gmail.com

Abstract- **Enclosure of the specially enabled in the IT revolution is both a social responsibility as well as a computational dispute in the rapidly advancing digital world today. The availability of text books for the visually challenged is limited due to the process of conversion. Generally the book in the English will be typed into Braille sheet with the Braille format using the Braille printer. The effort taken for this is very high because for each and every character their corresponding Braille character should be printed. This system enables the automated conversion of scanned images of book directly to the Braille format. The optical character recognition is primarily used for this purpose.**

Optical character recognition (OCR) method has been used in converting printed text into editable text. OCR is very popular and useful method in numerous applications. One of the important applications of this tool is its use in Braille Translation. Braille is used as the primary writing and reading system used by the visually challenged since the 19th century. Text-to-speech (TTS) is a type of speech synthesis application that is used to create a spoken sound version of the text in a computer document.

KEYWORDS – Optical Character Recognition, Image To Braille, OCR To Braille, OCR To TTS

I INTRODUCTION

Computer literacy opens new specialized opportunities for the visually impaired and low vision persons, thus enhancing their employability and efficiency, giving them improved proficiency and confidence at work. Until these technologies are available for Indian languages, persons with blindness in India will remain on the dark side of the digital divide and will be forced to remain dependent on others for basic tasks like writing and reading books, reading a letter or writing examinations.

Most of the access modern technology tools built for people with visually challenged and limited vision are built on the two basic building blocks of OCR software and Text-to-Speech (TTS) engines, and any information on these would be invaluable for people with vision impairment.

Optical character Recognition (OCR) is a conversion of scanned or printed text images, handwritten text into editable text for further processing. This technology allows recognition engine to recognize the text automatically. It is like combination of eye and brain of human body. An eye can view the text from the images but actually the brain processes as well as interprets that extracted text read by eye [1]. In development of computerized OCR system, few problems can occur. First: there is very little visible difference between some letters and digits for computers to understand. For example it might be difficult for the computer to differentiate between letter “o” and digit “0”. Second: It might be very difficult to extract text, which is embedded in very dark background or printed on other words or graphics. In 1955, the first commercial system was installed at the reader’s digest, which used OCR to input sales report into a computer and then after OCR method has become very helpful in computerizing the physical office documents. There exists lot of applications for OCR, which includes: Image text extraction from natural scene images, license plate recognition, extracting text from scanned documents etc.

Today many types of OCR software available in the markets like: Web OCR, Desktop OCR, Server OCR etc. Accuracy rate of any OCR tool varies from 71% to 98%. Many OCR tools are available as of now but only few of them are open source and free rest of them are proprietary. Tesseract is the one of the open source and free OCR software. It is written in the C++, so it becomes platform independent. So that it's possible to use in other applications in the form of Dynamic Link Library (DLL). So it can be easily added as the reference in the form of DLL in other application to use the functionality provided by Tesseract. Instead of restricting strictly to recognize only one language like English, Russian, Hindi or Chinese. It is also available in multilingual format that can recognize various languages provided their language model [2].

II RELATED WORKS

Optical character recognition is a system presented by Ravina Mithe, Supriya Indalkar, Nilam Divekar developed an android application which performs extraction of text from the images and convert it into speech. They primarily used Tesseract as a OCR recognition engine for the extraction of text from the images [3].

Human Machine Interface – A Smart OCR for the Visually Challenged is a system presented by Bindu Philip and R. D. Sudhaker Samuel is used to convert the scanned images of Malayalam language into text format then it is converted into speech and Braille format [4].

III RECOGNITION ENGINE

Tesseract OCR engine works in step by step manner as described in the architecture diagram as shown in figure 1. First step is Adaptive Thresholding, which converts the image into binary images. Next step is connected component analysis, which is used to extract the outlines of character. This system is very useful because it does the OCR of image with white text and black background. Tesseract was probably first to provide this kind of processing. Then after, the outlines are converted into Blobs [5].

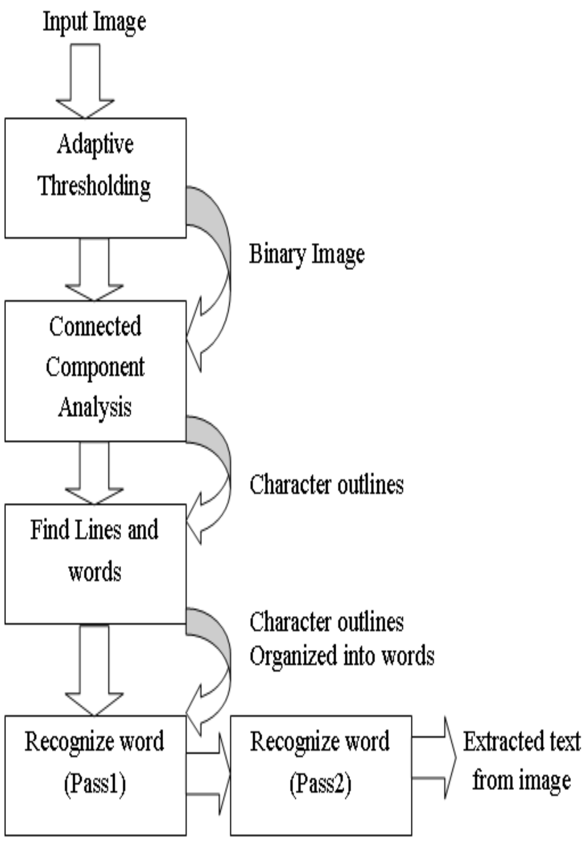

Fig. 1 Architecture of recognition engine

Blobs are prepared in the form of text lines, and the lines and regions are analyzed for some fixed area or equivalent text size. Text is divided into words using definite spaces and fuzzy spaces. Recognition of text from the image is then started as a two-pass process. In the first pass, an effort is made to recognize each word from the text. Each word passed satisfactory is passed to an adaptive classifier as training data. The adaptive classifier tries to recognize text in more precise manner. As some training data is given to adaptive it has learn something new so final phase is used to resolve various issues and to extract text from images.

IV BRAILLE

Braille is a tactile writing system used by the blind and the visually challenged. It is conventionally written with embossed paper. Braille-users can be able to read computer screens and other electronic supports with refreshable Braille displays [6]. They can able to write Braille with the original slate and stylus or type it on a Braille writer, such as a convenient Braille note-taker, or on a computer system that can print with a Braille embosser as shown in the Figure 2.

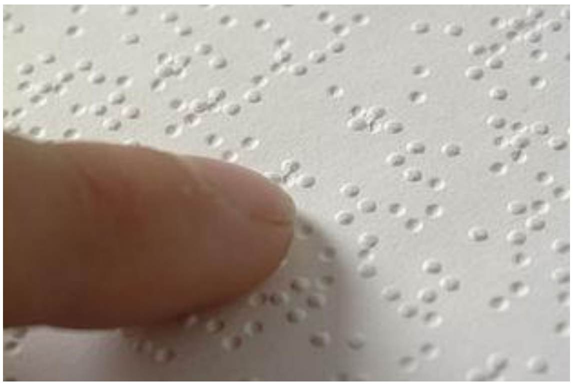

Fig. 2 Braille Sheet

Braille characters are small rectangular blocks called cells that contain tiny palpable bumps called raised dots. The number and planning of these dots distinguish one character from another. Since the a variety of Braille alphabets originated as transcription codes of printed writing systems, the mappings (sets of character designations) vary from one language to another language. Furthermore, in English Braille it can be of three levels of encoding: Grade 1, is widely used as a letter-by-letter transcription used for basic literacy; Grade 2, is used an addition of abbreviations and contractions; and Grade 3, various non-standardized personal shorthand's.

V SPEECH SYNTHESIZER

A text to speech (TTS) synthesizer is a system that can read text aloud automatically, which is extracted from Optical Character Recognition (OCR). A speech synthesizer can be implemented by both hardware and software [7]. Speech synthesis is the artificial production of human dialogue. A computer scheme used for this purpose is called a speech synthesizer. A text-to-speech (TTS) system converts normal language text into speech as per the architecture shown in the Figure 3. A synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output.

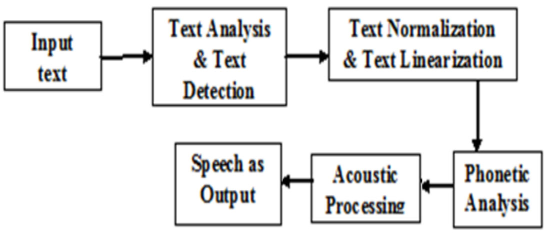

Fig. 3 Architecture of Speech Synthesizer

1 Text Analysis & Detection

The Text Analysis is part of preprocessing. It analyzes the input text and organizes into manageable list of words. Then it transforms them into full text. Text detection localizes the text areas from printed documents.

2 Text Normalization & Linearization

Text Normalization is the transformation of text into the form of pronounceable. Often text normalization is performed before the text is processed in some way, such as automated language translation or generating synthesized speech. The main purpose of this process is to identify punctuation marks and pauses between words. Mostly the text normalization process is done for converting all letters of lowercase or upper case, to remove punctuations, accent marks, stop words or "too common words "and other diacritics from letters.

3 Phonetic Analysis

It provides phonetic alphabets. The grapheme to phoneme conversion is done. It is actually a conversion of orthographical symbols into phonological symbols.

4 Acoustic Processing

It performs formant synthesis. It works intelligently and thus does not require any kind of database of speech samples. For speak out the text, it uses voice characteristics of a person.

VI PROPOSED METHODOLOGY

The scanned image is taken as input. The image should be scanned with at least 300 dpi for better recognition result. The image is then preprocessed like grayscale conversion, binary conversion and so on. The preprocessed image is then given to the recognition engine for extracting the text from the image. The extracted text is further converted into Braille format and audio. Figure 4 shows the overall architecture proposed in this system.

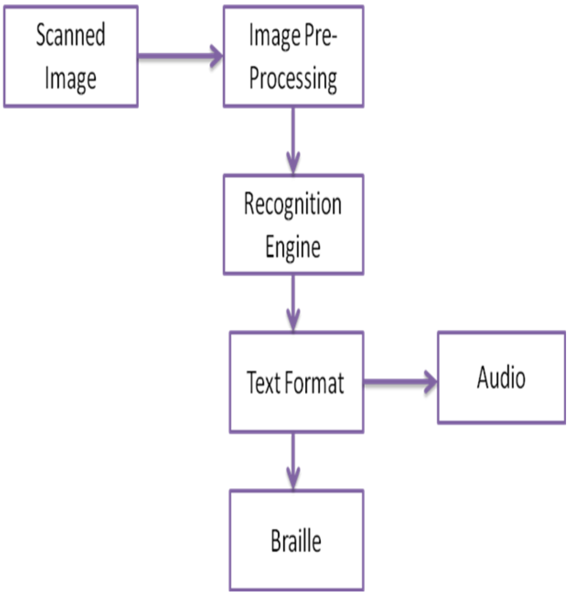

Fig. 4 Proposed Architecture

On using with the above architecture the process is taken in an effective way as expected. The process can be represent in an linear fashion as starting with the scanned image as it proceeds into the conversion of text followed by Braille finally the audio. The conversion of image to Braille is illustrated in the figure 5.

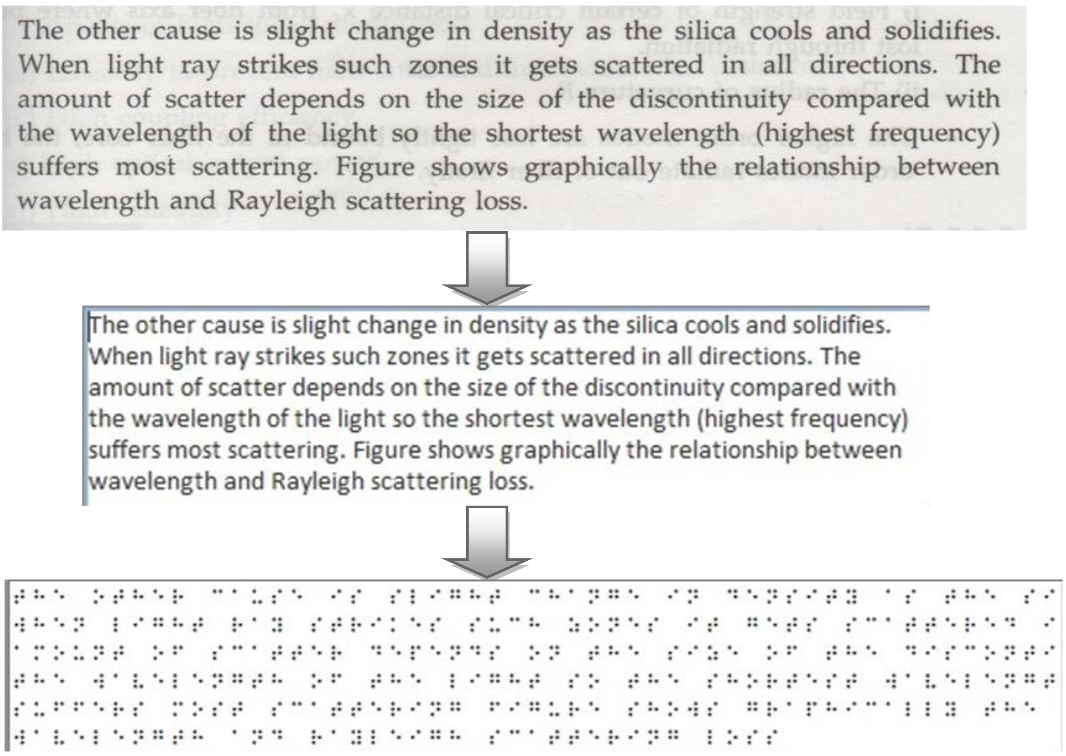

Fig. 5 Conversion of Image to Braille

VII RESULTS & CONCLUSION

A set of 10 pages is scanned and taken as the jpeg image format. Those images were scanned with different resolution for each images say 200, 300 and 400 dpi. All those images were given as input to the recognition engine and output is saved in individual file. The result is then tabulated as shown in table 1 with the total number of words in the image and those words which recognized incorrectly then the accuracy is calculated and represented in a graphical way for better observation.

Image ID	Total No. of Words	No. of Words recognized incorrectly			Accuracy		
		200 dpi	300 dpi	400 dpi	200 dpi	300 dpi	400 dpi
IM001	525	23	19	17	95.61905	96.38095	96.7619
IM002	469	15	12	10	96.80171	97.44136	97.8678
IM003	226	9	3	3	96.0177	98.67257	98.67257
IM004	241	36	13	10	85.06224	94.60581	95.85062
IM005	296	16	12	9	94.59459	95.94595	96.95946
IM006	292	7	6	5	97.60274	97.94521	98.28767
IM007	186	4	2	1	97.84946	98.92473	99.46237
IM008	287	10	7	5	96.51568	97.56098	98.25784
IM009	334	18	13	12	94.61078	96.10778	96.40719
IM010	349	29	27	24	91.69054	92.26361	93.12321

Table 1. Accuracy rates over various input images

The graph shown in figure 6 represents the accuracy level of all the images being taken as input. It is clearly seen that the accuracy level increases with the increase in the resolution. The overall accuracy is found to be 94.64% for the 200 dpi, 96.58% for 300 dpi and 97.17% for 400 dpi. On the experimental results the minimum resolution for scanning the image is found to be 300 dpi.

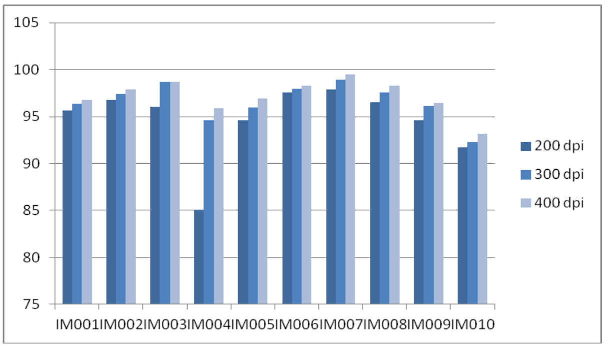

Fig. 6 Accuracy ratio between set of images

VIII FUTURE ENHANCEMENT

This system is capable of performing the OCR for English language and converting the English text to English Braille and audio for the same language. This work can be further enhanced to perform in several languages rather than only English and in short a multilingual system can be adopted.

IX REFERENCES

[1] Chirag Patel, Atul Patel, Dharmendra Patel, "Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study", International Journal of Computer Applications (0975 – 8887), October 2012

[2] Ray Smith, Daria Antonova, Dar-Shyang Lee, "Adapting the Tesseract Open Source OCR Engine for Multilingual OCR", ACM, 2009

[3] Ravina Mithe, Supriya Indalkar, Nilam Divekar, "Optical Character Recognition", International Journal of Recent Technology and Engineering (IJRTE), March 2013

[4] Bindu Philip and R. D. Sudhaker Samuel, "Human Machine Interface – A Smart OCR for the Visually Challenged", International Journal of Recent Trends in Engineering, November 2009

[5] R. Smith. "An overview of the Tesseract OCR Engine." Proc 9th Int. Conf. on Document Analysis and Recognition, IEEE, Curitiba, Brazil, Sep 2007, pp629-633.

[6] Dr.B.L.Shivakumar, M.Rajasena thipathi, "English to Braille Conversion Tool using Client Server Architecture Model", International Journal of Advanced Research in Computer Science and Software Engineering, August 2013

[7] William Walker, Philip Kwok, Paul Lamere, "FreeTTS Open Source Speech Synthesis"