TEXT RECITER FOR VISUALLY IMPAIRED

Nan Ahmed Siddiqui, Ali Akbar Siddiqui, Sadaf Salahuddin, M. Tahir Qadri

Abstract — Visual impairment refers to the defects of eye vision which can make it difficult for visually impaired people to see things and identify them as compared to normal persons or the capability of sight is disturbed in visually impaired people. Some can see things with little difficulty and some can be included in the category of blind people. So this was the purpose for us to build something for such people to gain knowledge with least limitation. Our project has the capability to assist the visually impaired to grow their academic qualifications and save themselves from fraud in any kind of documentation reading. Our proposed idea is synergism and implementation of two technologies OCR and espeak. The project is implemented quite easily by using a webcam which will be attached to the wrist of user through a ring type structure. The user just have to move its hand to the document which he/she wishes to read and the remaining task is done with the help of an algorithm (software) and a processor (hardware) to simultaneously carry out all the functions. The design is simple, easy to implement and the mandatory parts are software based.

Keywords: Optical Character Recognition (OCR), Tesseract, Text Reciter.

I. INTRODUCTION

Many people experience some type of visual problems in their lives. Some can no longer see objects far away. Others have problems reading small prints.

These types of conditions are often easily treated with eyeglasses or contact lenses. And some include color blindness.

But when one or more parts of the eye that are required to process images become damaged, severe or total loss of vision could occur.

In these types of cases, vision cannot be fully restored with medical treatment, surgery, or corrective lenses like glasses or contacts. In such matter a precise means of visual communication is required to better understand the surroundings.

Visual impairment is a term experts use to describe any kind of vision loss, whether it's someone who cannot see at all or someone who has partial vision loss.Some people are completely blind, but many others have what's called legal blindness. They haven't lost their sight completely but have lost enough vision that they would have to stand 20 feet from an object to see it as well as someone with perfect vision could see from 200 feet away.

Vanderheiden, Gregg C. has proposed an idea of "Touch screen for the vision-impaired"[1]. Hasser, Christopher J., and Marvin R. Roark proposed a "Tactile graphics display" for visually impaired computer users [2-4]. Noguchi, Atsushi has proposed an idea of "Hypertext control through voice synthesis" to aid the people is listing data without having to look at the screen or terminal [3-6]. Siegel, Steven H. proposed "Method for the auditory navigation of text" to point out the ways to help the people who are unable to read [4-10].

Our proposed system would allow visually impaired people to access books and documents that were previously beyond their reach and end their dependence on Braille system. It will also allow them to understand any printed text in more than 30 languages. Due to the cost effectiveness of this prototype system, it is within the reach of common man which was one of our primary goals. Experiments were first performed on PC platform helped us in identifying the resources best suited for an initial prototype. But as for now the complete system is equipped with a processor board and is fully functional, achieving the goal of recognizing text pointed by finger and then reciting it loudly to the user through a head phone. The headphone is part of hardware which is being used to deliver the output of the project. Overall components to test this project comprise of a

Manuscript Received 2-1-20 17; accepted: 2017; date of current version June- 2017
Noman Ahmed Siddiqui is withDepartment of Electronics Engineering, Sir Syed University of Engineering & Technology (email: noman_sid@yahoo.com)
Ali Akbar Siddiqui is with Department of Telecommunication Engineering, Sir Syed University of Engineering & Technology(email: ali124k@hotmail.com)
Sadaf Salahuddin is with Department of Electronics Engineering, Sir Syed University of Engineering & Technology (email: sadaf.salahuddin1@gmail.com)
M. Tahir Qadri is withDepartment of Electronics Engineering, Sir Syed University of Engineering & Technology (email: mtahirq@hotmail.com)

 ISSN 1816-613X

webcam, a printed document, processor and a headphone.

II. COMPARISON BETWEEN BRAILLE AND OUR SYSTEM

It may not sound so practical at first but it can be done by using the modern techniques along with extra efforts to make this idea works. To make our point more clear, we are giving a brief comparison of our idea with old method which is known as 'Braille'. Braille was difficult, time consuming and extra skills were required to learn through it. It didn't provide you any method to learn by yourself. In Braille, it is essential to use your sense of hearing as well as of touch, it gives the understanding of only specially written books for blind people. Our proposed system is easy to understand and no extra skills are required. It provides you a chance to learn by yourself without being dependent on anyone. It requires you to only use your sense of hearing. The overall cost of our system is very much less than the Braille system.

III. SYSTEM HARDWARE MODEL

In our proposed system, we have used Raspberry Pi to be our core processor. A video camera will acquire a frame from live video stream in which the used is pointing the finger at and the frame is passed to the on board processor for the execution. Raspberry Pi will identify the word from and correlate it with which is stored in its memory. After the identification the word identified will be send back to the onboard processor for speech pattern to recognize and finally be able to recite for the used through headphone.

The significance of our project can be understood by the imagination of how much change it can brought not in just an individual's life but to the whole society by allowing each and every one a fair chance irrespective of their disability to read and learn so that anyone can become a scientist, a doctor, an engineer or a Teacher even if he/she is visually impaired or even completely blind.

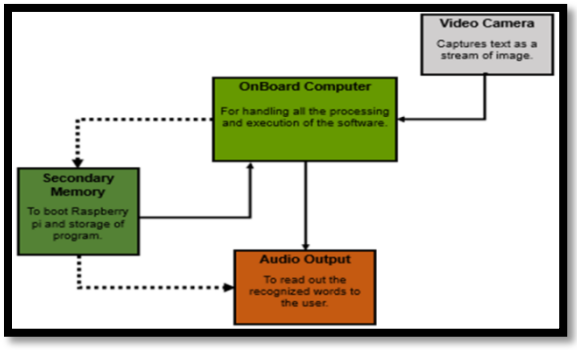

Figure.1**:** Block diagram of the system

IV. SOFTWARE MODEL

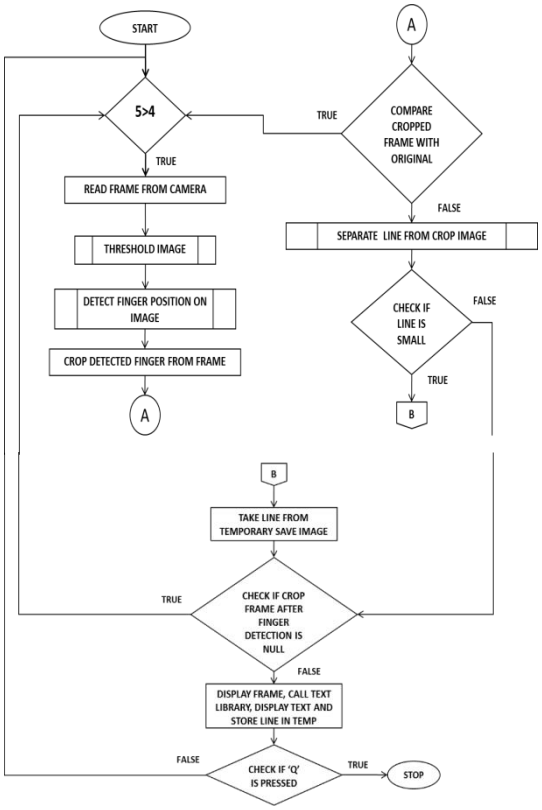

Figure.2: Flow diagram of the system.

Fig-2 represents the flow representation of our source code, the code start by first initializing camera to input video frames and Tesseract OCR Engine. The next step consists of a decision box that enters the program into an infinite loop after that we read the camera one frame at a time and call our finger detection function which is followed by the cropping of image just above the detected finger, both the original and cropped frame are compared for error detection in case of any error a new frame is captured, if there is no error present then line is extracted from the cropped image which contains the text being pointed by the user. The text line undergoes some error checking for the small line or whether it is empty or not, or even the line is suitable for character recognition. If all these error checks are passed then the data is outputted and the code continues to grab the next frame unless the button 'q' is pressed which will cause the code to exit. The button 'Q' is the ending of the sentence or the worse that the reader is pointing.

V. RESULTS

As the user will have to put his finger on the text which he or she wants to read so the program will first detect the finger and then the line that finger is pointing. Since the text will comprise of multiple lines of text so it is necessary for the user to reach

ISSN 1816-613X

to the first line then to the second and so on therefore our project is capable to separate the lines from each other and at a time it will process only one line which is being pointed by the user. The algorithm can detect the lines individually so it will identify the user when he/she reaches the first line that the first line is being pointed by him/her. If the user doesn't go to the start of line so it will never be easy to understand the meaning of whole line so the algorithm has the capability to detect the start of line so that the user can always find from where to start reading.

Obviously if the user know where the line is starting from, then he/she must be able to find that where the line ends to move to the next line and continue reading in a nice flow without any gap or hesitation that if there is something else on the line or not. The algorithm is capable of converting the images from one format to other format like jpeg, jpg, bitmap etc. Because the input images can be of any format and it will be an extra advantage to convert them before processing as the processing of different formats vary in levels, some are easy to process and some are complex. Text inputs can vary in different types so we had to focus on the processing of text very keenly and the techniques which can manipulate the text according to our needs. For this reason we use the software Tesseract which is responsible for all the handling of text data in our project and it is working more efficiently than we expected.

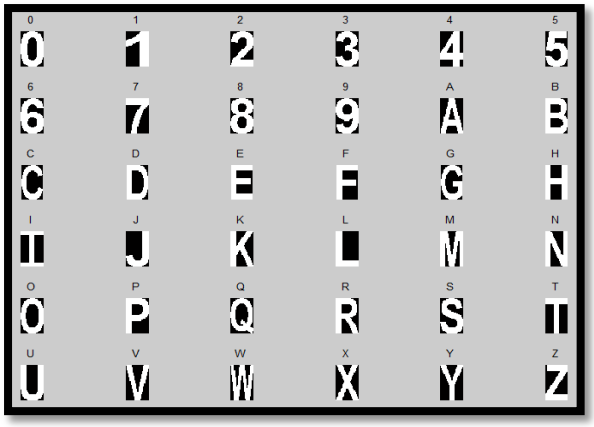

Figure.3: Reference Characters & Digits for OCR Algorithm

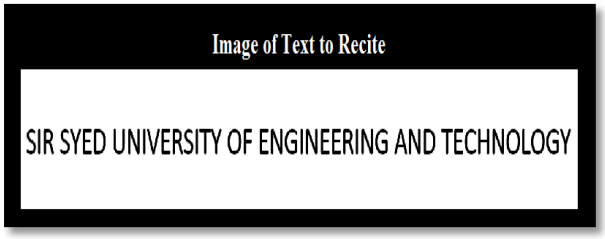

Figure.4: Complete Sentence to Recite

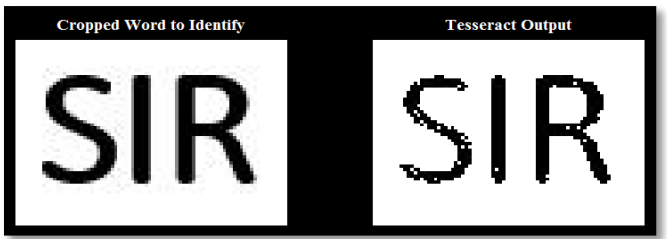

Figure.4: (a) Cropped Image from The Text (SIR). (b) Tesseract OCR Engine Output

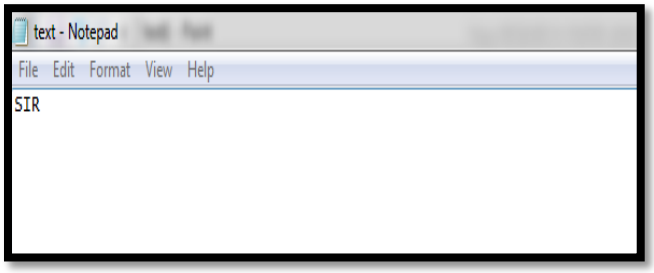

Figure.5: Identified Text using OCR, displayed on Notepad (SIR).

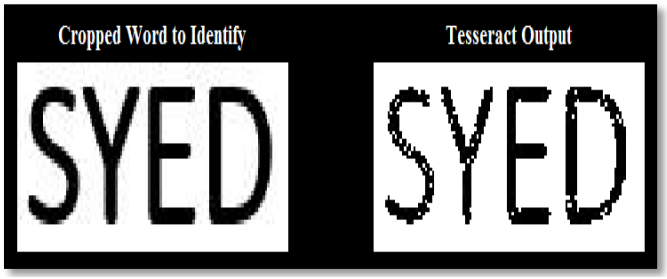

Figure.6: (a) Cropped Image from The Text (SYED). (b) Tesseract OCR Engine Output

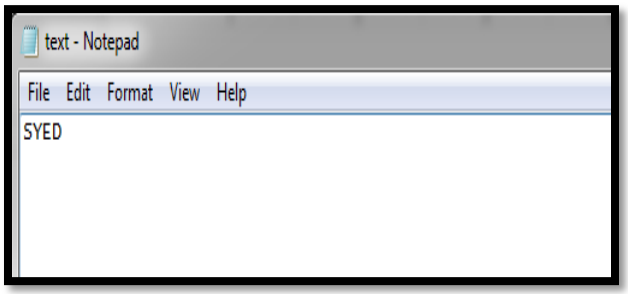

Figure.7: Identified Text using OCR, displayed on Notepad (SYED)

VI. CONCLUSION

Usually many ideas are based on Security or for luxuries for the people, but our proposed work is based to assist the common people who are blind or visually impaired. The reason behind this idea for us was the disability of visually impaired people who want to learn and make progress in their lives like others by educating themselves and learning different skills. So by help of this device the visually impaired can also learn and enhance

ISSN 1816-613X

their abilities to do work with normal people. The project is also able to recite the other languages book for example French, Urdu and more than 30 other languages.

Our software would allow such people to access books that were previously beyond their reach and end their dependence on the complex Braille system. It will also allow them to understand any printed text in more than 30 languages. Due to the cost effectiveness of this project, it is within the reach of common man which was one of our primary goals.

ACKNOWLEDGMENT

The authors would like to thank Sirsyed University of engineering & technology, Karachi, Pakistanfor their support in the completion of this research work.

REFERENCE

[1] Vanderheiden, Gregg C. "Touch screen for the vision-impaired." U.S. Patent No. 6,384,743. 7 May 2002.
[2] Hasser, Christopher J., and Marvin R. Roark. "Tactile graphics display." U.S. Patent No. 5,736,978. 7 Apr. 1998.
[3] Noguchi, Atsushi. "Hyper text control through voice synthesis." U.S. Patent No. 5,983,184. 9 Nov. 1999.
[4] Siegel, Steven H. "Method for the auditory navigation of text." U.S. Patent No. 7,181,692. 20 Feb. 2007.
[5] Andrews, H. C. and Hunt. B. R.[1977].Digital Image Restoration, Prentice Hall Englewood Cliffs, N.J.
[6] Andrews, H. C. [1970] Computer Techniques In Image Processing,AcademicPressNew York.
[7] Swaroop, C H.[August 4, 2013] “A Byte of Python” Creative Commons Attribution-ShareAlike 3.0.
[8] Maria Petrou and Costas Petrou.[January 2011].“Image Processing, The Fundamentals 2nd Edition”.
[9] Rafael C.Gonazles (University of Tennessee) and Richard E.Woods. [March 2003] “Digital Image Processing Third Edition”.
[10] Gary Bradski and Adrian Kaehler.[October 4, 2008] Computer Vision with the OpenCV Library 1st Edition. O'Reilly Media.

 ISSN 1816-613X