ISSN: 0975–9646

Bharati Ainapure et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (3) , 2016, 1195-1198

Printed Text Assimilator for Visually Impaired: Implementation

Bharati Ainapure, Nitin Sankpal , Pranav Ranadive, Rohan Shiroor, Satish Vairagar
B.E. Computer

Abstract **– We proposed a camera mounted assistive text assimilator model to help blind and dyslexia patients read printed text from a pervasive device objects in their day-to-day lives. To isolate the objects from the uneven colored backgrounds or other plain surroundings in the camera-perspectives, we first propose an effective and prominent static method to define Interested Region of Area (IRA) in the non-initialized video which will then be allowed to take a capture of the real-time text. In order to make it more accurate we have used a better design module which was named as Table CAM, help to take entire frame at a time and send to the Text-to-Speech module in order to make subject listen the assimilated text. Based on survey on various technological aspects we have implemented the design with working model. This paper demonstrates the usage along with all the required perspective of the implementation. The primary motive of this technology is precise scanning and text-to-speech transformations in order to achieve efficient output.**

Index Terms**: Assistive devices, Binarization, Braille, OCR-Optical Character Recognition, mobile devices and handheld devices, text reading, text-to-speech**

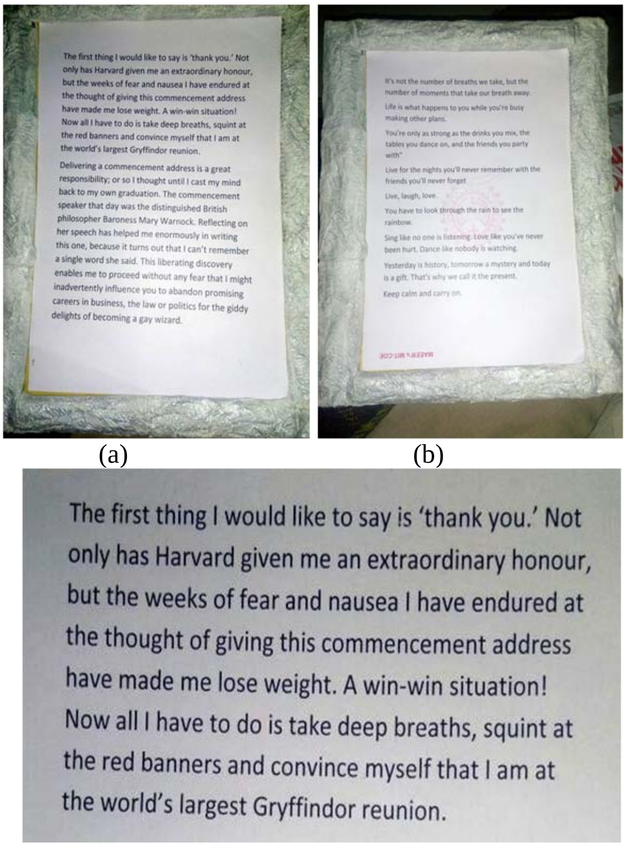
I INTRODUCTION

Over 85 million people are having visually impairments alone in India itself among them approximately 30 million are blind – that can affect their ability to perform activities of day-to-day living. Adults with visual impairments can continue to work and pursue a better range of careers because of the use of computers and other devices.

Currently there are a few systems that have some propose that can be useful for the same perspective model for portable use, but they cannot handle the data transfer rate that may help to produce a better result at the end user side. For example Assistive Text and product labeling for blind people which are able to read the absurd text on the products in a market cart. The foremost reason to make aside this type of implementation is that users need to shake the product first and only then it will turn that text into OCR and send to next module for processing. It may become a complex task for blind people to manually doing all of these tasks.

Accessing printed text in a digital and real world context is a major challenge for the visually impaired. A primary study with visually impaired people reveals numerous difficulties with existing technologies including problems with bordering, focus, accuracy and efficiency. People with a vision impairment also naturally based on tactile feeling through their fingers to read Braille and retrieve information about the environment. We are proposing a new approach: a printed text assimilator which will be outputted into audio streams so that it can be easily be audible to the blind people. This paper contains all the information about the proposed system with their benefits and drawbacks as well.

In assistive technologies introduced for blind persons, it is very challenging for users to position the object of interest as per the camera vision and its perspective which includes vertical view from the camera lens to the frame.



(a) (b) (c)

Figure 1. Examples of printed text inside the frame container with different light conditions.(a) At normal Light Conditions, (b) Low Light Condition, (c) Text to be considered.

We are more focused on the pervasive devices that can be assistive and somehow automated as it will be more convenient for visual impairments to rely on the device:

A. The first section will provide us with the detailed information about the system which we have been developed. The section gives an idea about how those systems will be helpful to the target people in order to provide a desired result.

B. Second section will describe about the binarization technique we have used in the system in order to make the input to the OCR more accurate to read and process.

C. At the last section paper will describes about the user end output section which will explain about the text to speech with the help of some of the examples we have considered.

II. PROPOSED SYSTEM

Figure 2. First prototype of the proposed system consisting of a computer system with OCR and Text-to-speech processor, and a table CAM with camera mounted on the top inside.

This paper presents a prototype system of printed text assimilator. Mentioned in the following Figure 2, the system model consists of some functional components: printed text processor, table CAM for capturing, box frame container and text-to-speech processor. The table CAM component captures the picture in the frame and sends it for processing. Next task will be binarization of image so that it will increase the accuracy. In our prototype, it corresponds to a camera mounted on the head of the table CAM. The printed text processor component is used for deploying proposed prototype algorithms, including 1) Otsu's algorithm or method for binarization of image, 2) Text-to-speech module will help to convert the text from OCR into audio streams that can be processed by the freetts.

As this paper is based on the assistive technology, we are considering only those aspects of the development in prototype model. This simple hardware configuration ensures the portability of the assistive text reading system. Figure 3. Depicting the overall working of the prototype system. A Frame sequence V is captured by camera mounted on the table CAM for blind users, which focused on the box frame vertically below to the table CAM consists of sample printed text papers. To extract text information from the box frame printed text, static object provides convenience to handle precise capture of Interested Region of Area (IRA) S by just placing the printed text paper inside frame.

$$S = \frac{1}{|V|}\sum_{i} R(V_i, F)$$

Where V_i denotes the i-th frame in the frame box captured sequence, $|V|$ denotes the number of frames, F denotes the foreground color from the camera perspective, and R depicts the calculated foreground object at each frame in the box frame. Interested Region of Area is determined by the average of foreground masks(see details in next sections). [3] Preliminary algorithms are efficient and reasonably accurate since they provide more intuitive and precise control over scanning and text-to-speech; increased spatial understanding of the text layout; reduced camera framing, focus, and lighting issues.

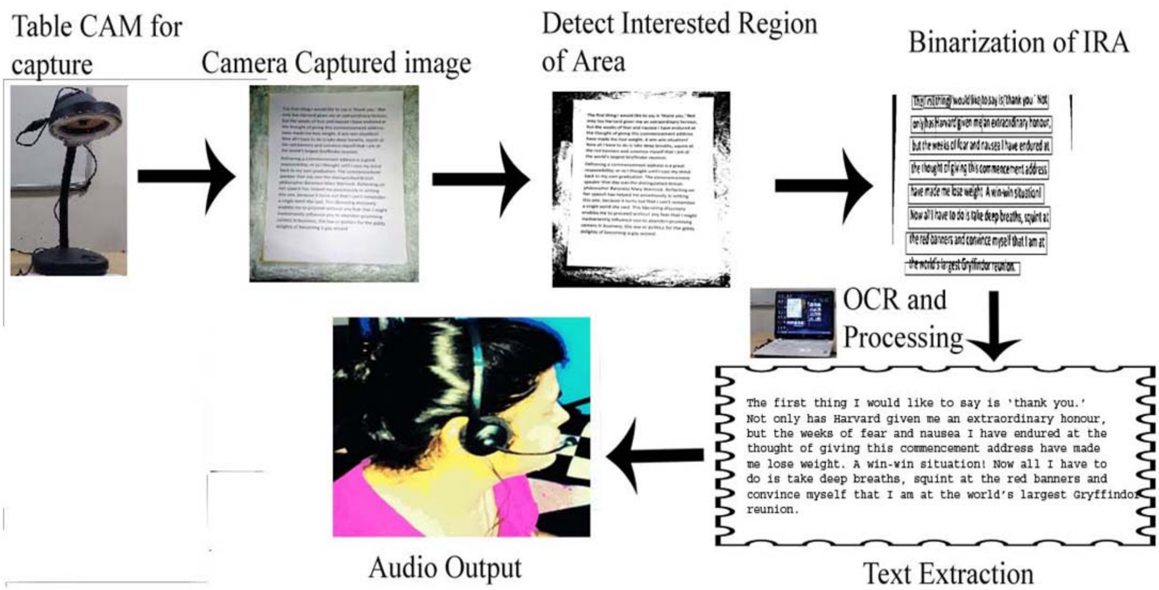
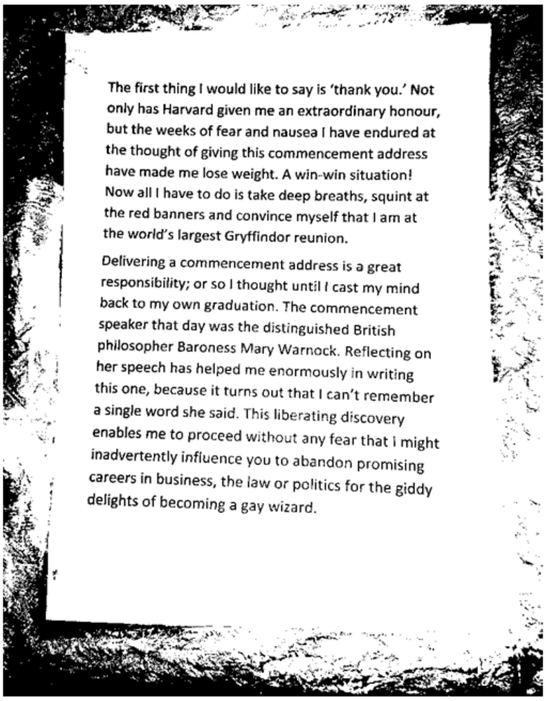


Figure 3. Flowchart for proposed prototype model for extracting text and read it in audio streams.

III. BINARIZATION TECHNIQUE

In order to increase the accuracy of the reading text by OCR we have binarized the image using otsu's method or algorithm. In Otsu's method we need to search for the threshold value that will minimizes the class variance within that class. Operation must be used to detect edges inside an image and edges are very much critical task in image manipulation. So, to binarize the image, we have used Otsu's method which is based on class variance. The algorithm maintains typically two classes of pixels of bi-modal histogram and then it will help to calculates the optimum threshold value which separates those two classes so that their combined internal class variance is minimum.


The first thing I would like to say is 'thank you.' Not only has Harvard given me an extraordinary honour, but the weeks of fear and nausea I have endured at the thought of giving this commencement address have made me lose weight. A win-win situation! Now all I have to do is take deep breaths, squint at the red banners and convince myself that I am at the world's largest Gryffindor reunion.

Delivering a commencement address is a great responsibility; or so I thought until I cast my mind back to my own graduation. The commencement speaker that day was the distinguished British philosopher Baroness Mary Warnock. Reflecting on her speech has helped me enormously in writing this one, because it turns out that I can't remember a single word she said. This liberating discovery enables me to proceed without any fear that I might inadvertently influence you to abandon promising careers in business, the law or politics for the giddy delights of becoming a gay wizard.

Figure 4. Binarized image after capturing from table CAM

The main advantage of it will be 1) Speed, as we are using only histograms and arrays of small length) and 2) easiness of the implementation. The first thing to binarize an image is to convert it to gray scale by using luminance method. In order to convert the image to gray scale we have used one of the simplest method of averaging. We need to take all the color components i.e Red, Green and Blue and divide them by number of color types i.e. three

$$p_{new} = \frac{R + G + B}{3}$$

Where p_{new} is the new pixels in the image which are calculated by reducing the intensity of R, G, and B. After converting image to gray scale, we need a threshold value. This can be done manually, but with this approach we normally do not get the accuracy as such. Otsu's method will finds that binary threshold that we are looking for. This will take the iterations among possible threshold values and calculate measure of spread for the pixel. Finds out the pixels is whether fall in background or foreground.

IV. END USER AND AUDIO OUTPUT

The automation of IRA detection and text optimization were independently examined and calculated as unit tests to ensure effectiveness and robustness of entire proposed system.

The Blind person will seat before the table CAM and bring the printed text paper near our frame that will make the triggering device to activate and send trigger signal to the computer program so that camera will start capturing image and send it to processor i.e. OCR processor. The resolution of the captured image is 640 × 480. There were two conditions we have added for testing and it is giving us different stimulation of camera input. The user need to put the object inside the box frame and all remaining work will be done by the table CAM equipped with camera. Text recognition is performed with the help of open source off-the-shelf OCR. However, our experiments show that OCR generates better accuracy with the text greater than 18 pixels font size, also increase the performance to send it over text-to-speech module.

The recognized text codes are stored temporarily into the script file. Then we employ the freetts java API to load these files and display the audio streams output of text information. Blind users can now listen speech at specific rate, volume and tone.

The focus of the our method is on large text found in outdoor environments, such as story books, bills with large fonts.

V. CONCLUSION

In this paper, we have described a prototype model to read printed text assimilator on the assistive devices for visually impaired, more focused on blind people. In order to solve the common aiming problem for blind users, we have proposed a static method with the use of table CAM similar to table lamp. User needs to just put their object inside the box frame and all other task will be done by the table CAM. This method can effectively distinguish the interested region of area from background and binarization removed the problem of reading multi-color text image. To extract text regions from multi-color foreground we have proposed a novel text binarization method with the help of Otsu's method. Open source off-the-shelf OCR is used to perform word recognition on the interested text region and transform them into audio stream output for blind users.
Our Future work will extend our algorithm for increase accuracy of OCR and audio streams. We will also work on the text font size below 18 pixels with more robust block patterns for text extraction. We will also make the automation of the box frame so that user does not need to move the object manually each time.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their constructive and insightful reviews that improved the quality of this system prototype.

REFERENCES

[1] "Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Persons ", Chucai Yi, *Student Member*, *IEEE*, YingLiTian, *Senior Member*, *IEEE*, Aries Arditi ,IEEE 2013

[2] "The Design and Preliminary Evaluation of a Finger-Mounted Camera and Feedback System to Enable Reading of Printed Text for the Blind", Lee Stearns1, Ruofei Du1, Uran Oh1, Yumeng Wang2, Leah Findlater3, Rama Chellappa4, Jon E. Froehlich, 2011

[3] C. Yi and Y. Tian, "Text string detection from natural scenes by structure based partition and grouping," In *IEEE Transactions on Image Processing*, 2011

[4] D. Dakopoulos and N. G. Bourbakis, "Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey," IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev., vol. 40, no. 1, pp. 25–35, Jan. 2010.

[5] D. Morris, T. S. Saponas, and D. Tan, "Emerging input technologies for always-available mobile interaction," Found. Trends Human--Computer Interact., vol. 4, no. 4, pp. 245– 316, 2010.

[6] T. S. Saponas, "Supporting Everyday Activities through Always-Available Mobile Computing," University of Washington, 2010.

[7] "10 facts about blindness and visual impairment", World Health Organization: Blindness and visual impairment,2009. www.who.int/features/factfiles/blindness/blindness_facts/en/index.htm

[8] Advance Data Reports from the Nation al Health Interview Survey,2008 .http://www.cdc.gov/nchs/nhis/nhis_ad.htm