

# Text Detection and Recognition with Speech Output for Visually Challenged Person

<sup>1</sup>Ms.Rupali D. Dharmale, <sup>2</sup> Dr. P. V. Ingole

Department of Electronics and Telecommunication Engineering,  
G.H.Raisoni College of Engineering and Management, Amravati.

## Abstract

*Reading text from text images and text boards is a difficult task for visually challenged persons. The proposed system reads the text encountered in scene images and text boards with the aim to provide assistance to the visually challenged persons. In this work, an approach has been attempted to extract and recognize text from scene images and convert that recognized text into speech. An android based demo system is developed to show the effectiveness of our proposed system of text detection and recognition with speech output for visually challenged person.*

**Keywords:** Android, OCR, Mobile application, Visually challenged person, Text to speech (TTS).

## 1. Introduction

Every year, the number of visually challenged persons are increasing due to eye diseases, age related causes, uncontrolled diabetes, traffic accidents and other causes. Mobile applications that give support to the visually challenged persons have become an necessary device in visually challenged person's life. Recent development in mobile, computer vision, availability of digital cameras make it feasible to assist these visually challenged person by developing camera based applications that combine computer vision tools with other existing beneficial products such as OCR system. Reading is obviously necessary in today's society. The main focus of our project is that visually challenged person can get text information which is present in the text boards, instructions on traffic sign boards and hoardings in audio form. With this point of view, the application design for a camera based reading system that remove text from text board and identify the text characters and strings from the captured image and finally, text will be converted into audio. This system allows the user to photograph a text image such as Stop sign board on road side and click on speak button and hear the text aloud and gives real time feedback.

Extracting text information from captured image is difficult task due clutter background and non text outliers [1], and further, text consists of dissimilar words where every word may contain different characters in a variety of fonts, styles, and sizes, resulting in large intra-variations of text patterns. Even if the problems of clutter background and text segmentation were to be ignored for the moment, there are several other reasons such as (i) font style and thickness; (ii) background as well as foreground color and texture; (iii) camera position which can introduce geometric distortions; (iv) illumination and (v) image resolution. Optical character recognition is the electronic conversion of photographed images of printed text into computer-readable text. A text-to-speech system converts normal language text into speech. A Text-to-speech (TTS), as it is called, is usually meant to help visually challenged people.

## 2. Literature Review

In this section, we present some earlier work for supporting visually challenged persons with text to speech technology. A number of reading assistance system have been designed for visually challenged [2], [3], [4], [5], [7], [8]. Hangrong Pan et al. designed a computer vision-based system to detect route number of bus and recognize information of the coming bus from images captured by a camera at a bus stop [6]. Sherine Sebastian et al. proposed a method to detect panels and to recognize the information inside image for blind persons accessing unfamiliar environments [9].

## 3. Implementation Details

In this section, we introduce the implementations details. The below figure 3.1 shows the GUI which process all steps of text detection and recognition with speech output.



**Figure 3.1:** Main GUI Layout



**Figure 3.2:** Captured Image from Camera

This is the first step of text detection and recognition process. The input image is given to further processing.



**Figure 3.3:** Saved Image

The text detection, recognition and extraction is done with the help of Tesseract OCR algorithm. The Tesseract OCR algorithm has the following steps:

- 1] Input image is a grayscale or color image.
- 2] Adaptive thresholding performs the reduction of grayscale image to binary image. In an image, there are black pixels and white pixels.
- 3] Connected-component labelling - Tesseract searches throughout the image, identifies the foreground pixels, and marks them as blobs or possible characters.
- 4] Line finding - lines of text are found by analyze the image space adjacent to possible characters. This does a Y projection of the binary image and finds locations having a pixel count less than a particular threshold. These areas are potential lines, and are further analyzed to confirm.
- 5] Baseline fitting - finds baselines for each of the lines. After each line of text is found, it examines the lines of text to find fairly accurate text height across the line. This process is the first stage in determining how to identify characters.
- 6] Fixed pitch detection - the other part of setting up character detection is finding the fairly accurate character width. This allows for the correct incremental removal of characters as Tesseract walks down a line.
- 7] Non-fixed pitch spacing delimiting- characters that are not of the same width or of a width that agrees with the surrounding region are reclassified to be processed in an alternating manner.
- 8] Word recognition - after finding all of the possible character blobs in the document, it does word recognition word by word, on a line by line basis. Words are then passed through a contextual and syntactical analyzer which ensures correct recognition.

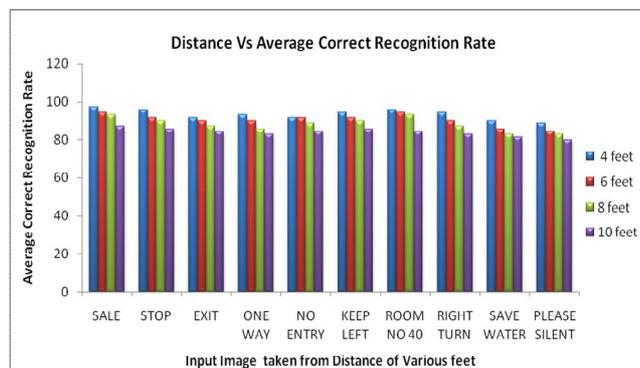


**Figure 3.4:** Recognized and Extracted Text

Android TTS API (Application Programming Interface) is used for the text to speech conversion to produce speech synthesis within the application, allowing the mobile device to produce speech output.

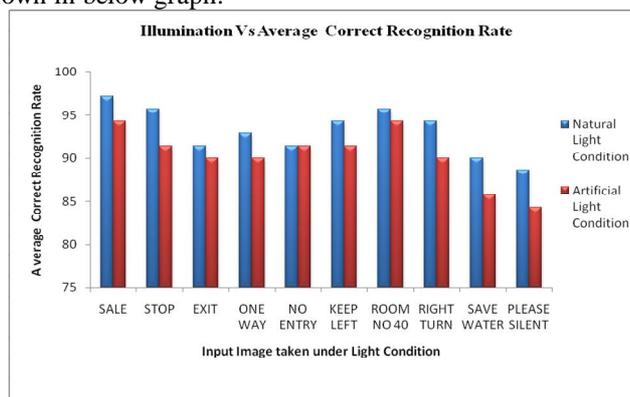
#### 4. Experimental Results and Discussion

Experiments were performed to test the text detection and recognition system which is developed as an application on android mobile phone. Test data used in the experiments consists of pictures of text boards taken at surroundings which include scenes of different text boards. All the test images were taken by Samsung Galaxy GT-I8262 embedded five megapixel camera. We photographed a set of images on the basis of distance and light condition and evaluated the performance of system. All of the text strings in captured images are in horizontal and of same size and same font. We captured the image from the four various distances such as from 4 feet distance, 6 feet distance, 8 feet distance, 10 feet distance and observed the results. For this we had taken 10 test images and performed experiments 70 times on one image and calculated the percentage of correct recognition for each text image. The percentage of correct recognition is calculated as the ratio between the number of correct recognition and number of experiments.



**Figure 4.1:** Distance Vs Average Correct Recognition Rate

Here we captured the text image by mobile camera from distance of 4 feet under natural light condition and artificial light condition. For this we had performed experiments 70 times on one image and calculate the percentage of correct recognition. The results are shown in below graph.



**Figure 4.2:** Illumination Vs Average Correct Recognition Rate

## 5. Conclusion

We have developed a system of text detection and recognition with speech output on Android mobile platform. Compared with a PC platform, the mobile phone platform is portable and more suitable to use. The input image is captured from mobile camera and extracts text information from image. It detects text regions from captured image and recognizes text. Then obtained text is converted to speech which can provide assistance to visually challenged people. This system is implemented as demo system of text detection and recognition which has audible output on mobile Samsung Galaxy GT-I8262 device with Android version 4.1.2. The designed text detection and recognition system with speech output for visually challenged person found to be highly useful and comfortable for visually challenged person. In future work it is proposed to improve the accuracy rate of text detection and text recognition by improved algorithms. It is possible to extend this work to handle non-horizontal text strings.

## Referances

- [1] Chucai Yi, Yingli Tian, "Scene Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration," IEEE Transactions on Image Processing, Vol. 23, No. 7, July 2014.
- [2] Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, "Improved text-detection methods for a camera-based text reading system for blind persons," IEEE in Proceedings of Eighth International Conference on Document Analysis and Recognition, pp. 257 – 261, Vol.1, ISSN: 1520-5263, 2005.
- [3] Marwan A. Mattar, Allen R. Hanson, Erik G. Learned-Miller, "Sign Classification for the Visually Impaired," IEEE Workshop on Computer Vision Applications for the Visually Impaired, 2005.
- [4] Kumar J.A.V., Visu A., Raj M.S., Prabhu M.T., Kalaiselvi V.K.G., "A pragmatic approach to aid visually impaired people in reading, visualizing and understanding textual contents with an automatic electronic pen," IEEE International Conference on Computer Science and Automation Engineering (CSA), Page(s): 623-626, Vol.4, 2011.
- [5] Oi-Mean Foong and Nurul Safwanah Bt Mohd Razali, "Signage Recognition Framework for Visually Impaired People," International Conference on Computer Communication and Management Proc .of CSIT vol.5 IACSIT Press, Singapore, 2011.
- [6] Hangrong Pan, Chucai Yi, Yingli Tian, "A primary travelling assistant system of bus detection and recognition for visually impaired people," IEEE International Conference on Multimedia and Expo Workshops (ICMEW), San Jose CA, Page(s): 1-6, 15-19 July 2013.
- [7] Adil Farooq, Ahmad Khalil Khan, Gulistan Raja, "Implementation of a Speech Based Interface System for Visually Impaired Persons," Life Science Journal, 2013.
- [8] Chucai Yi, Yingli Tian and Aries Arditi, "Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for Blind Persons," IEEE/ASME Transactions on Mechatronics, Vol. 19, No. 3, pp. 808, June 2014.
- [9] Sherine Sebastian and Priya S., "Text Detection and Recognition from Images as an Aid to Blind Persons Accessing Unfamiliar Environments," Asian Research Publishing Network (ARPN) Journal of Engineering and Applied Sciences, ISSN 1819-6608, Vol. 10, No. 17, September 2015.

## AUTHORS



**Ms. Rupali D. Dharmale** received B.E. degree in Electronics and Tele comm. engineering from Sant Gadge Baba Amravati University, Maharashtra, India, in 2013. She is currently pursuing the M.E. degree in Electronics and Telecommunication at G.H. Rasoni College of Engineering and Management, Amravati, Sant Gadge Baba Amravati University India.



**Dr. Prashant V. Ingole** received his B.E. degree in 1989 in Electronics Engineering and M.E. degree in Automatic Control and Robotics from College of Engineering Badnera, Amravati University, Amravati and Faculty of Engineering and Technology M.S. University Baroda respectively. He received his PhD degree in 2010 from Visvesvaraya National Institute of Technology, Nagpur in the area of Medical Image Processing and Image Retrieval. He worked as Lecturer, Assistant Professor and Professor at P.R.M.I.T. & R., Badnera (Formerly College of Engineering, Badnera). Currently, he is working as Principal of G.H. Rasoni College of Engineering and Management, Amravati. He has more than 30 research papers in his credit in National and International Conferences and Journals. His research interests include Digital Signal Processing, Soft Computing, Image Processing, and Embedded System. He is a fellow of IETE and IEI and a Life Member of ISTE.