

Text Detection and Recognition with Speech Output in Mobile Application for Assistance to Visually Challenged Person

¹Ms.Rupali D. Dharmale, ² Dr. P. V. Ingole

Abstract- Now a day's use of mobile is broadly increased every person possesses a mobile phone, in which lots of application run. Using Android mobile phones we can help the visually challenged people by providing easy reading of text boards and printed text information in the form of audio. Reading text from printed text images and text boards is a challenging task for visually challenged persons. The proposed system extracts and recognizes text from scene images and converts that recognized text into speech. This application is very helpful and handy for visually challenged person. The novelty of this work is to convert image containing text into speech.

Index Terms- Android, OCR, Text reading, Text to audio conversion, visually impaired person.

1. Introduction

The number of visually impaired persons is increasing due to uncontrolled diabetes, age related causes, eye diseases, traffic accidents, and other causes. Cataract is leading cause of blindness and visual impairment. Mobile applications that provide the support to visually challenged person have become an essential device in visually challenged person's life. Recent advances in mobile technology, digital camera, computer vision and camera based application make it possible to support visually challenged persons by developing camera based application that combine computer vision with other existing technology such as optical character recognition (OCR) system. With the

rapid development of camera-based applications on smart phones and handy devices, understanding the pictures taken by these devices has gained growing attention from the computer vision community in recent years which will be helpful for these individuals. The main focus of our research is that the visually challenged person can get information about printed text, text boards, scene text, hoardings, and instructions on traffic sign boards in audio form. With this point of view, the system design for a camera based reading system that extract text from textual board and identify the text characters from the captured image and finally, textual information will be converted into speech.

To detect text information from image, there exist many practical difficulties, such as non-uniform backgrounds, due to the large variations in character font, size, texture, color, background, orientations, and many other reasons. Text detection from scene/text camera images is possible due to high resolution camera. For extracting text information from image, algorithms are required. However extracting text information from captured text image is difficult due to two main factors: 1) jumbled backgrounds with noise, text and non-text part. 2) Random text patterns such as character, fonts, sizes etc [1]. The frequency of occurrence of text in image is very little, and limited number of text characters is separate from background outliers. To solve these difficult problems captured image text is divided into two processes: text detection and text recognition. Text detection is used to detect image region containing text characters. It aims to take out non- text background outliers [3]. Text recognition is to convert pixel based text into readable code. Optical character recognition is the electronic conversion of images captured by a digital camera of printed text into readable text. OCR has a good performance when recognizing machine-printed text in camera-based document analysis. Optical Character Recognition, or OCR, is a technology that converts different types of printed documents, such as scanned paper

documents or images captured by a digital camera into readable data.

2. Previous Work

In this section, we present some previous research works for assisting visually challenged people with text to speech technology. A number of handy reading assistants have been designed specifically for the visually challenged [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. Michael R.T.F. et al. proposed a system which operates the mobile devices without using the keypad [12]. In [14], a camera-based assistive text reading system to read text labels and product packaging from hand-held objects. Text detection is to detect regions in an image that contain text characters. [1]. Methods of feature descriptor can broadly be classified as Histogram of the oriented gradient (HOG) descriptor, Scale invariant feature transform (SIFT), Speeded up robust features (SURF), Gradient location and oriented histogram (GLOH) [1]. These are very popular feature descriptor used in computer vision and image processing for the purpose of object detection.

3. Proposed Methodology

In literature review, various text to speech systems are discussed for visually challenged persons but there exists some limitations. The objective of this work is

1. To model systems which could perform text to speech conversion for visually challenged person.
2. To study existing techniques for text detection and text recognition from scene images/text boards.
3. To evaluate accuracy of text to speech conversion.

The flowchart for proposed methodology is shown in following figure.

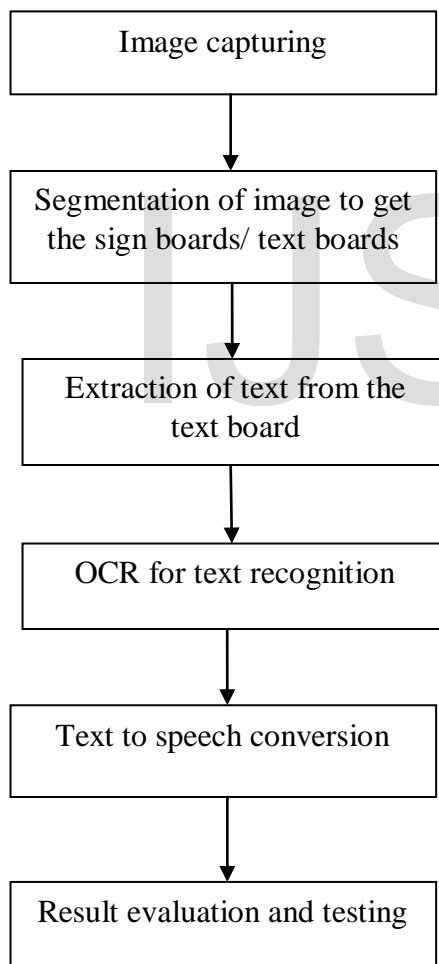


Figure 3.1: - Flowchart for proposed methodology

Initially, an input image is captured from camera. Then a segmentation algorithm is applied to perform segmentation of the desired part of image from sign boards/text boards. Image segmentation is an essential process for most image analysis techniques. Then Extraction of text from text board is performed by using image processing technique. The text recognition is proposed to be done by optical character recognition. Optical character recognition (OCR), is the electronic conversion of photographed images of typewritten or in print text into computer-readable text. Then obtained text is converted into speech which is output and finally result evaluation and testing, as shown in figure 3.1.

4. Implementation

Android is an open source and Linux-based operating system targeted for mobile devices such as smart-phones and tablet computers. Applications are generally developed in Java programming language using the Android software development kit (SDK). If used correctly, the SDK, together with Eclipse (the officially supported IDE) and JDK (Java Development Kit) is capable to deliver modern software for Android devices. The Android SDK (software development kit) provides the API (Application programming interface) libraries and developer tools essential to build, test, and debug apps for Android. This application is created in

eclipse. Android Development Tools (ADT) is a plug-in for the Eclipse integrated development environment (IDE) that is designed to provide an integrated environment for Android application.

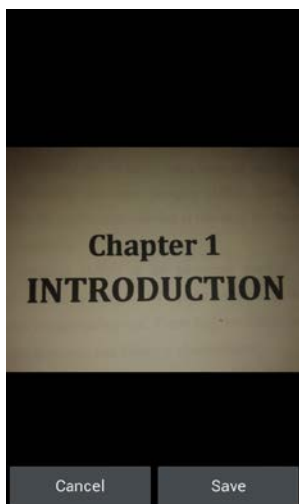
5. Experimental Result

Results are shown in following Screenshots.



Screen shot 5.1: Main GUI layout

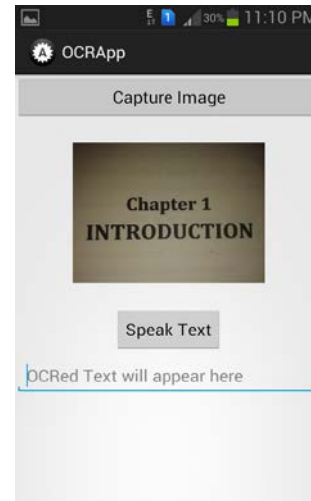
The above screen shot 5.1 shows main screen on mobile.



Screen shot 5.2: Captured image from camera

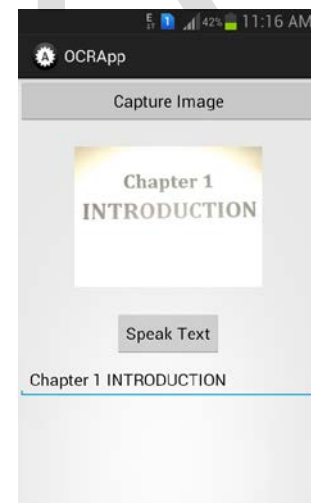
The above screenshot 5.2 shows the captured image from camera to by android mobile phone.

The first step is capture image from camera and save the image.



Screen shot 5.3: saved image

The above screen shot 5.3 shows saved image. In this step image processing will be done for retrieve the text from image. Here the text and non text background will be separated.



Screen shot 5.4: Text to speech conversion

The above screen shot 5.4 shows text to speech conversion. In this step the text appear in text box and after that the obtained text is audible.

Here obtained text is “Chapter 1 INTRODUCTION” it is voiced as output by proposed system.

6. Conclusion

A text detection and recognition with speech output system was successfully demonstrated on Android platform. This system is very handy and useful for the visually challenged persons. Compared with a PC platform, the mobile platform is portable and more convenient to use. This system will be helpful for visually challenged persons to access information in written form and in the surrounding. It is useful to understand the written text messages, warnings, and traffic direction in voice form by converting it from Text to voice. It is found that this system is capable of converting the sign boards and other text into speech.

References

- [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] J. Zhang and R. Kasturi, “Extraction of text objects in video documents: Recent progress,” in *Proc. 8th IAPR International Workshop DAS*, pp5-17, Sep. 2008.
- [3] C. Yi and Y. Tian, “Text string detection from natural scenes by structure-based partition and grouping,” *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2594–2605, Sep. 2011.
- [4] P. Blenkhorn, D.G. Evans “Using speech and touch to enable blind people to access schematic diagrams” science direct, *Journal of Network and Computer Applications*,1998.
- [5] Hideyuki Yoshida, Toshiki Kindo, “A newspaper reading out system with an adaptive information atering technology to support visually impaired people”, *IEEE*, 1999.
- [6] 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.
- [7] Shehzad Muhammad Hani, Lionel Prevost “Texture Based Text Detection in Natural Scene Images: A Help to Blind and Visually Impaired Persons”, *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Assistive Technology for All Ages CVHI*, 2007.
- [8] 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.

[9] Oi-Mean Foong and Nurul Safwanah Bt Mohd Razali, "Signage Recognition Framework for Visually Impaired People", 2011 International Conference on Computer Communication and Management Proc .of CSIT vol.5 IACSIT Press, Singapore,2011.

[10] Krishnan K.G., Porkodi C.M., Kanimozhi K. "Image recognition for visually impaired people by sound". IEEE International Conference on Communications and Signal Processing (ICCSP), Melmaruvathur, Page(s):943 – 946, 3-5 April 2013.

[11] 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.

[12] Michael R.T.F. ,RajaKumar B. , Swaminathan S.,Ramkumar M. "A novel approach: Voice enabled interface with intelligent voice response system to navigate mobile devices for visually challenged people",IEEE International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT), Tiruvannamalai, Page(s): 1 – 4, 7-9 Jan. 2013.

[13] Adil Farooq, Ahmad Khalil Khan, Gulistan Raja, "Implementation of a Speech

Based Interface System for Visually Impaired Persons", Life Science Journal, 2013.

[14] 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.

IJSER