RESEARCH ARTICLE OPEN ACCESS

Text Detection and Recognition with Speech Output for Visually Challenged Person: A Review

[1]Ms.Rupali D. Dharmale, [2]Dr. P. V. Ingole
[1] (Department of Electronics and Telecommunication, G. H. Raisoni college of Engineering and Management, Amravati, Maharashtra, India)
[2](Principal G. H. Raisoni college of Engineering and Management, Amravati, Maharashtra, India)

ABSTRACT
Reading text from scene, images and text boards is an exigent task for visually challenged persons. This task has been proposed to be carried out with the help of image processing. Since a long period of time, image processing has helped a lot in the field of object recognition and still an emerging area of research. The proposed system reads the text encountered in images and text boards with the aim to provide support to the visually challenged persons.
Text detection and recognition in natural scene can give valuable information for many applications. In this work, an approach has been attempted to extract and recognize text from scene images and convert that recognized text into speech. This task can definitely be an empowering force in a visually challenged person's life and can be supportive in relieving them of their frustration of not being able to read whatever they want, thus enhancing the quality of their lives.
Keywords: Optical character recognition (OCR), Text detection and recognition, Text to speech conversion, Visually challenged person.

I. INTRODUCTION

Every year, the number of visually challenged persons is increasing due to eye diseases diabetes, traffic accidents and other causes. Therefore applications that provide support to the visually challenged persons have become an important tool. Recent developments in computer vision, digital cameras, and computers make it possible to assist these persons by developing camera-based products that merge computer vision technology with other existing beneficial products such as optical character recognition (OCR) systems. When a visually challenged person is walking around, it is important to get text information which is present in the scene/text boards. Reading is obviously necessary in today's society. Printed text is all over in the form of reports, receipts, bank documents, restaurant menu cards, classroom handouts, product packages**,** instructions on medicine bottles, etc. As an important form of communication, text is widely used in our daily life. For example, different sign boards, directions, shop names etc contain textual and/or symbolic information that is perceived by a human being to facilitate knowledge of environment and perhaps also help in his navigation. The need to read textual and/or symbolic information becomes essential in the case of blind or visually challenged persons. With this point of view, the system which detect the text from textual/symbolic board and recognize the text characters from the captured scene text image and finally, textual and/or symbolic information will be converted into speech.

To extract text information from image, text detection and recognition algorithms are necessary. However extracting scene image's text is a not easy task due to two key factors: 1) cluttered backgrounds with noise and non-text outliers, and 2) diverse text patterns such as character types, fonts, and sizes. The frequency of occurrence of text in scene image is very small, and a limited number of text characters are embedded into difficult non-text background outliers [1]. However, it is difficult to model the structure of text characters in scene images due to the lack of discriminative pixel-level appearance and structure features from non-text background outliers. Further, text consists of different words where each word may contain different characters in various fonts, styles, and sizes, resulting in large intra-variations of text patterns. To solve these difficult problems, scene text extraction is separated into two processes [2]: text detection and text recognition.

Detection of text and classification of characters in scene images is a challenging visual recognition difficulty for visually challenged people. Text detection is used to localize image regions containing text characters and strings. It aims to remove most non-text background outliers[3]. Text recognition is to convert pixel-based text into readable code. It aims to accurately distinguish different text characters and properly composed text words.

OCR is optical character recognition. OCR is used to recognize words**.** It can recognize characters, words and sentences without any mistakes, making the software more perfect. Also, OCR software has a high rate of recognition. OCR is the electronic conversion of photographed images of typewritten or printed text into computer-readable text.

A text-to-speech (TTS) system converts normal language text into speech. A Text-to-speech (TTS), as its called, is usually meant to help visually challenged people.

II. PREVIOUS WORK

2.1 Existing systems for blind and visually challenged persons incorporating text detection.

In this section, we present a general review of previous work on text detection and recognition. There exists some research works for helping visually challenged people with text to speech technology. A number of portable reading assistants system have been designed particularly for visually challenged persons.

P. Blenkhorn, D.G. Evans implemented a computer-based system that allows blind users to read, create and edit one type of schematic diagram, specifically data flow diagrams used in software engineering, is presented, together with the mapping from the original diagram to a suitable generic, tactile diagram [4]. Hideyuki Yoshida, Toshiki Kindo presented a newspaper reading out system to support visually impaired people. They built an adaptive newspaper reading out system that is sorting headlines in order of user's priority. The system consists of three modules: information filtering module, speech recognition module and text-to-speech synthesis module [5]. Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, implemented a system that reads the text encountered in natural scenes with the aim to provide support to visually impaired persons. This paper describes a novel text-detection method geared for small text characters. This method uses Fisher's discriminant rate (FDR) to make a decision whether an image area should be binarized using local or global thresholds [6]. Shehzad Muhammad Hanif, Lionel Prevost implemented a texture based technique to detect text in grey level natural scene images. It is a wearable system to make possible navigation and to assist the blind and visually impaired persons in real world. It has three parts, a bank of stereovision, a processing unit for visual perception and a handheld tactile surface. The textual/symbolic information interpretation module to the vision system of the Intelligent Glasses will recognize the text from the captured scene and textual and/or symbolic information will be displayed on the handheld tactile [7]. Kumar J.A.V. , Visu A. , Raj M.S. , Prabhu M.T. implemented an automated text to audio converting pen . If a person would like to read/understand any portion of text that text is converted to an audio signal. This audio signal is transmitted to the person's ears through wireless technology such as ZigBee [8].

Oi-Mean Foong, Nurul Safwanah Bt Mohd Razali presented a signage recognition framework for Malaysian Visually Impaired People. Their proposed framework captures an image of a public signage and transforms it into a text file using Otsu's OCR method. The text file reads by a speech synthesizer that tells the visually impaired people what the image is. This framework does not need huge database of the signage but only the character database [9]. Krishnan K.G., Porkodi C.M., Kanimozhi K. successfully presented a method where a blind person can get information about the shape of an image through speech signal. The novelty of this work is to covert the image to sound using the methodology of edge detection [10]. Hangrong Pan, Chucai Yi, Yingli Tian designed a computer vision-based system to detect and recognize bus information from images captured by a camera at a bus stop. This system is able to notify the visually impaired people in speech the information of the coming bus, and detect the route number and other related information which is depicted in the form of text. For bus detection, histogram of the oriented gradient (HOG) descriptor is in use to extract the image based features of bus facade. Cascade SVM model is applied to train a bus classifier to recognize the existence of bus in sliding windows. In bus route no recognition they design a text detection algorithm on the basis of layout analysis and text learning and then recognize the text codes from detected text regions for audio announcement [11]. Michael R.T.F. , RajaKumar B., Swaminathan S. ,Ramkumar M. proposed a system that will be helpful for the visually challenged people. This model provides the opportunity to visually challenged person to operate the mobile devices without using the keypad [12]. Adil Farooq, Ahmad Khalil Khan, Gulistan Raja implemented human-computer interface system with a complete text recognition and speech processing capability. The method uses windows text to speech conversion and image recognition (OCR) technique to analyze and extract textual information from digital scanned images. Their research uses an open source engine Asprise OCR for text extraction and is expressed in audible system. The implementation was done in Microsoft visual studio using C sharp [13]. Chucai Yi, Yingli Tian and Aries Arditi proposed a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their everyday lives. To separate the object from cluttered backgrounds or other surrounding objects in the camera view, they first propose an efficient and effective motion based method to define a region of interest (ROI) in the

video by asking the user to shake the object. This method extracts moving object region by a mixture-of-Gaussians based background subtraction method. In the extracted ROI, text localization and recognition are conducted to obtain text information. Text characters in the localized text regions are then binarized and recognized by off-the-shelf optical character recognition software. The recognized text codes are output to blind users in speech [14].

2.2 Text detection and text recognition methods.

Text detection is to localize image regions containing text characters and string [1]. Methods of text detection, can broadly be classified as gradient features based, color segmentation based and texture features base, histogram of the oriented gradient (HOG) descriptor, layout analysis basis and text feature learning base. Text recognition is to convert pixel based text into readable code.

III. CONCLUSION

Today's world is moving towards digitalization. In recent years, digital cameras and camcoders are increasingly popular and they have shown potential as other imaging devices. The researchers working in document analysis and recognition have changed their orientation and instead of working with conventional scanner captured document images, they are concentrating on analysis of images taken from camera.

We have reviewed some text detection and text recognition methods. In review following interesting application were found such as newspaper reading out system, signage recognition and conversion to text, automatic electronic pen, camera based text reading system, travelling assistant system of bus detection and recognition.

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.