iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** III **Month of publication:** March 2019

DOI: http://doi.org/10.22214/ijraset.2019.3030

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

Portable Camera based Assistive Text and Product Label -Reading from Hand-Held Object for Visually Impaired Person

Aishwarya Ghuguskar[1], Sweta Fating[2], Pranjali Pongde[3], Karuna Gedam[4], Avinash Suryawanshi[5]
[1, 2, 3, 4] *Students, Department of ETC, KDK College of Engineering, Umrer, India*
[5]*Assistant Professor, Department of ETC, KDK College of Engineering, Umrer, India*

Abstract: ***Text detection and recognition plays a significant role in human life. It provides an extensive application in image analysis. Text information appears a real word object like documents, labels on grocery bottles, scene images and license plates etc. It is challenging for visually impaired user to find text region in image. This problem should not ignore because we cannot assume that the camera captured image contains region having text only. The main objective of this proposed system is to read text in camera captured image for assisting visually impaired users. While performing this task, capture image undergoes through different image processing techniques. Firstly, input image is binarized. Geometric and stroke width filtering helps to eliminate non-text region in a binarized image. Cropped text region in a complex scene is binarized and given to optical character recognition for character recognition. Visually impaired user users will get the output in a form of speech. Experimental result shows that proposed method offers better performance in text detection and recognition.***
Keywords: Binarize, stroke width, optical character recognition (OCR), connected component, geometric filtering, visually impaired, text recognition, detection.

I. INTRODUCTION

The national census of India has estimated about 21.9 million disabled people in country. Out of which 48.5% people are visually impaired. Many days to day activities of visually impaired people may become complicated and challenging. Not every printed or written document that a visually impaired person wants to read is available in Braille language. Disabilities of visually impaired people have made them limited access to documents, computers, Internets, etc. Examples of printed texts includes street signs, grocery labels, reports, classroom handouts, instructions on medicine bottles, bank statements, product labels etc. Now a days there is a need of basic education to visually impaired children and adults. It can be given by a teacher but it requires a large individual effort and different techniques are also necessary.

Automatic text detection and recognition from camera captured images receives an increasing attention because of potential applications in image retrieval, video annotation, target location and robotic navigation. Analysis of camera captured images are possible as large number of increased resolution digital cameras are available.

Text is present all over in real world.

It is not only present on images of report, documents etc. but also appear on various real-world objects. We can capture the image with the help of image capturing devices such as camcorders, digital cameras, cameras in personal digital assistants (PDA) and mobile phones.

It becomes possible to capture image with such devices as these devices are small, handy and light in weight. Along with document images, it captures text images present on real world objects such as buildings, road signs, license plates, shops, products etc. Text detection and recognition finds all areas in an image or video which is considered as a text by human. Text location mark the boundaries of the areas on which text is located and output will be in the form of characters present in that image.

It is allowable for real world images and video processing which is captured by a standard camera or a mobile phone camera and each detected area can be further processed by a computer for conversion of extracted content in an image into a readable text format. Automatic text extraction becomes extremely challenging due to non-uniformity in background, differences in text due to variations in style, colour, size, orientation, and alignment, as well as low image resolution, low contrast and complex background. Real world texts are generally written in different fonts and languages. Even arrangement of text does not follow rules of printed documents.

Fig.1.1. Some sample images containing text on objects such as Book, Products, Natural Scene Images etc.

II. LITERATURE REVIEW

Every year, the number of visually impaired persons is increasing due to eye diseases diabetes, traffic accidents and other causes. There are about 200,000 persons with acquired blindness in Japan. Therefore, computer applications that provide support to the visually impaired persons have become an important theme. We have already developed a pen-based character input system for blind persons using a PDA. On this system, people with acquired blindness remember the shape and writing order of Japanese characters and they can use this system as a notepad and as an E-mail terminal anytime, anywhere. This application essentially works as communication tool. However, such a device does not solve all of the problems encountered by a blind person willing to go outside unaccompanied. When a visually impaired person is walking around, it is important to get text information which is present in the scene. For example, a 'stop' sign at a crossing without acoustic signal has an important meaning. In general, way finding into a man-made environment is helped considerably by the ability to read signs. The system capable of localising and reading aloud text embedded in natural scene images can be very helpful for blind and visually impaired person-providing information useful in everyday life, it increases their confidence and autonomy. Even though the currently available optical character recognition (OCR) programs are fast and accurate most of them fail to recognised text embedded in natural scene image. The goal of the algorithm describe in this paper is to localised text-like image regions and pre-process them in the way that will make OCR work more reliably, the approach described in the paper is based on colour image segmentation and segment shape analysis. Preliminary test has shown that the proposed algorithm offers satisfactory detection rate and is pretty robust to typical text distortions, such as slant, tilt and bends.

III. METHODOLOGY AND IMPLEMENTATION

This paper presents of assistive text reading. As illustrated in fig the system framework consists of three main function components:

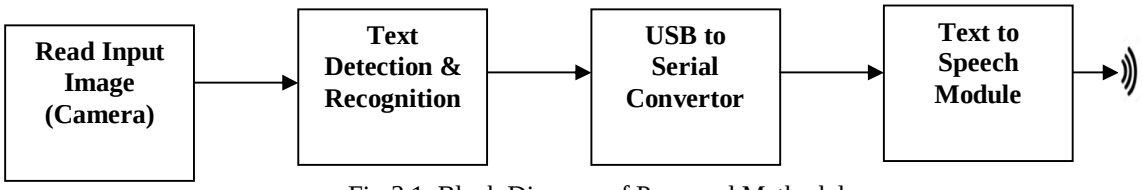


Fig 3.1. Block Diagram of Proposed Methodology

A. Image Capture

The image capture component captures image with the help of web camera attached to the system. Here, Image is captured with the help of web cam (Logitech C207 HD). The image from the webcam is taken in RGB format.

B. Data Processing

The data processing component is used for processing captured image for extracting text regions. Fast and robust image processing techniques are used for improving image quality and to detect text region in an image. Text localization algorithm is used to obtain text regions in image. Text recognition is performed to transform text information into understandable codes. We can use laptop or desktop computer as a processing device.

C. Audio Output

The audio output component helps to inform the visually impaired user of recognized text codes in the form of audio or speech. Recognized texts are passed serially to text to speech conversion module through serial convertor. By using text to speech module, text information is converted into speech or audio. Visually impaired user can hear the audio output with help of headphones or speaker attached to the system.

IV. MECHANISM OF DATA PROCESSING

The data processing involves the following steps:

A. *Binarisation*

In binarization, conversion of a gray scale image (having pixel value in between 0 to 255) into binary image (having pixel value 0 and 1) is carried out by selecting a threshold value that separates the foreground and background. Each pixel is compared with the threshold value. If it is greater than the threshold then value is set to be 1 else the value is 0. Binarization can be performed in two ways i.e. global or local.

In global method, one intensity value is applied to the whole image. Global method is useful in the images where contrast distribution of foreground and background is uniform.

The histogram analysis of image (document) generally shows two peaks i.e. a higher peak represents white background and a small peak represents the foreground. In local method or adaptive thresholding methods different intensity value is applied to different parts of the image. Local method is mainly used in degraded images (documents) where there exists illumination problems or background noise. Among various binarization techniques Otsu, Niblack, Savula are commonly used.

B. *Geometric Filtering*

In binary image, CC i.e. connected components in foreground region are letter candidates. Non-text area in an image is removed by applying some geometric rules on connected component.

In geometric filtering, the objects which are very small or large are rejected. Generally, aspect ratio of text characters is closely equal to one therefore CCs which does not follow this property can be neglected. Regions having aspect ratio greater than 3 are removed from the image. Aspect Ratio is ratio of width and height or it is ratio of longer side to shorter side of a rectangle.

Aspect Ratio = Longer side / Shorter side

C. *Stroke Width Filtering*

Stroke width extraction is used for removing non-text regions which remains after filtering the geometric properties of connected components. Text characters has stroke with constant or variable orientation.

Stroke width is determined which is based on the distance transform. The stroke width information is given at each pixel of the original CC with any shape of stroke. In our algorithm, we apply the Euclidean distance transform for labelling each pixel of foreground with the distance to its nearby background pixel.

D. *Text Line Formation*

Large number of non-text regions are eliminated from the previous steps. Mostly text characters remain in image after removal of connected components. Connected regions exists in binary image can be detected using connected component labelling. Characters which remained in previous steps are grouped into single connected component. It can be done using morphological operations. Only those regions remain in binarized image which have area greater than the threshold value. Here we have taken threshold value as 7500 pixels. Letters are grouped together for identifying horizontal text lines. Here Otsu's method [18] is used for cropped text region. It performs the binarization of text regions. This binarized text is applied to optical character recognition (OCR) which recognize the texts present in binarized image and it displays the output text.

E. *Optical Character Recognition (OCR)*

Optical character recognition (OCR) is the electronic or mechanical conversion of images into machine encoded text. It helps to identify numeric digit or alphabetic character in an image. Once the character is recognized, it is transformed into ASCII code. OCR helps to read text of various font sizes. But it cannot properly recognize handwritten texts. OCR identifies text present in an image and display output text.

F. *Text to Speech Conversion (Hardware)*

We used text to speech module for conversion of text to audio output. Text to speech module is mainly based on basics of phonics of English. Recognized texts are passed serially to text to speech module with the help of Serial to USB convertor.

V. RESULT

To evaluate the effectiveness or efficiency of our proposed system, we apply this algorithm on various captured images. We have collected images of book covers, product labels, natural scene images, documents etc. Logitech C207 HD web camera with autofocus is connected to laptop with the help of USB connection. Figure 7.1 shows snapshot of experimental system.

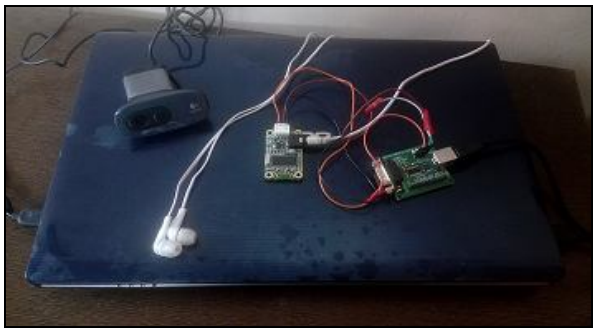


Fig.5.1. Snapshot of Experimental System

A. Graphical User Interface (GUI)

GUI is the main part of software application for interaction between human and computer. User can interact with our system with the help of graphical user interface. *Once the user gets GUI, he can control it with the help of some keys on the keyboard. Here, 'Capture Image' button can be control by key 'F'. 'Text Region' button can be controlled by key 'G'. 'Recognize' button can be controlled by key 'H'. 'TTS' button can be controlled by key 'J'. We can also make GUI more effective by giving voice instructions on different buttons.* GUI of this project is as shown below.

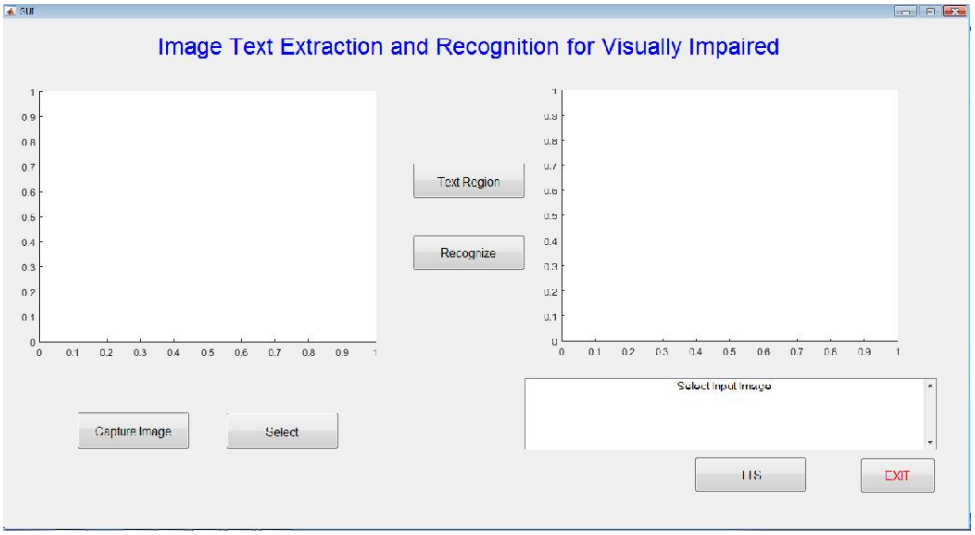
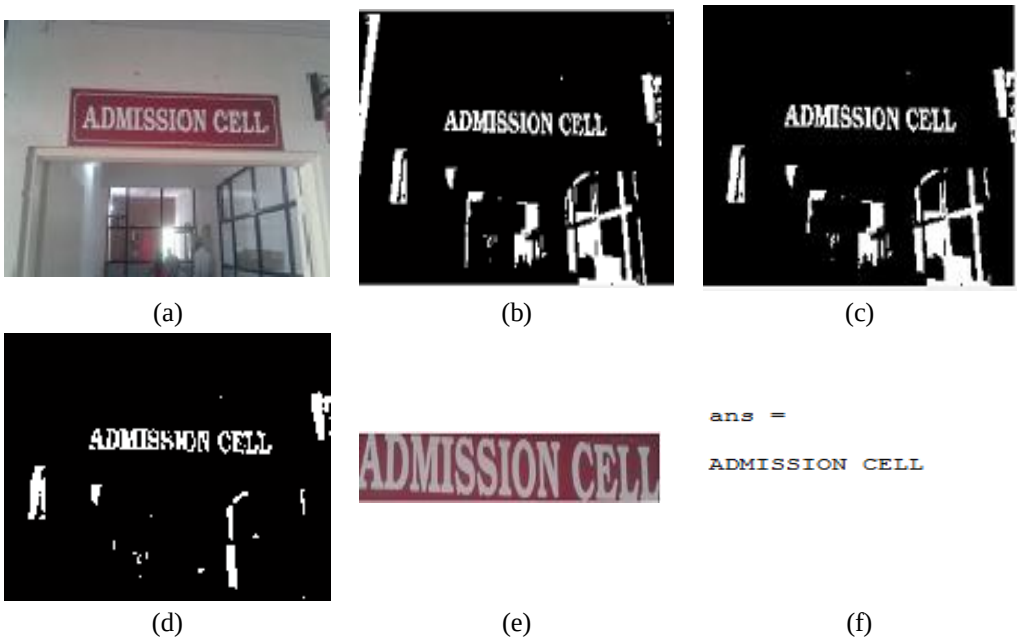


Fig.5.2. Graphical User Interface

B. Text Detection & Recognition Process

The captured image which has some text regions along with non - text regions also. Captured image is processed first for removing non-text regions and getting text regions only. Fig. (b) shows binarized image after pre-processing. Filtering is performed to eliminate some connected components (non-text regions) using region properties. Specific threshold values are used for various region properties. Threshold value of Eccentricity is 0. 995.Threshold value of Euler number is -4 and threshold value of solidity is 0.4. It helps to remove the regions which does not fit into the properties of text. Fig. (c) shows image after geometric filtering. Characters which shows too much variation in stroke are removed from image. Stroke width filtering helps to eliminate large number of non-text region from image. Fig (d) shows image after stroke width filtering. Morphological opening and closing helps to get bounding box of text region. Fig. (e) shows text region which is cropped from original image. Otsu's binarization method is applied on extracted text region and this binarized image **is** given to optical character recognition i.e. OCR for recognizing characters in binarized image. Fig. (f) shows result of this proposed system provided in MATLAB.



(a) (b) (c)

(d) (e) (f)

Overview of text detection. (a) Original Image. (b) Image after binarization. (c) Image after region filtering. (d) Image after stroke width filtering. (e) Cropped text region. (f) OCR result.

This algorithm is tested on various natural images to detect text information present in those images. These images include products, book covers and also various scene images. Fig. 5.2 shows some results showing text information present in images we have taken.

VI. CONCLUSIONS

We have proposed a system to read text on book covers, documents, product labels, scene images etc. Non-text regions from cluttered or complex background are removed by applying geometric rules and stroke width filtering. Text recognition is done with OCR i.e. optical character recognition. User will get the output in speech or voice with the help of text to speech converter module. Performance of proposed method is degraded when deal with distorted or blur and low resolution or low contrast images.

VII. FUTURE ENHACEMENT

We have described a system to read printed text on various images for assisting visually impaired people. This method can effectively extract the text regions from images with complex background. Our future work will extend our algorithm to handle curved or non-horizontal text string. In future we will enhance this work with higher accuracy and large number of input samples. Also, we will try to increase portability of system.

REFERENCES

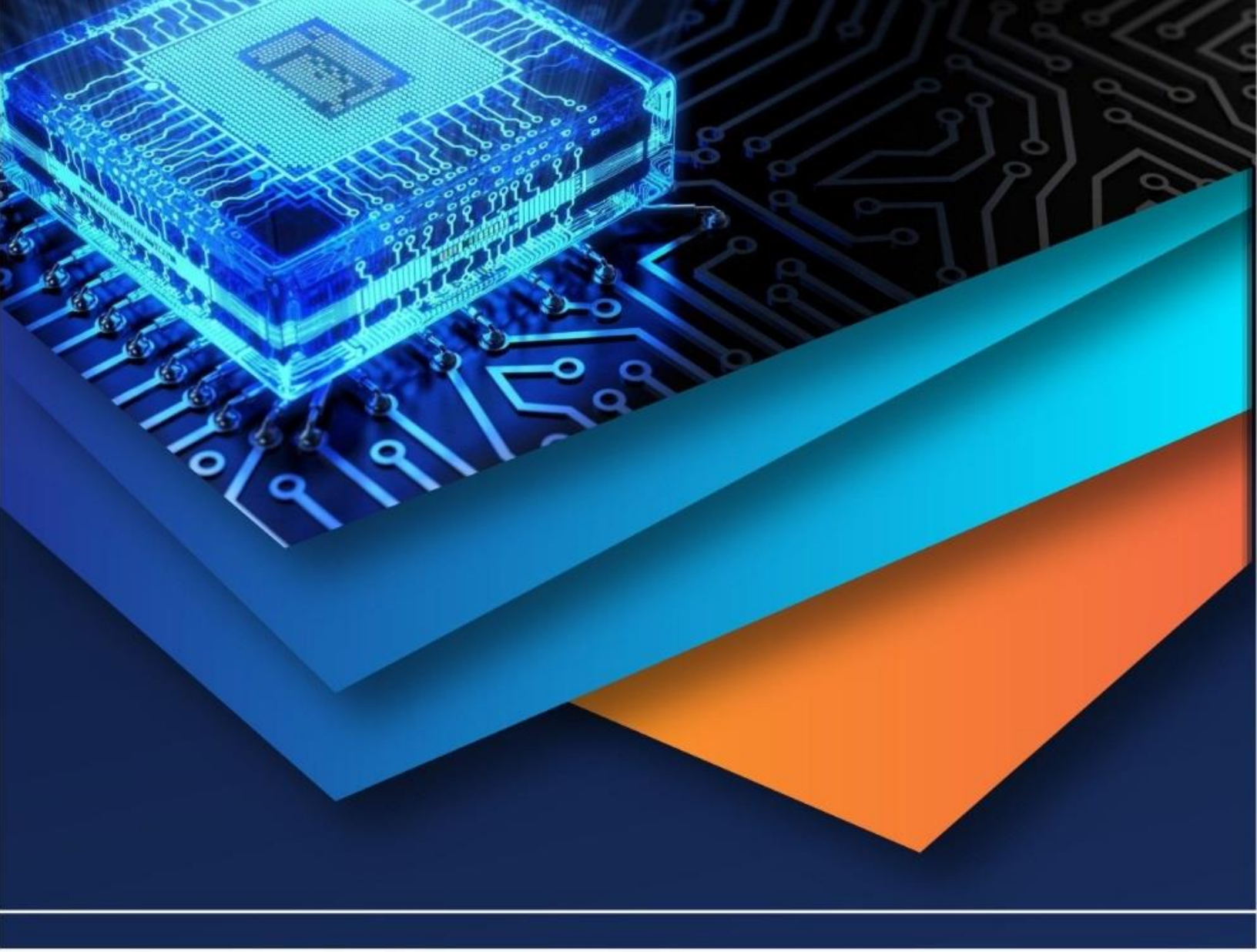
[1] Nobuo Ezakai, Marius Bulacu, Lambert Schomeker “Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons” Proceedings of the 17th International Conference on Pattern Recognition (ICPR’04)1051-4651/04 $ 20.00 IEEE

[2] Marcin Pazio, Ryszard Kowalik “Text Detection System for The Blind” 15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007, copyright by EURASIP

[3] National Census of India, [Online]. Available: http://censusindia.gov.in/Census_And_ You /disabled_population.aspx

[4] ScanTalker, Bar code scanning application to help Blind Identify over one million products(2006).[Online].Available:http://www.freedomscientific.com/fs_news/PressRoom/en/ 2006/ScanTalker2-Announcement_3–30–2006.asp

[5] KReader Mobile User Guide, knfb Reading Technology Inc. (2008). [Online]. Available: http://www.knfbReading.com



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY