A Portable Camera-Based Assistive Text Reader for Blind Persons

Pooja Belunkhi[1], Dr. Ravindra Patil[2]

[1]Student, Department of Electronics, DKTE Society's, Textile and Engineering Institute, Ichalkaranji Kolhapur Maharashtra India-416115
[2]Professor, Department of Electronics , DKTE Society's, Textile and Engineering Institute, Ichalkaranji Kolhapur Maharashtra India- 416115

Abstract— *A portable camera-based assistive text reader structure is projected to assist a blind persons to read text labels and titles from objects/products in their day-to-day life. The cluttered backgrounds or other surrounding of objects in the camera view are isolated by using a region of interest (ROI). In extracted region of interest, text localization and text recognition will be done to acquire text information. Text characters in the localized text regions are then converted into binary format and recognized by trained optical character recognition software. The recognized text codes are output to blind users in speech. The proposed framework is implemented on Raspberry pi board.*

Keywords— assistive device; Blindness; Open CV; optical character recognition (OCR); Raspberry pi;

1. Introduction

Globally, the impacts of vision loss extend far beyond the frequency of blindness and vision impairment. For the 39 million people who are blind and the 246 million who have a vision impairment, vision loss often represents a lifetime of inequality [W1]. Even in a developed country like the United States, the 2008 National Health Interview Survey (NHIS) reported that an estimated 25.2 million adult Americans (over 8%) are blind or visually impaired [W2]. This number is increasing rapidly as the baby boomer generation ages. Reading is vital in daily life. Printed text is available everywhere in bank statements, form of receipts, product labels, restaurant menus, on medicine bottles, etc. The capability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and social self-sufficiency. A camera-based text reader frame assist blind persons to read text labels in their day-to-day life.

Recent researches in computer vision and portable computers make it possible to help blinds by making camera-based products that combine computer vision technology along with other commercial products such OCR systems. Also while screen readers, optical aids, and video magnifiers can help out blind users and those with low visualization can access documents, there are few devices that can provide good access to common hand-held objects such as product labels, and objects printed with text. K-Reader Mobile Applications which are intended specifically for blind users which performs good when detecting text information from receipts, fliers, and many other documents. But these systems/ device fail to give an economic solution of the problem and are accessible on specific platforms. No smart phones have planned for blind person until now. Thus convenience of the Mobile application is a different question. Also, these systems cannot read text from commercial products with complex background, non-flat surfaces (text printed on cylindrical surface for e.g. medicinal bottles) and other packaged products. Although a number of assistive reader systems have been designed specifically for blinds, the proposed system presents a portable and economic solution.

Although a number of reading assistants have been designed specifically for the visually impaired, to our knowledge, no existing reading assistant can read text from the kinds of challenging patterns and backgrounds found on many everyday commercial products. As text information can appear in multiple scales, fonts, colors, and orientations. To help out the blind people to read text from hand held objects, system has conceived of a camera-based text reading frame to track object of interest within the camera view and extract print text information from the object. Our proposed algorithm can effectively handle complex background and multiple patterns, and extract text information from hand-held objects.

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labeling. For example, portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products through speech and Braille. But main drawback is that it is very difficult for blind users to get the position of the bar code and to correctly point the bar code reader at the bar code. Some reading-assistive systems such as pen scanners might be employed in these and similar situations. Such systems integrate OCR software to offer the function of scanning and recognition of text and some have integrated voice output. However, these systems are generally designed for and perform best with document images with simple backgrounds, standard fonts, a small range of font sizes, and well-organized characters rather than commercial product boxes with multiple decorative patterns. In solving the task at hand, to extract text information from backgrounds with multiple and variable text patterns, here propose a text localization algorithm that combines rule based layout analysis and learning-based text classifier training. These, in turn, generate representative and discriminative text features to distinguish text characters from background outliers.

2. Framework and Algorithm overview

2.1 System Block Diagram

The block diagram of the proposed framework is one of the techniques which work out as a portable assistive text reader. The proposed framework is designed on single board computer –Raspberry pi, in order to provide an economic and potable solution. . The image processing module is done. The text image to the text file is converted using a text localization algorithm. The implemented block is as shown in Fig.1

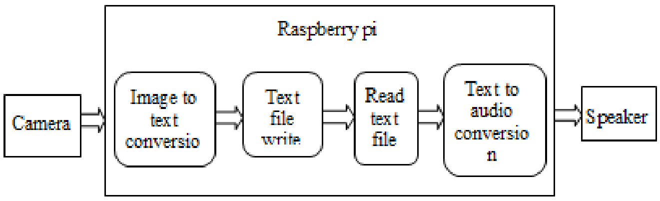

Fig.1. Block diagram of image to Speech conversion

The system consists of Raspberry pi, Camera, SD card and personal computer. Those all components are connected by USB adaptors. Raspberry pi is the key element in processing module. The raspberry pi will also do the image processing on the image like image segmentation, noise removal, boundary detection etc.

2.2 Methodology of Implementation.

Text-to-speech converter module contains of two main modules, the image processing module and the voice processing module (Fig.2). The Image Processing System captures image with the help of camera and converts the image into text. In Fig.2 first block is image processing module which converts .jpg to .txt form and second block is voice processing module which converts .txt to .wav/.speech form.

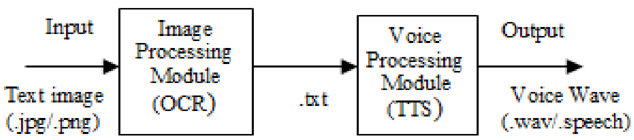

Fig.2. Stages of Text-To-Speech system

2.3 Text recognization Algorithm

The implementation of the image to text is the first module which converts text image file(.jpg) into the text file(.txt). The methodology for image to text conversion is as shown in Fig.3.

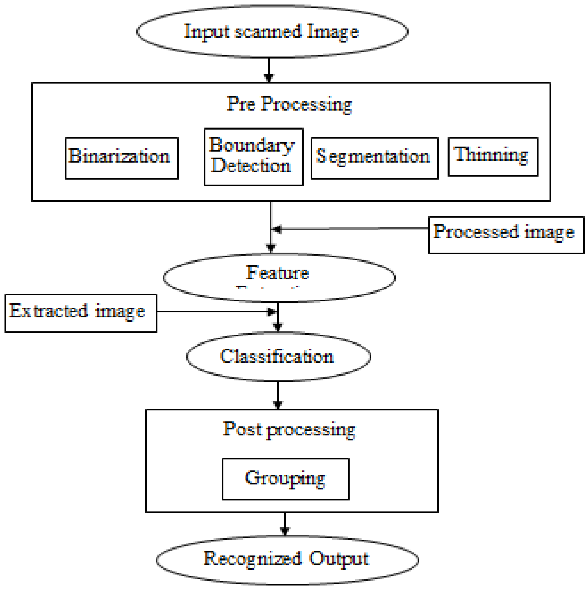

Fig.3.flow chart of text recognition system

A typical text recognition algorithm consists of the following basic components:

1) Input scanned Image
2) Pre-processing
3) Feature Extraction
4) Classification
5) Post- processing

1) Input scanned Image

The two required command line arguments are:

- --image: The path to the image to be OCR'd.
- --reference: The path to the reference OCR image.

This image contains the digits 0-9 in the OCR-A font. First of all load reference OCR image then convert it to grayscale and do thresholding + inverting on it. In each of this operation I have restored or overwrite ref. The reference image is as shown in fig 1.1.

Fig.1.1: reference image

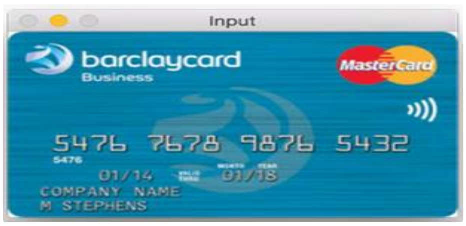

Fig.1.2: Input image

Fig.1.3: gray scaled image

2) Finding contours for the OCR image

To find contours in the reference OCR image (i.e. the outlines/boarder of the digits) and sort them from left to right direction, and initialize a dictionary to map digit name to the ROI. The contours from left-to-right is sorted as well as initialize a dictionary, digits, which maps the digit name to the region of interest At this point, looping has been done through the contours, extract, and associate ROIs with their corresponding digits.

3) Extracting region of interest (ROI)

The region of interest is extracted from ref (the reference image) by using the bounding rectangle parameters. This ROI contains the digit. Then resize each ROI to a fixed size of 57×88 pixels. To ensure *every* digit is resized to a fixed size in order to apply template matching for digit recognition.

4) Extraction small elements and details

The input image is gray scaled. Next step is to perform a morphological operation on the input image. With the help of rectKernel and converted gray input image perform a Top-hat morphological operation and store the output result as Top-hat. The Top-hat operation gives light regions against a dark background the resulting image is as shown in Fig 4.1

Fig4.1. Top hat resultant image

The next step is to isolate the digits and compute a Scharr gradient of the Top-hat image in the *x*-direction. After computing the absolute value of each element in the gradX array, then scale the values into the range [0-255]. To do this computes the minVal and maxVal of gradX followed by scaling equation (i.e., min/max normalization). The last step is to convert gradX to a uint8 which has a range of [0-255] .The result is shown in Fig5.1.

Fig 5.1 sobel operated image

5) Closing the gaps:

Closing operation is done to close the gaps between the digits. The rectKernel is used. An Otsu thresholding is performed and binary threshold of the gradX image followed by another closing operation. The result of these steps is shown in Fig.6.1.

Fig 6.1. Result of Otsu and Binary Threshold

Reference image) by using the bounding rectangle parameters. This ROI contains the digit. Then resize each ROI to a fixed size of 57×88 pixels. To ensure *every* digit is resized to a fixed size in order to apply template matching for digit recognition.

6) Extraction small elements and details

The input image is gray scaled. Next step is to perform a morphological operation on the input image. With the help of rectKernel and converted gray input image perform a Top-hat morphological operation and store the output result as Top-hat. The Top-hat operation gives light regions against a dark background the resulting image is as shown in Fig 4.1

Fig4.1. Top hat resultant image

7) Edge detection:

The next step is to isolate the digits and compute a Scharr gradient of the Top-hat image in the *x*-direction.

After computing the absolute value of each element in the gradX array, then scale the values into the range [0-255]. To do this computes the minVal and maxVal of gradX followed by scaling equation (i.e., min/max normalization). The last step is to convert gradX to a uint8 which has a range of [0-255] .The result is shown in Fig5.1.

Fig 5.1 sobel operated image

8) Closing the gaps:

Closing operation is done to close the gaps between the digits. The rectKernel is used. An Otsu thresholding is performed and binary threshold of the gradX image followed by another closing operation. The result of these steps is shown in Fig.6.1.

Fig 6.1. Result of Otsu and Binary Threshold

9) Grouping and display of result

The contours are found and stored them in a list. Then initialize a list to hold the digit group locations. Now let's loop through the contours while filtering based on the aspect ratio of each, allow to pruning the digit group locations from other, irrelevant areas of the credit card.

Fig 7.1. Grouping of the digits

3. Results:

Fig.1. Input image

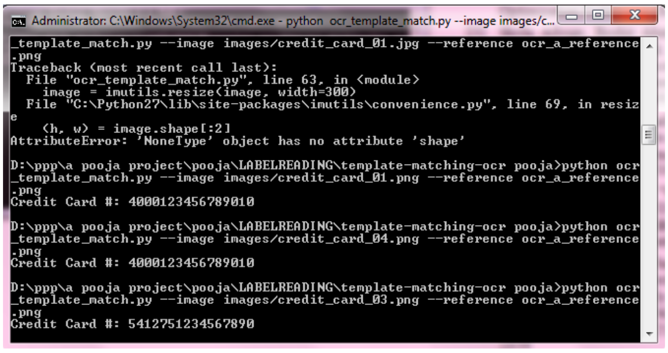

Fig.2. Command Prompt Window

Fig.3. Resultant images with detected digits (in red)

Fig.4. Walmart Card Fig.5. OpenSky Card

Fig 6. Platinum Card Fig 7.US Bank Card

Fig 8. Cerdit Card

Fig 9. Bed Bath & Beyond Card

4. Conclusion and Future work

In this paper, I have described the prototype system to read digits on the card and it will help the blind persons to read the contents. For solving the general problem of the blind user, I have designed the system using text reorganization algorithm. This method can easily distinguish the digits from the card and extract the interested region. The text localization algorithm is used to recognize the digits from the cards. This proposed framework can handle to the noisy, blur, reflection and complex background images. I have done digit recognition for 20 cards. The achieved results are as shown in results there is around 99.99% accuracy for matching but in Fig.7. US Bank card it gives near about 75% accuracy due to mismatching.

The future work will be extending for the recognition of the string character and paragraphs of text books. The future framework will recognize more than three blocks and also improvement in algorithm to handle a horizontal text string.

5. Reference:

[1] 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, JUNE 2014. pp: 808-817

[2] Samruddhi Deshpande, Ms. RevatiShriram "Real Time Text Detection and Recognition on Hand Held Objects to Assist Blind People", 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT) ©2016 IEEE. pp: 1020-1024

[3] Shyla Afroge, Boshir Ahmed, Firoz Mahmud "Optical Character Recognition using Back Propagation Neural Network "2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE) 8-10 December 2016, Rajshahi-6204, Bangladesh ©2016 IEEE. pp: 1-4

[4] RobiahHamzah, Mohammad Izzat Mohamad Fadzil "Voice4Blind: The Talking Braille Keyboard to Assist the Visual Impaired Users in Text Messaging", 2016 4th International Conference on User Science and Engineering (i-USEr) ©2016 IEEE. pp: 256-270

[5] Sanjana.B, J.RejinaParvin "Voice Assisted Text Reading System for Visually Impaired Persons Using TTS Method" IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 6, Issue 3, Ver. III (May. - Jun. 2016), www.iosrjournals.org

[6] KwangIn Kim, Keechul Jung, and JinHyung Kim "Texture-Based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 12, DECEMBER 2003. pp:1631-1639

[7] R. Lienhart and A. Wernicke, "Localizing and segmenting text in images and videos," Circuits and Systems for Video Technology, vol. 12, no. 4, 2002, pp. 256 –268.