Tech – NETRA (New Eyes To Read Artifact)

Priyanka Bacche, Apurva Bakshi, Karishma Ghiya, Priyanka Gujar

Abstract — Nowadays printed text appears everywhere like product names, restaurant menus, instructions on bottles, signed boards etc. Thus blind people need some assistance to read this text. In this paper, we propose a camera based portable device which helps blind people to read printed text from hand held objects. User has to take video by shaking product in front of camera of device. To recognize the product from complex background we perform background subtraction of successive frames and define region of interest (ROI) i.e. product. Then we perform text area localization algorithm to localize the text region on region of interest. Then, the optical character recognition (OCR) is used to convert the text from text regions and then converted to voice output.

Index Terms— blind people, optical character recognition (OCR), region of interest (ROI), text area localization

I. INTRODUCTION

According to a survey there are 45 million blind people exists today [1]. Recent developments in computer technology make it feasible to assist these individuals by developing camera-based products. Blind people need some portable assistance to read this printed text. To achieve this few systems had been designed. Portable barcode readers help blind people to identify barcode information of different products via speech and Braille [2]. But major drawback of such systems is to find exact position of barcode on that product.

Another portable reading assistant, KReader Mobile, when installed on cell phone allows users to read documents like mails and receipts [3]. But this software requires the document to be placed on a clear and flat surface and background should mostly be filled with text. Even though it can help read black print on white background, it cannot read colored text or text on complex backgrounds. Also most voice pens for blind assistance require manual localization of text lines in documents and objects in hand. However, in camera-based scene images, manual text localization is impractical, especially for blind people. Therefore, algorithms of automatic text localization are required for better results [4].

Manuscript received Feb, 2014.

***Priyanka Bacche** , PVGCOET, Pune University ,Pune, Maharashtra, India,*

***Apurva Bakshi** , PVGCOET, Pune University ,Pune, Maharashtra, India,*

***Karishma Ghiya** , PVGCOET, Pune University ,Pune, Maharashtra, India,*

***Priyanka Gujar** , PVGCOET, Pune University ,Pune, Maharashtra, India,*

This work is mainly to help blind people to shop independently. The portable device Tech-NETRA assists blind people to read the printed text on products.

II. PROPOSED SYSTEM

The proposed system consists of a portable device which has camera and processing capability. The input to the device is video captured by the user and the device gives audio output as name of the product. The main steps in this process are explained by the Fig 1.

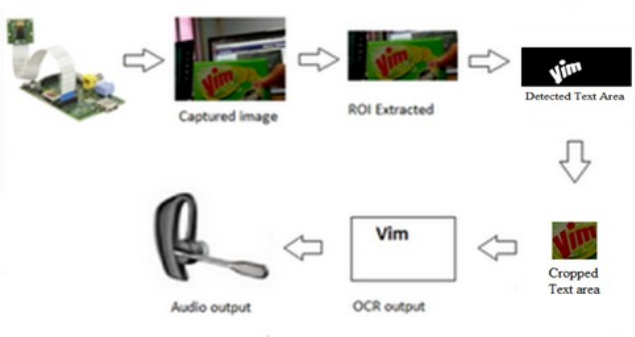


Fig -1: Proposed System

User has to place the product in front of the portable device and shake it. Then the camera mounted on the device will capture video of the product and process it. Finally the recognized text is converted to the voice output and provided to the user through earphone.

The portable device used here is Raspberry Pi model B. It is having 750 MHz processor and 3mm jack for audio output. The camera module of Raspberry pi is used for capturing HD video. The Debian operating system is installed on device and whole system is implemented using C++ language.

III. METHODOLOGY

A. Video Processing and Background Subtraction

The main task performed by device is video processing to extract ROI from video. The video limit is set to around 5s automatically. User has to shake product in front of the camera for first few seconds and then he should keep product stable for few seconds at the end of the video.

Background Subtraction in video is done by method of mean and variance by comparing each new frame with previous one and comparing number of pixels which are changed [5]. The learning rate in this process is set to -1. After some time number of background pixels which are stable are made black and pixels which are moving and which are foreground pixels are marked as white.

ISSN: 2278 – 7798

After processing some frames in video, a frame is taken for analyzing product area i.e. ROI. The foreground pixels which are marked white are collected in vector of points and from these connected components the bounding box is constructed. The points of bounding box are used to crop the ROI from last frame of the video which is stable frame so that in ROI product will be visible clearly.

Fig -2: A frame of the video

Fig -3: The cropped ROI from video

B. Text area localization and text extraction

Next task performed by portable device is localizing the text area and text extraction . There are two methods available for text extraction

1. Region based method
2. Textual based method

Region based method focuses on high contrast between text and background whereas textual based method focuses on some textual properties which distinguish text from its background.

For our system we use region based method. The following steps are performed for extracting text area.

1. The color RGB image is first converted into grey scale image. Then that image is converted into binary using OSTU's thresholding.

2. The obtained binary image is given to the canny edge detector. The minimum and maximum thresholds are set to 60 and 180 respectively.

3. The edge image is dilated using square structuring element of size 2X2. The dilation is performed for obtaining contours.

FIg -4: Edge detected image

4. The contours are labeled from image. The extent value for contour is calculated using below formula:

$$\text{Extent value} = \frac{Height * Width}{Area}$$

Where,

Height and Width are height and width of contour and Area is number of pixels in contour area. The contours for which extent value is less than 2.13 and height is greater than 9 and also area > 370 are recognized as text area. The new blank image is taken and only these contours are drawn and filled on that image [6].

Fig- 5:Filled Image

C. Text recognition and Audio output

The localized text area is cropped from original image (Fig 3) and given to command line OCR. The OCR returns text file which is then given to Text-to-Speech converter. That will produce audio output.

IV. Result

The detailed results are shown in Table I.

V. Conclusion

In this paper, we proposed a Portable system for blind people to read product labels. Above system is tested for various cases and it gives appropriate results for horizontally aligned text labels on complex colored backgrounds .It is also

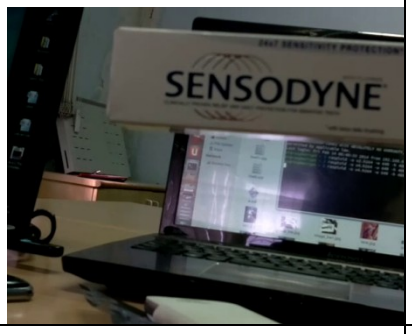
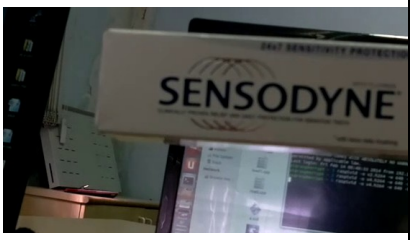
Processed frame of video	Detected object	Detected product name
		
		
		

Table I: Results

able to recognize product from complex backgrounds. Results may not be accurate if lighting effects on text is not uniform.

ACKNOWLEDGMENT

We would like to seize the opportunity to express our sincere gratitude towards Mr. Sagar Deshmukh, Senior Manager of 'Techlead Software Engineering Pvt. Ltd.', for providing us with his invaluable support and sponsoring us for this noble project.

We also want to thank our HOD Prof. Mr. G. V. Garje , our project guide, Prof. Mrs. M. S. Pokale and other computer department faculties for their invaluable co-operation and guidance.

REFERENCES

[1] World Health Organization. (2009). 10 facts about blindness and visual impairment[Online].Available:www.who.int/features/factfiles/blindness/blindness_facts/en/index.html

[2] Scan Talker, 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

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

[4] Chucai Yi, Student Member, IEEE, Yingli Tian, Senior Member, IEEE, and Aries Arditi, "Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for Blind Persons", IEEE/ASME TRANSACTIONS ON MECHATRONICS, 2013

[5] Zoran Zivkovic, "Improved Adaptive Gaussian Mixture Model for Background Subtraction", In Proc. ICPR, 2004

ISSN: 2278 – 7798

[6] Satish R. Damade, Ranjana S. Zinjore, "Morphological Approach for Text Extraction from Document Image", Proceedings of SARC-IRAJ International Conference, 14th July 2013

Priyanka U. Bacche was born in Pune, India on 14.11.1992 and currently pursuing the under graduate course, B.E. (Computer Engineering) in Pune Vidyarthi Griha's College Of Engineering, Pune, India

Apurva A. Bakshi was born in Pune, India on 13.10.1992 and currently pursuing the under graduate course, B.E. (Computer Engineering) in Pune Vidyarthi Griha's College Of Engineering, Pune, India

Karishma A. Ghiya was born in Pune, India on 17.07.1993 and currently pursuing the under graduate course, B.E. (Computer Engineering) in Pune Vidyarthi Griha's College Of Engineering, Pune, India

Priyanka R. Gujar was born in Pune, India on 23.06.1993 and currently pursuing the under graduate course, B.E. (Computer Engineering) in Pune Vidyarthi Griha's College Of Engineering, Pune, India