ISSN:0975-9646

S. Babu Renga Rajan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (6) , 2015, 5173-5176

Product Label Reading From Handheld Objects Using Portable Camera For Visually Challenged Persons

Prof. S. Babu Renga Rajan[1], Mrs.B. Shanmuga Sundari[2], Mrs. A. Haseena Beevi[3]

[1-] Professor & Head, Dept of CSE, PET Engineering College, Vallioor
[2, 3-]Assistant Professor, Dept of IT, PET Engineering College, Vallioor

Abstract- This system propose a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives. To isolate the object from cluttered backgrounds or other surrounding objects in the camera view, we 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 acquire text information. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Adaboost model. 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. We explore user interface issues and assess robustness of the algorithm in extracting and reading text from different objects with complex backgrounds.

Index Terms — **ROI (Region of Interest), Hand-held objects, Optical Character Recognition (OCR), text reading, text region localization, Background Subtraction (BGS)**

I. INTRODUCTION

In worldwide the visually impaired peoples are 314 million, in that 45 million are visual impairment which was released by "World Health Organization" in 10 facts regarding blindness. The estimation of The National Health Interview Survey 25.2 million adult Americans (over 8%) are blind or visually impaired [1,2]. This number is increasing rapidly as the baby boomer generation ages. Reading is obviously essential in today's society. Printed text is everywhere in the form of reports, receipts, banking statements, restaurant men us, classroom handouts, product packages, instructions on medicine bottles, etc. And while optical aids, video magnifiers, and screen readers can help blind users and those with low vision to access documents, there are few devices that can provide good access to common hand-held objects such as product packages, and objects printed with text such as prescription medication bottles. The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and foster economic and social self-sufficiency so here we are going to propose a system that it useful to blind people.

A number of portable reading assistants have been designed specifically for the visually impaired. *KReader Mobile* runs on a cell phone and allows the user to read mail, receipts, fliers, and many other documents [3]. However, the document to be read must be nearly flat, placed on a clear, dark surface (i.e., a non cluttered background), and contain mostly text.

Fig. 1. Examples of printed text from hand-held objects with multiple colors, complex backgrounds, or nonflat surfaces.

(a) (b) (c) (d)

Fig.2. Example of text localization and recognition from camera captured images. (a) camera captured images. (b) Localized text regions (marked in blue). (c) Text regions cropped from image. (d) Text codes recognized by OCR.

As shown in Fig. 1, such text information can appear in multiple scales, fonts, colors, and orientations. To assist blind persons to read text from these kinds of hand-held objects, we have conceived of a camera-based assistive text reading framework to track the 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 both hand-held objects and nearby signage, as shown in Fig. 2.

II. ALGORITHM OVERVIEW

This paper presents a model system of assistive text reading. As illustrated in Fig. 3, the system framework consists of three functional components[1]:

1. scene capture
2. data processing and
3. audio output.

The scene capture component collects scenes containing objects of interest in the form of images or video.

In our prototype, it corresponds to a camera attached to a pair of sunglasses. The data processing component is used for deploying our proposed algorithms, including 1) object-of-interest detection to selectively extract the image of the object held by the blind user from the cluttered background or other neutral objects in the camera view; and 2) text localization to obtain image regions containing text, and text recognition to transform image-based text information into readable codes. We use a laptop as the processing device in our current prototype system. The audio output component is to inform the blind user of recognized text codes. A Bluetooth earpiece with microphone is employed for speech output.

Fig. 3. shows three functional components for Scene capture, data processing and audio output

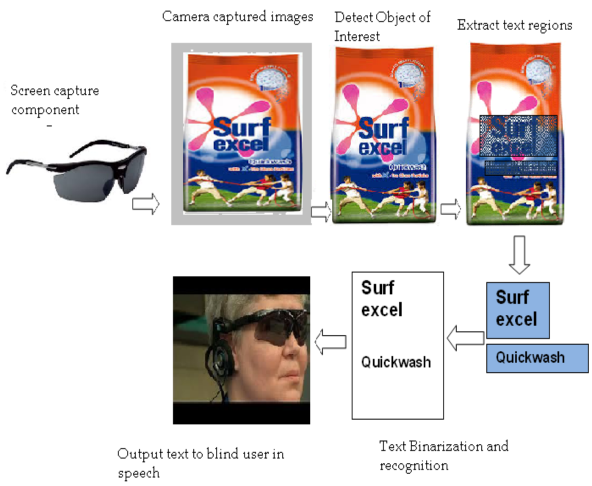

Fig. 4. Flowchart of the proposed framework to read text from hand-held objects for blind users.

A frame sequence V is captured by a camera worn by blind users, containing their hand-held objects and cluttered background. To extract text information from the objects, motion based object detection is first applied to determine the user's object of interest S by shaking the object while recording video

$$S = \frac{1}{|V|}\sum_i R\,(V_i, B)$$

where Vi denotes the ith frame in the captured sequence, $|V|$ denotes the number of frames, B denotes the estimated background from motion-based object detection, and R represents the calculated foreground object at each frame.

Next, our novel proposed text localization algorithm is applied to the object of interest to extract text regions. At first, candidate text regions are generated by layout analysis of color uniformity and horizontal alignment

$$X^c = argmax_{s \in S} L(s)$$

where $L(\cdot)$ denotes the suitability responses of text layout and XC denotes the candidate text regions from object of interest S. Then, a text classifier is generated from a Cascade-Adaboost learning model, by using stroke orientations and edge distributions of text characters as features.

$$X = H[X^c] = H[argmax_{s \in S} L(s)]$$

where H denotes the Cascade-Adaboost classifier and X denotes the localized text regions.

After text region localization, off-the-shelf OCR is employed to perform text recognition in the localized text regions. The recognized words are transformed into speech for blind users.

III. OBJECT REGION DETECTION

To ensure that the hand-held object appears in the camera view, we employ a camera with a reasonably wide angle in our prototype system (since the blind user may not aim accurately).

To extract the hand-held object of interest from other objects in the camera view, we ask users to shake the hand-held objects containing the text they wish to identify and then employ a motion-based method to localize the objects from cluttered background. Background subtraction (BGS) is a conventional and effective approach to detect moving objects for video surveillance systems with stationary cameras. To detect moving objects in a dynamic scene, many adaptive BGS techniques have been developed.

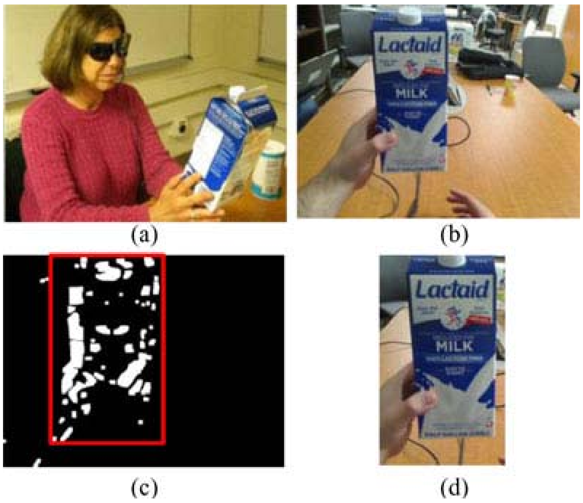

Fig. 5. Localizing the image region of the hand-held object of interest. (a) Capturing images by a camera mounted on a pair of sunglasses. (b) Example of a captured image. (c) Detected moving areas in the image while the user shaking the object (region inside the bounding box). (d) Detected region of the hand-held object for further processing of text recognition.

As shown in Fig. 5, while capturing images of the hand-held object, the blind user first holds the object still, and then lightly shakes the object for 1 or 2 s. Here, we apply the efficient multiple Gaussian-mixture-based BGS method to detect the object region while blind user shakes it. More details of the algorithm can be found in [1].Once the object of interest is extracted from the camera image, the system is ready to apply our automatic text extraction algorithm.

IV. Automatic Text Extraction

As shown in Fig. 6, we design a learning-based algorithm for automatic localization of text regions in image. In order to handle complex backgrounds, we propose two novel feature maps to extracts text features based on stroke orientations and edge distributions, respectively. Here, stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Adaboost based text classifier [10].

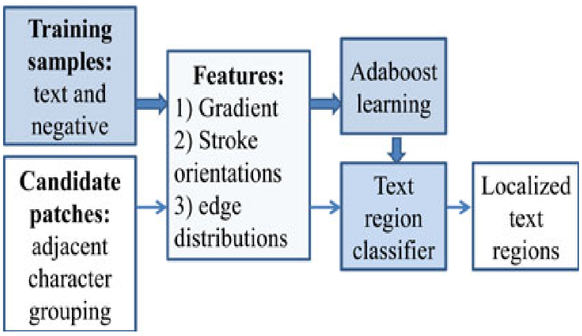

Fig. 6. Diagram of the proposed Adaboost-learning-based text region localization algorithm by using stroke orientations and edge distributions.

V. Text Recognition And Audio Output

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text regions. A text region labels the minimum rectangular area for the accommodation of characters inside it, so the border of the text region contacts the edge boundary of the text character[9]. However, our experiments show that OCR generates better performance if text regions are first assigned proper margin areas and binarized to segment text characters from background.

The recognized text codes are recorded in script files. Then, we employ the Microsoft Speech Software Development Kit to load these files and display the audio output of text information.

Blind users can adjust speech rate, volume, and tone according to their preferences[12, 13].

Fig. 7 Examples of blind persons capturing images of the object in their hands.

VI. Conclusion And Future Work

In this paper, we have described a prototype system to read printed text on hand-held objects for supporting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. This method can effectively distinguish the object of interest from background or other objects in the camera view. To extract text regions from complex backgrounds, we have proposed a novel text localization algorithm based on models of stroke orientation and edge distributions. The corresponding feature maps estimate the global structural feature of text at every pixel. Block patterns project the proposed feature maps of an image patch into a feature vector. Adjacent character grouping is performed to calculate candidates of text patches Prepared for text classification. An Adaboost learning model is employed to localize text in camera-based images. Off-the-shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users.

Our future work will extend our localization algorithm to process text strings with characters fewer than three and to design more robust block patterns for text feature extraction. We will also extend our algorithm to handle non horizontal text strings in complex backgrounds.

REFERENCES

[1] Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons , Chucai Yi, Student Member, IEEE, Yingli Tian, Senior Member, IEEE, and Aries Arditi

[2] C. Yi and Y. Tian, “Assistive text reading from complex background for blind persons,” in roc. Int. Workshop Camera-Based Document Anal. Recognit., 2011, vol.LNCS-7139, pp. 15–28

[3] Advance Data Reports from the National Health Interview Survey (2008).[Online].Available: http://www.cdc.gov/nchs/nhis/nhis_ad.htm.

[4] B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with strokewidth transform,” in Proc. Comput. Vision Pattern Recognition, 2010, pp. 2963–2970.

[5] 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.

[6] International Workshop on Camera-Based Document Analysis and Recognition(CBDAR 2005, 2007, 2009, 2011). [Online]. Available: http://www.m.cs.osakafuu.ac.jp/cbdar2011/

[7] J. Zhang and R. Kasturi, “Extraction of text objects in video documents: recent progress,” in Proc. IAPR Workshop Document Anal. Syst., 2008, pp. 5–17.57

[8] K. Kim, K. Jung, and J. Kim, “Texture-based approach for text detection in images Using support vector machines and continuously adaptive mean shift algorithm,” IEEE Trans. Pattern Anal. Mach. Intel., vol. 25, no. 12, pp. 1631–1639, Dec. 2003.

[9] L. Ma, C. Wang, and B. Xiao, “Text detection in natural images based on multi-scale edge detection and classification,” in Proc. Int. Congr. Image Signal Process., 2010, vol.4, pp. 1961–1965.

[10] N.Nikolaou and N. Papamarkos, “Color reduction for complex document images,”Int. J. Imaging Syst. Technol.,vol. 19, pp. 14–26, 2009

[11] H. Shen, K. Chan, J. Coughlan, and J. Brabyn, “A mobile phone system to find crosswalks for visually impaired pedestrians,” *Technol. Disability*, vol. 20, no. 3, pp. 217–224, 2008.

[12] M. Shi, Y. Fujisawab, T. Wakabayashia, and F. Kimura, “Handwritten numeral recognition using gradient and curvature of gray scale image,” *Pattern Recognit.*, vol. 35, no. 10, pp. 2051–2059, 2002.

[13] P. Shivakumara, T. Phan, and C. L. Tan, “A gradient difference based technique for video text detection,” in *Proc. Int. Conf. Document Anal. Recognit.*, 2009, pp. 66–70.

[14] C. Stauffer andW. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” presented at the IEEE Comput. Soc.Conf.Comput. Vision Pattern Recognit., Fort Collins, CO, USA, 1999.

[15] S. Shoval, J. Borenstein, and Y. Koren, “Auditory guidance with the Navbelt: Acomputerized travel for the blind,” *IEEE Trans. Syst.,Man, Cybern. C. Appl. Rev.*, vol. 28, no. 3, pp. 459–467, Aug. 1998.

AUTHORS

Prof. S. Babu Renga Rajan : Prof and Head of Computer Science in PET Engineering College, Vallioor, doing PhD in Data Mining.
Interested areas are Data Mining and software Engineering, Software Quality Assurance

Mrs.B. Shanmuga Sundari : Assistant Professor of Information Technology in PET Engineering College, Vallioor. Interested areas are Network Security, SOA and web Services.

Mrs. A. Haseena Beevi: Assistant Professor of Information Technology in PET Engineering College, Vallioor. Interested areas are Network Security