IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 4, Issue 6, June 2016

Portable Camera Based Assistive Product Label Reading for Blind and Visually Impaired Individuals

Ashwini V. Mhaske[1], Priti V. Chaudhari[2], Panchashila G. Sabale[3], Amit B. Sanghai[4], Pravin B. Patil[5], Rashmi S. Dshmukh[6]

M.Tech student, Department of E&TC, GCOE, Jalgaon India [1-6]

Abstract: We propose a camera-based assistive text reading framework to help blind persons to read text labels from hand-held objects in their day to day lives. In this paper Camera acts as main vision to capture the image of product packaging and hand held objects. To isolate the object from complex backgrounds, we first propose an effective motion-based method to define a region of interest (ROI) in the image. In the extracted ROI, text localization and recognition are conducted to acquire text information. Then text characters are recognized by off-the-shelf optical character recognition (OCR) software. Using text to speech convertor the extracted texts are output in audio output.

Keywords: camera-based assistive text reading, motion-based method, text localization and recognition, off-the-shelf optical character recognition

I. INTRODUCTION

In worldwide there are 314 million visually impaired people and blind, out of which 45 million are visual impairment which was released by "World Health Organization" in 10 facts regarding blindness. The valuation of The National Health Interview Survey 25.2 million adult Americans are blind or visually impaired. The valuation of The National Census of India there are 21.9 Million disabled people in the country ,out of which more than 15 million people are blind[1][2].

Reading is obviously necessary in today's society. Printed text appears everywhere in the form of receipts, bank statements, reports, restaurant menus, classroom notes, product labels, instructions on medicine bottles, etc.Optical aids, screen readers and video magnifiers can help blind users and those with low vision to access documents, there are few devices which provide good access to common hand-held objects such as productlabels, and objects printed with text such as prescription medication bottles. The ability of people who are blind or those who have significant visual impairments to read printed labels and product packages will enhance their independent living and foster economic and social self-sufficiency so here we are going to propose a system that it useful to blind people.

II. EXISTING SYSTEM

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labelling. For example, portable bar code readers designed that helps blind people to identify different products in an extensive product database can enable users who are blind to access information about these products. But a big limitation is that it is very hard for blind person to find the position of the bar code and to correctly point the bar code reader at the bar code. Some reading systems such as pen scanners, mobile readers might be employed in these similar situations. OCR software is integrated by these systems to offer scanning and recognition function of text; also some systems have integrated voice output. However, these systems perform best with document images with simple backgrounds, a small range of font size, standard fonts, and well-organized characters rather than hand held product packages with multiple decorative patterns. Most of the OCR software cannot directly handle scene images with complex backgrounds.

A number of portable reading systems have been designed specifically for the visually impaired "K-Reader Mobile" runs on a cell phone which allows the visually impaired person to read mail, receipts, fliers, and many other documents [14]. But these documents must be flat, placed on a clear, dark surface. In addition, "K-Reader Mobile" accurately reads black print on a white background. However, it has problems in recognizing colour text or text with colour background.

Although a number of reading assistants systems have been designed specifically for the blind person, but still no existing reading assistant can read text from the complex backgrounds found on many everyday commercial products. Fig.1. shows different examples of printed text from hand-held objects with multiple colours, complex backgrounds, or non flat surfaces.

Fig.1. Examples of printed text from hand-held objects with

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 4, Issue 6, June 2016

III. LITERATURE SURVEY

Different text extraction methods are studied as Text localization and Text recognition in natural scene images of real-world scenes. In [3], a survey was done on several ongoing researches on camera based document analysis such as text detection, extraction, enhancement, recognition and its applications. In [4], [5] methods based on sliding windows are discussed which are more robust to noise, but they have high computational complexity as in this input whole image is scanned with windows of multiple sizes. C. Yi and Y. Tian et.al., proposed a method of adjacent character grouping to calculate the image patches that contain fragments of text strings [6]. Rule based and learning based methods are also proposed for text extraction. Learning based methods model text structure and extract representative text features to build text classifiers. L.Ma et.al., [7] performed classification of text edges by using histograms of oriented gradients and local binary patterns as local features on the support vector machine model. In [8] a finger worn device containing a button camera and microcontroller is implemented. This device assists the visually impaired by reading paper-printed text. Majid Mirmehdi et.al [9] proposed a mobile head mounted device for detecting and tracking text. A real-time text detection algorithm is used for text detection and extraction. Zhu et.al [10] proposed an algorithm for video text detection, text localization and text extraction approach in videos. Christin Wolf et.al, proposed a method for contented based image in multimedia documents [11]. In [12] a mixture-of- Gaussians-based background subtraction technique is used to determine the region of interest in video and moving object region is extracted. Then, text localization and recognition algorithms are used to acquire text details. In [13] a camera based assistive text reading framework is proposed that helps blind persons to read text labels from hand-held objects in their daily lives.

IV. PROPOSED SYSTEM

To help blind persons to read text labels from different patterns with complex backgrounds found on many everyday products of Hand-held objects, have to conceive of a camera-based assistive text reading framework to get the object of interest within the camera view and extract printed text information from the object. The algorithm used in this system can handle clutter background and different patterns. Also this proposed system can extract text information from hand-held objects. In existing systems it is very challenging for blind user to position the object of interest within the camera's view. But, there are still no acceptable solutions. In many stages this problem is approached. In this framework the product object should be appears in the camera's view.

This system going to develop a motion based method to extract the object i.e. region of interest (ROI) from the captured image. After that system framework will perform text recognition only on that region of interest. The text in captured images is mostly surrounded by noise, text characters appear with different scale, size, fonts and colours. Hence this is very difficult task to localize objects and ROI form captured image.

For the text orientations, in this paper assumes that text strings in scene images keep approximately horizontal. Different algorithms have been proposed for text localization in natural scene images. We can categories them in 2 parts as: Rule- Based and Learning-Based. To solve this task, to extract text information from complex background and multiple patterns, here we propose a text localization algorithm which combines rule based analysis and learning based classifier training. This algorithm can efficiently extract text information from hand held objects.

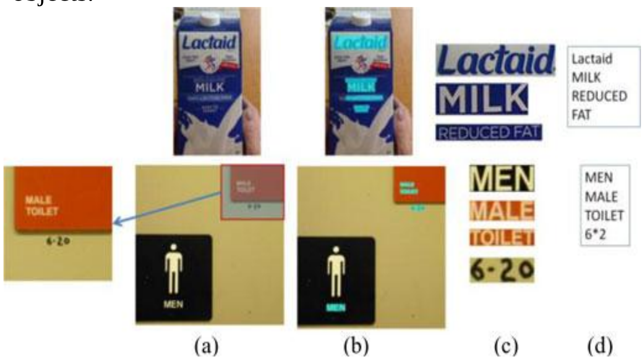

Fig.2. Two examples of text localization and recognition From camera captured Images. (Top) Milk box. (Bottom) Men bath room signage. (a) Camera captured Images. (b) Localized text regions (marked in blue). (c) Text regions Cropped from image. (d) Text codes recognized by OCR.

V. FRAMEWORK AND ALGORITHM OVERVIEW

This paper presents a prototype framework to assist blind people to read text labels from handheld objects. As shown in Fig. 3, the system consists of 3 functional components as:

- Scene Capture
- Data Processing
- Audio Output

Here, the scene capture component collects the images and video containing object of interest. Web camera is used to capture the images of the product, here a camera is attached to pair of sunglasses. The captured images are in RGB.

The data processing component is used for deploying the proposed algorithms, which includes following processes

- Object of interest detection to extract the object from captured image held by user from complex background in the camera view.
- Text Localization to obtain text region containing text information and finally Text Recognition to obtain readable codes from text information.

Finally, the audio output component is to inform the blind user of recognized readable text codes. A Bluetooth earpiece or headphones with mini microphone is used for audio output. This simple hardware configuration shows the portability of the proposed system. Fig. 4 illustrates a work flowchart of the prototype system.

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 4, Issue 6, June 2016

Fig.3. Snapshot demo system including three functional Components for Scene capture, data processing, and audio output.

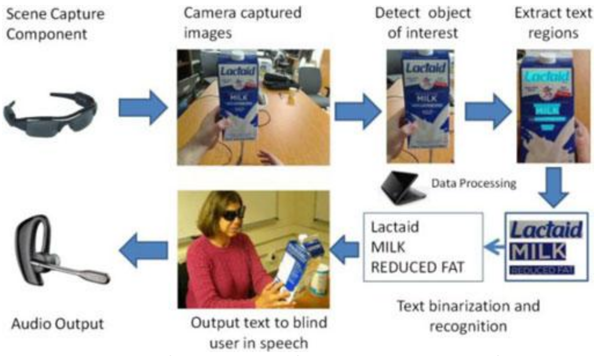

Fig.4. Flowchart of the proposed framework to read text from hand-held objects for blind users.

A frame sequence v is captured by a web camera held by user containing their hand held products and background. The motion based method is applied to extract the text information from object.

$$S = \frac{1}{|V|} \sum_{i} \mathcal{R}(V_i, B) \qquad (1)$$

here Vi represent the ith frame in the captured sequence, |V | represents the number of frames, *B* is the estimated background from motion- based object detection, and *R* denotes the calculated foreground object at each frame. Next, to obtain the object of interest and to extract the text regions, a novel text localization algorithm is proposed. Using layout analysis of horizontal alignment and color uniformity, the candidate text regions are generated.

$$X^C = \text{argmax}_{s \in S} L(s) \qquad (2)$$

Here L (*) represents the suitability responses of text layout and Xc is the candidate text regions from object of interest S.

Then, from a Cascade-Adaboost learning model a text classifier is generated, by using edge distributions and stroke orientations of text characters as features.

$$X = H\left[X^C\right] = H\left[\text{argmax}_{s \in S} L(s)\right] \qquad (3)$$

Where H [.] denotes the Cascade classifier and X is the localized text regions. After text region localization, OCR is applied to perform text recognition in the localized text regions. The recognized text codes are transformed into speech for blind users.

The main contributions used in this prototype system are:

- A motion based algorithm to solve the aiming problem for user by shaking the object of interest for short period.
- A text localization algorithm to extract text regions from clutter background and multiple patterns.
- A portable camera based assistive text reading framework to assist blind users reading text from hand held objects. Algorithms of the system are evaluated over images captured by users using above described techniques.

VI. TEXT RECOGNITION AND AUDIO OUTPUT

In the audio output part the blind user will get speech output of recognized text information. A Bluetooth earpiece or headphones is used for audio output. The task of text recognition is performed by off the shelf OCR to output of informative words of from text localized text regions. The text characters in the recognized text regions are accommodated in minimum rectangular area , hence the border of the text regions contacts the boundary of the text characters. The OCR performs better if text regions are aligned with proper margins and to segment text characters from clutter background. The height and width of each text localized region is enlarged by 10 pixels respectively. Then binarization of text region is performed by OTS's method, where text margin areas are considered as background. Then these recognized text codes are recorded as script file using matlab software. Here blind user can adjust speech rate,tone,volume according to their performance.

VII. **RESULT AND DISCUSSION**

The accuracy of text recognition from image depends on the text extraction. The main problem is segmentation of isolated characters in text extraction. The text regions have been successfully extracted irrespective of the text font and size using the proposed system. The proposed method has provided a comprehensive model of text extraction and recognition in images. In order to evaluate the performance of the proposed method, 10 images have been tested on the system. The proposed system is applied to extract the text regions from background and recognize the characters from extracted text regions. Different stages of system output of a sample image are shown in Fig. 4-a to 4-e.

DOI 10.17148/IJIREEICE.2016.4664

IJIREEICE
ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526
INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 4, Issue 6, June 2016

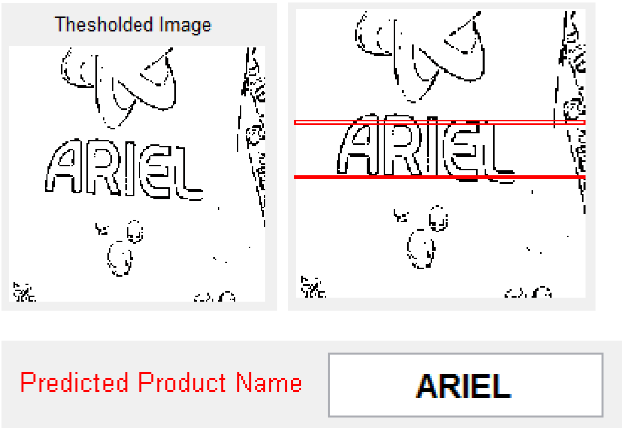

Fig.4-a. Sample Input Color Image, Fig.4-b. Gray Image, Fig.4-c. Thresholded Image, Fig.4-d. Text Region Detection, Fig.4-e. Recognized product name

A. PERFORMANCE MEASURES

Metrics used to evaluate the performance of the system are Precision, Recall rates and F-Score. Precision and Recall rates have been computed based on the number of correctly detected characters in an image in order to evaluate the efficiency and robustness of the system and the Metrics are as follows:

- False Positives (FP) / False alarms are those regions in the image which are actually not characters of a text, but have been detected by the algorithm as text.
- False Negatives (FN)/ Misses are those regions in the image which are actually text characters, but have n ot been detected by the algorithm. Correctly detected and recognized characters are True Positives (TP).
- Incorrectly detected and recognized characters are True Negatives (TN).
- Precision rate (P) is defined as the ratio of correctly detected characters to the sum of correctly detected characters plus false positives as represented in equation below.

$$P = \frac{TP}{TP + FP} * 100$$

- Recall rate (R) is defined as the ratio of the correctly detected characters to sum of correctly detected characters plus false negatives as represented in equation below.
- F-score is the harmonic mean of recall and precision rates. It is represented in equation as,

$$F - score = \frac{2 * P * R}{P + R}$$

- Accuracy (A) is defined as the ratio of the sum of correctly and incorrectly detected and recognized characters to sum of correctly and incorrectly detected and recognized characters, false positives and false negatives. It is represented in equation as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100$$

Precision and recall rates are useful as measures to determine the accuracy of proposed algorithm in locating correct text regions and eliminating non-text regions and recognizing correct text characters from images. Precision and recall rates, f-score and accuracy for different images used in this system are analyzed to determine the success and limitations. The comparison is based on the accuracy of the results obtained with precision and recall rates and F-Score.The average performance of the proposed system is shown in Table 6.1.

Table 6.1: Performance Measures of The Proposed System

Images	Precision Rate (%)	Recall Rate (%)	Fscore (%)	Accuracy (%)
1	100	100	100	100
2	59.13	100	58.33	47.37
3	100	100	100	100
4	93.33	100	96.55	96.55
5	93.33	100	96.55	94.12
6	53.17	100	80	90.91
7	71.43	93.13	83.33	71.43
8	41.78	100	57.14	45.45
9	100	100	100	100
10	100	100	100	100

The average precision rate of the proposed system is 81.27 %, which means that there is less number of false positives. The system produces average recall rate 99.313%, as only some relatively weak text regions in the images are missed.The proposed method has been tested on various types of images and the experimental results show that proposed method outperforms the other methods.

VIII. CONCLUSION

In this paper, a prototype system framework to read printed text on hand held object for assisting blind person is described. In order to solve the common aiming problem, we have proposed a motion based technique to detect the object of interest where blind person has to shake the object for short period. This method can efficiently detect the object of interest from clutter background in the camera vision. A novel text localization algorithm is proposed to localize text region. Off-the-shelf OCR is used to perform text recognition on text localized regions and then recognized text codes are transformed to speech for blind person.

REFERENCES

[1] National Census of India- http://censusindia.gov.in/Census_And_You/disabled_population.aspx

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

[3] K. Jung, K.I. Kim, A.K. Jain, "Text information extraction in images and video: a survey", Pattern Recognition, 37 977-997S, 2004.

[4] X. Chen and A. L. Yuille. "Detecting and readingtext in natural scenes." Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 2:366–373, 2004.

[5] R. Lienhart and A. Wernicke, "Localizing and segmentingtext in images and videos," Circuits and Systems for VideoTechnology, vol. 12, no. 4, pp. 256 –268, 2002.

[6] C. Yi and Y. Tian, "Text string detection from natural scenes by structurebased partition and grouping," IEEE Trans. Image Process., vol. 20, no. 9,pp. 2594–2605,Sep. 2011.

[7] L. Ma, C. Wang, and B. Xiao, "Text detection in natural images based onmulti-scale edge detection and classification," in Proc. Int. Congr. ImageSignal Process., 2010, vol. 4, pp. 1961–1965.

[8] Nanayakkara, S., Shilkrot, R., Yeo, K. P., and Maes, P." EyeRing: a finger-worn input device for seamless interactions with our surroundings". In Augmented Human ,2013.

[9] Carlos Merino-Gracia,, Karel Lenc,Majid Mirmehdi, " A Head-Mounted Device for Recognizing Text in Natural Scenes "Lecture Notes in Computer Science Volume 7139, pp 29-41,2012.

[10] Chengjun Zhu, yuanxin Quyang, Lei Gao, Zhenyong Chen, Zhang Xiong."An Automatic Video Text Detection, Localization and Extraction Approach". School of computer Science and Technology, Beihang University, SITIS - Pages 166-175-2006.

[11] Christin Wolf, Jean-Michel Jolion, and Francoise Chassaing."Text localization, enhancement and binarization in multimedia documents". In Int. Conf. on Pattern recognition, pages 1037–1040, Quebec City, Canada., August 2002.

[12] Rajkumar N, Anand M.G, Barathiraja N, "Portable Camera-Based Product Label Reading For Blind People", International Journal of Engineering Trends and Technology (IJETT) – Volume 10 ,Issue No : 1 - Apr 2014.

[13] Chucai Yi, Student Member, IEEE, YingLi Tian, Senior Member,IEEE, Aries Arditi," Portable Camera-based Assistive Text and ProductLabel Reading from Hand-held Objects for Blind Persons" IEEE/ASMETransactions on Mechatronics.

[14] 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