ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Text and Context Label Perusal from Hand-Held Objects Using Transportable Digicam For Visually Challenged Folks

R. Ramesh, M.Srinivasulu

M. Tech, Dept. of ECE, SITS, JNT University Aanthapur, Tirupati, AP, India.

Assistant Professor, Dept. of ECE., SITS, JNT University Aanthapur, Tirupati, AP, India.

ABSTRACT: We recommend a digital camera-founded assistive text studying framework to help blind humans study textual content material labels and product packaging from hand held objects of their everyday lives. To isolate the item from cluttered backgrounds or special surrounding objects in the digicam view, we first endorse an effective and powerful movement centered procedure to outline a neighborhood of curiosity (ROI) within the video via utilising asking the user to shake the thing. This approach extracts moving object area with the help of a mixture-of-Gaussians situated old previous subtraction process. Within the extracted ROI, text localization and realization are carried out to gather text understanding. To robotically localize the textual content regions from the thing ROI, we recommend a novel textual content localization algorithm by means of studying gradient aspects of stroke orientations and distributions of area pixels in an Adaboost model. Textual content characters within the localized text areas are then binarized and famous by way of off-the-shelf optical persona cognizance (OCR) application. The famous text codes are output to blind purchasers in speech. Efficiency of the proposed textual content localization algorithm is quantitatively evaluated on ICDAR-2003 and ICDAR-2011 potent learning Datasets. Experimental end result expose that our algorithm achieves the state-of-the-arts. The proof-of-proposal prototype can be evaluated on a dataset gathered utilising 10 blind men and women, to investigate the effectiveness of the approach's hardware. We discover patron interface issues, and examine robustness of the algorithm in extracting and studying textual content from special objects with complicated backgrounds..

KEYWORDS: blindness; assistive instruments; text learning; hand held objects, text neighborhood localization; stroke orientation; distribution of discipline pixels; OCR;

I. INTRODUCTION

In today's society,reading is obviously considered as an essential task. The Printed text is present everywhere in the form of receipts,classroom handouts, bank statements, reports, restaurant menus, product packages, instructions present on medicine bottles etc. The capability of the people who are blind or they have the significant visual impairments for reading the printed labels and product packages which may enhance the independent living and foster economic and also social self-sufficiency.

Now-a-days, there already exists a few systems which have some promise to portable use, but they are not able to handle the product labeling. For example, the portable bar code readers are designed for helping the blind people to identify the different products that are present in an extensive product database may enable the users who are blind for accessing of information regarding these products via speech and Braille. The limitation is that it is regarded as very hard to blind users for finding the position of the bar code and in order to correctly point the bar code reader present at the bar code. Some of the reading assistive systems known as pen scanner may be employed with these similar situations. Even though a number of reading assistants that has been specifically designed to the visually impaired, to the knowledge, no reading assistant that exists may read the text from different kinds of challenging patterns and the backgrounds that are found on many of the everyday commercial products.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

A number of transportable studying assistants were designed chiefly for the visually impaired [13, 20, 24, 29]. KReader cell runs on a phone telephone and allows the consumer to read mail, receipts, fliers and plenty of other files [13]. Nonetheless, the document to be learn have to be practically flat, placed on a transparent, darkish surface (i.E., a non-cluttered heritage), and incorporate mostly textual content. In addition, KReader cellular thoroughly reads black print on a white background, but has problems recognizing coloured text or textual content on a colored historical past. It cannot learn textual content with intricate backgrounds, textual content printed on cylinders with warped or incomplete snap shots (such as soup cans or medicine bottles). Moreover, these systems require a blind person to manually localize areas of curiosity and text areas on the objects in most instances.

Even though a quantity of reading assistants had been designed chiefly for the visually impaired, to our abilities, no present reading assistant can learn textual content from the varieties of challenging patterns and backgrounds determined on many daily business products. As proven in Fig. 1, such text knowledge can appear in a couple of scales, fonts, colors, and orientations. To aid blind individuals to read textual content from these types of hand held objects, we've got conceived of a digicam-founded assistive text reading framework to monitor the thing of curiosity inside the digicam view and extract print text expertise from the thing. Our proposed algorithm can without problems manage complex history and multiple patterns, and extract textual content understanding from each handheld objects and regional signage, as shown in fig.

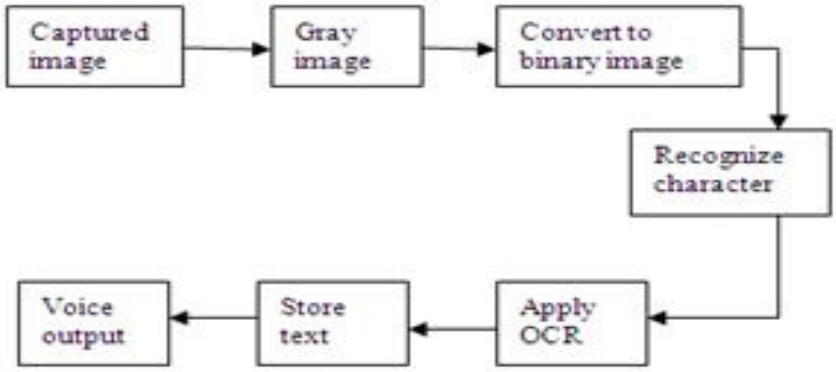

Fig:1. Block Diagram of Text Reading

In assistive studying methods for blind persons, it is rather difficult for customers to position the item of interest within the center of the digicam's view. As of now, there are still no perfect options. To be certain the hand-held object appears within the digicam view, we use a digicam with sufficiently wide angle to accommodate customers with only approximate aim. This will usually effect in different textual content objects showing within the camera's view (for example whilst searching at a supermarket). To extract the hand-held object from the camera snapshot, we advance a movement-founded approach to obtain a region of interest (ROI) of the article. Then we participate in textual content consciousness simplest in this ROI.

It is a difficult crisis to routinely localize objects and textual content areas of interest from captured photos with complex backgrounds, in view that text in captured images is surely surrounded with the aid of quite a lot of background outlier 'noise', and text characters ordinarily appear in a couple of scales, fonts, and colours.
For the text orientations, this paper assumes that text strings in scene photos maintain approximately horizontal alignment. Many algorithms were developed for localization of text areas in scene pictures. We divide them into two classes, rule-based and finding out-based.

Rule-founded algorithms apply pixel stage photograph processing to extract textual content knowledge from predefined textual content facets such as persona dimension, side ratio, facet density, persona structure, and colour uniformity of textual content string, etc. Phan et al. [23] analyzed part pixel density with the Laplacian operator and employed maximum gradient differences to establish textual content regions. Shivakumara et al. [31] used gradient difference maps and carried out international binarization to acquire textual content regions.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Epshtein et al. [7] designed stroke width transforms to localize textual content characters. Nikolaou et al. [21] applied colour discount to extract text in uniform colors. In ref. [5], colour situated text segmentation is performed by way of a Gaussian combo model for calculating a self belief price for textual content regions. This variety of algorithm tries to define a universal characteristic descriptor of text. Finding out-centered algorithms, then again, model textual content structure and extract consultant textual content elements to construct textual content classifiers. Chen et al. [4] awarded 5 varieties of Haar-headquartered block patterns to teach textual content classifiers in an Adaboost finding out model. Kim et al. [12] viewed text as a certain texture and analyzed the textural elements of characters through a support vector computer (SVM) mannequin. Kumar et al. [14] used Globally Matched Wavelet filter responses of textual content structure as features. Ma et al. [18] carried out classification of text edges by means of utilising histograms of oriented gradients and local binary patterns as neighborhood points on the SVM model. Shi et al. [30] employed gradient and curvature features to model the grey scale curve for handwritten numeral consciousness under a Bayesian discriminant perform. In our research team, we've got beforehand developed rule-situated algorithms to extract text from scene portraits [38-40]. A survey paper about laptop-vision-situated assistive technologies to support humans with visible impairments will also be determined in ref. [20].

In solving the challenge at hand, to extract text expertise from difficult backgrounds with multiple and variable text patterns, we here endorse a textual content localization algorithm that mixes rule-centered layout evaluation and studying-founded textual content classifier coaching, which define novel function maps headquartered on stroke orientations and area distributions. These in flip generate representative and discriminative textual content points to distinguish textual content characters from historical past outliers.

II. RELATED WORK

A.The Connected regions approach:

The main idea behind this approach is that a letter may be considered as a homogeneous region (by using the restrictions) and hence it could be very useful for dividing the frame into the homogeneous regions. For computation of such a division, a split-and-merge algorithm is seemed to be very adequate. Its concept is that if there is a non-homogeneous region is present, then it is divided into four regions. If the two adjacent regions are of homogeneous then they may be merged. By using some size characterizations for the text (not so big and not so small) the inadequate regions can be deleted. The same process may be executed for the various frames and then the results may be temporally integrated for keeping only the elements that are present in all of the frames.

B. The Edge detection approach:

The main idea behind this approach is that the text will contrast a lot along with background. The text may have a well-defined edge which makes it possible for clear identification of the text. Hence, by using the concepts of edge detection it seems to be a good idea. The edge detection that is based on the certain predefined critical angles is known as threshold angles. The lines which coincide by these thresholds were identified as the edges of text. After the computation of edge lengths, the length number for edges present in x and y direction can be calculated and if it is greater than a particular threshold then it may be regarded as a text area. Then each of the text area can be binarized by using the luminance. For binarizing thresholds are used. The intermediate luminescence is considered as the threshold value and the areas that are brighter than the threshold is represented as white and the other areas are represented as black. Hence in final result, the text may be in white and the background may be in black (or it may be the inverse). Lastly, the similar process can be executed for the various frames and the results may be temporally integrated for keeping only the elements that are present in all of the frames.

C. Extraction of Text from Video.

Hence, the methods that depend on page models and the top-down page segmentation methods were not appropriate—it is required for usage of more locally adaptive bottom-up methods by less dependence on the regularity, the uniform font size and the availability of multiple lines etc. Most of the document analysis methods used for text detection will depend on an initial high quality binarization step taken from the original gray-scale image. During this case the background**:** The Text recognition present in document images has become an active research area to some time. However, the text recognition in case of broadcast quality digital video is considered as a problem for requiring

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

various approaches. Unlike the document images, the video frames will tend to have the text that is not in orderly columns but it is in widely scattered areas and fewer are of separated lines. Typically, the video frames are of noisy, low-resolution and full-color by interlace artifacts.

III. PROPOSED SYSTEM

- The proposed algorithm may effectively handle the complex of background and multiple patterns and it can extract the text information from both the hand-held objects and the nearby signage.
- In case of assistive reading systems for the blind persons, it is a very challenging task to the users for positioning the object of interest present within the center of the camera's view. Still now, there exist no acceptable solutions.
- This problem is approached in stages. In order to make sure the hand-held object will appear in the camera view, a camera with sufficiently wide angle is used for accommodating the users by only the approximate aim.
- Often, this will result in the other text objects that appear in the camera's view (for example, during shopping at a supermarket.

In day-to-day life, many of the electrical and electronic circuits and kits are frequently used that are designed by using the embedded systems technology. The electrical and electronics engineering students and also the electronics and communications engineering students required to design the final year electronics projects in order to gain hands on experience by the real time embedded systems and also for fulfilling the criteria of engineering graduation. The engineering for final year electronics projects were designed by using embedded systems and its applications. The computers, laptops, mobile phones, digital electronic systems, tablets and other electrical and electronic gadgets are also designed by using embedded systems. Hence, what is embedded system and the applications of embedded systems are described below.

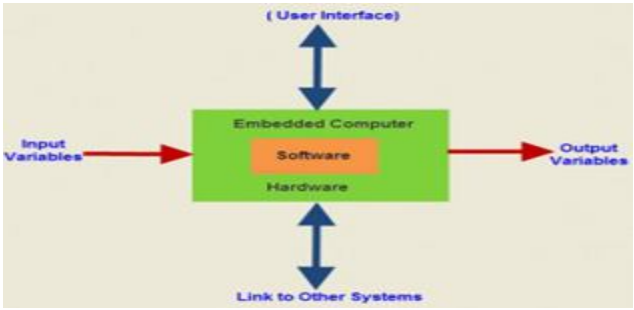

Fig .2: Block diagram Of Embedded System

A. *Embedded System*

The electronic system that integrates the hardware circuitry along with the software programming techniques in order to provide the project solutions is defined as embedded systems. With the use of this embedded system technology, the complexity regarding the circuits may be reduced for a great extent that will further reduce the cost and its size. The Embedded system was primarily developed by Charles Stark for reduction of the size and the weight of the project circuitry.

Basically, an embedded system is an electronic system which can be programmed or non-programmed for operation, organization and to perform the single or multiple tasks according to the application. In case of real time embedded systems, all of the assembled units will work together according to the program or by set of rules or the code that is embedded into the microcontroller. But, with the use of this microcontroller programming techniques only a limited range of problems will be solved.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

An embedded system is defined as an integration of hardware and software, the software that is used in the embedded system is a set of instructions that is termed as a program. The microprocessors or the microcontrollers that are used in the hardware circuits of the embedded systems were programmed for performing specific tasks by considering the set of instructions. Primarily, these programs are written by using any of programming software such as Proteus or Lab-view by using any of the programming languages like C or C++ or embedded C. Then the program will be dumped into the microprocessors or the microcontrollers which are used in the embedded system circuits.

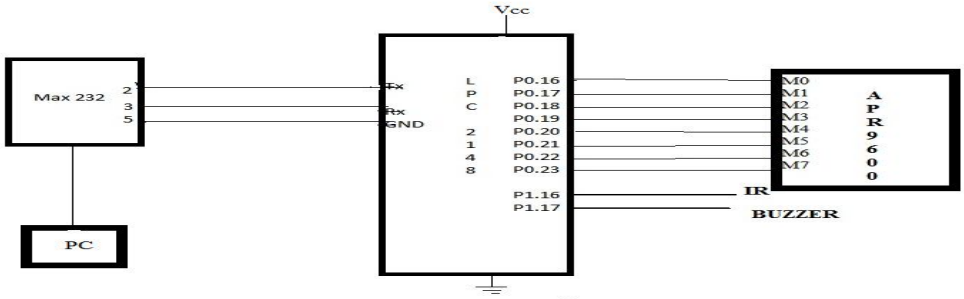

Fig.3.Circute Diagram Of Text

The text data captured from camera and processing through MATLAB is given to LPC 2148 microcontroller board through RS 232 serial communication and then it processing and drive the voice module APR 9600, which produce voice corresponding sound stored for the text send from MATLAB. When obstacle is present except label objects, the ir sensor is activated and the microcontroller drives the buzzer given sound alert to the blind.

Fig4:camera-captured images; localized text regions (marked in blue); text regions cropped from image; text codes recognized by OCR. The top-right portion of the bottom image that contains text is also shown in a magnified callout, for clarity.

It is a difficult crisis to mechanically localize objects and text areas of interest from captured pictures with elaborate backgrounds, seeing that text in captured pictures is without doubt surrounded by using quite a lot of heritage outlier 'noise', and textual content characters on the whole show up in more than one scales, fonts, and colors. For the text orientations, this paper assumes that text strings in scene images maintain approximately horizontal alignment. Many algorithms have been developed for localization of text areas in scene portraits. We divide them into two categories, rule-headquartered and learning-situated.

Rule-centered algorithms observe pixel degree photo processing to extract textual content understanding from predefined text aspects equivalent to character size, part ratio, facet density, character constitution, and colour uniformity of textual content string, etc. Phan et al. [23] analyzed side pixel density with the Laplacian operator and

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

employed highest gradient differences to determine text areas. Shivakumara et al. [31] used gradient difference maps and carried out international binarization to receive textual content regions. Epshtein et al. [7] designed stroke width transforms to localize text characters. Nikolaou et al. [21] applied colour reduction to extract textual content in uniform colours. In ref. [5], color established text segmentation is performed by way of a Gaussian combination model for calculating a self belief price for textual content areas. This style of algorithm tries to define a common function descriptor of text.

Learning-headquartered algorithms, however, model textual content constitution and extract consultant text aspects to construct textual content classifiers. Chen et al. [4] provided 5 types of Haar-founded block patterns to educate textual content classifiers in an Adaboost finding out mannequin. Kim et al. [12] considered textual content as a certain texture and analyzed the textural features of characters by a support vector desktop (SVM) model. Kumar et al. [14] used Globally Matched Wavelet filter responses of textual content constitution as aspects. Ma et al. [18] performed classification of textual content edges by using histograms of oriented gradients and nearby binary patterns as local facets on the SVM model. Shi et al. [30] employed gradient and curvature facets to model the gray scale curve for handwritten numeral realization under a Bayesian discriminant function. In our research workforce, we have beforehand developed rule-situated algorithms to extract text from scene pictures [38-40]. A survey paper about computer-vision-founded assistive applied sciences to support people with visual impairments can also be located in ref. [20].

C. Adaboost studying of text

Elements situated on the function maps of gradient, stroke orientation and aspect distribution, a textual content classifier is knowledgeable from an Adaboost studying mannequin. Picture patches with fixed size (peak forty eight pixels, width ninety six pixels) are accumulated and resized from pics taken from the ICDAR-2011 powerful studying competitors [11] to generate a coaching set for finding out facets of textual content. We generate confident coaching samples by using scaling or slicing the bottom reality text areas, in step with the side ratio of width to height . To train a powerful text classifier, we make certain that almost all optimistic coaching samples will incorporate 2 to 4 textual content characters. We construct a relationship between the width-to-height part ratio and the number of characters of floor-truth textual content areas. It indicates that the ground-fact areas with 2 to 4 text characters have width-to-peak ratios between zero.Eight and a couple of.5, at the same time the ones scale back than zero.8 most likely have less than 2 characters and those higher than 2.5 in general have greater than 4 characters.

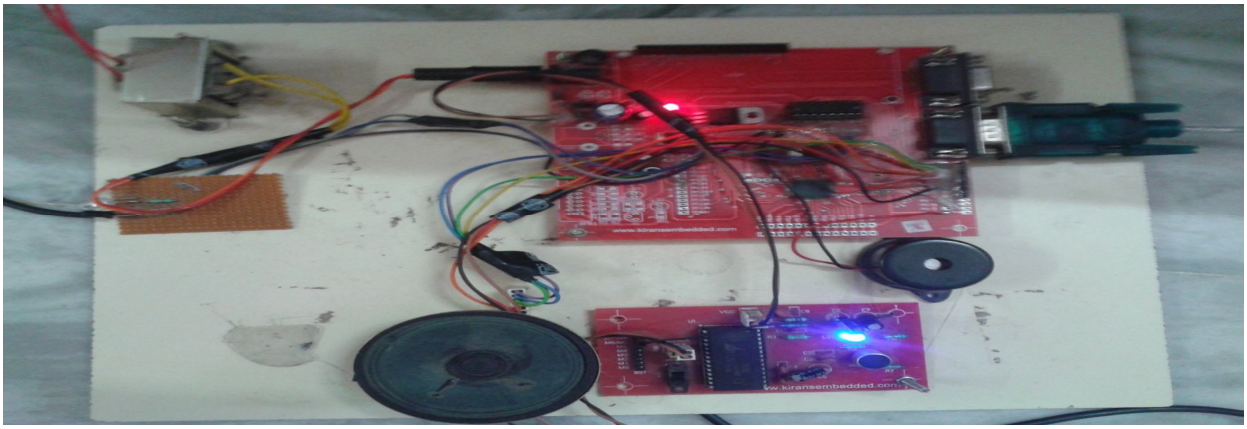

Fig: 5. A snapshot of our demo system, including three functional components for scene capture, data processing and audio output.

The automated ROI detection and text localization algorithms had been independently evaluated as unit assessments to make certain effectiveness and robustness of the whole procedure. We subsequently evaluated this prototype process of assistive textual content reading using photographs of hand held objects captured by way of 10 blind shoppers in individual. Two calibrations have been applied to arrange for the procedure scan. First, we advised blind customers to challenge handheld object inside the digital camera view. Considering it's tricky for blind customers

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

to check out their held objects, we employed a digital camera with a reasonably vast perspective. In future packages, we can add finger point detection and monitoring to adaptively instruct blind customers to try the thing.

Subsequent, we evaluated the consumer-captured dataset of object textual content. The dataset was as soon as manually annotated through labeling the areas of the item of targets and the textual content areas inside the object of curiosity areas. In our algorithm evaluation, we outlined a neighborhood as adequately detected if the ratio of the overlap discipline of a detected vicinity and its floor reality regional is a minimum of 3/4. Experiments established that 225 of the 312 ground reality textual content areas were hit by the use of our localization algorithm. By using utilising using the same evaluation measures as above experiments, we bought precision zero.Fifty two, bear in mind 0.Sixty two, and f-measure 0.Fifty two on this dataset. The precision is lower than that on the robust learning Dataset. The pictures within the consumer-captured dataset have curb resolutions and further compact distribution of text expertise, in order that they generate low-high-best subject maps and text boundaries, which outcomes in wrong spatial layouts and text structure elements.

Fig. 6. Examples of training samples with width-to-height ratio 2:1. The first two rows are positive samples and the other two rows are negative samples.

IV. PROTOTYPE EVALUATION

The automated ROI detection and text localization algorithms had been independently evaluated as unit tests to ensure effectiveness and robustness of the whole method. We subsequently evaluated this prototype method of assistive text reading using pictures of hand-held objects captured by way of 10 blind customers in man or woman. Two calibrations were applied to arrange for the system experiment. First, we suggested blind users to situation handheld object within the digital camera view. On the grounds that it's complex for blind users to try their held objects, we employed a camera with a reasonably wide angle. In future programs, we will add finger point detection and tracking to adaptively instruct blind users to attempt the thing.

Subsequent, we evaluated the user-captured dataset of object textual content. The dataset was once manually annotated via labeling the areas of the object of pursuits and the textual content areas inside the object of curiosity areas. In our algorithm analysis, we defined a neighborhood as properly detected if the ratio of the overlap field of a detected vicinity and its ground truth neighborhood is a minimum of 3/4. Experiments confirmed that 225 of the 312 ground fact textual content areas have been hit by means of our localization algorithm. By using utilizing the same evaluation measures as above experiments, we bought precision zero.52, take into account 0.Sixty two, and f-measure 0.52 on this dataset. The precision is lower than that on the potent studying Dataset. The photographs in the consumer-captured dataset have scale down resolutions and extra compact distribution of text information, so that they generate low-high-quality area maps and text boundaries, which effect in improper spatial layouts and text structure elements.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Fig:7. Some results of text localization on the user-captured dataset, where localized text regions are marked in blue.

V. CONCLUSION AND FUTURE ENHANCEMENT

In this project, it has been described a prototype system for reading the printed text based on hand-held objects to assist the blind persons. For solving the common aiming problem to blind users, a method is proposed which can be effectively distinguished the object of interest from the background or to other objects present in the camera view. For extracting the text regions that are obtained from complex backgrounds, a novel text localization algorithm is proposed based on the models of stroke orientation and the edge distributions. The corresponding feature will map the estimate for global structural feature of text at every of the pixel. The Block patterns project the proposed feature which maps for an image that is patched into the feature vector. The Adjacent character grouping that is performed for calculating the candidates for the text patches that are prepared to the text classification. The Off-the-shelf OCR is used for performing the word recognition present on the localized text regions and to transform it into audio output to the blind users.

The future work will extend the localization algorithm for process text strings along with the characters that are fewer than three and for designing the more robust block patterns to text feature extraction. It also extends this algorithm for handling non-horizontal text strings. Furthermore, it will address the significant human interface issues that are associated for reading the text by the users of blind.

REFRENCES

1. Chucai Yi, YingliTian, 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.
2. C. Yi and Y. Tian, "Assistive text readingfrom complex background for blind persons," in *Proc. Int. Workshop Camera-Based Document Anal.Recognit.*, 2011, vol. LNCS-7139, pp. 15–28.
3. X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in Proc. Comput. Vision Pattern Recognition., 2004, vol. 2, pp. II-366–II-373.
4. S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi, "Text extraction and document image segmentation using matched wavelets and MRF model," IEEE Trans Image Process., vol. 16, no. 8, pp. 2117–2128,Aug. 2007.
5. S. Shoval, J. Borenstein, and Y. Koren. Auditory Guidance with the Navbelt-a Computerized Travel for the Blind. In *IEEE Trans. on Systems, Man, and Cybernetics, Part C*, Vol. 28, No. 3, 1998.
6. C. Stauffer and W.E.L. Grimson, "Adaptive Background mixture Models for Real-time Tracking", In *CVPR*, Vol. 2, ISBN: 0-7695-0149-4, 1999.
7. X. Chen, J. Yang, J. Zhang and A. Waibel, "Automatic detection and recognition of signs from natural scenes," In *IEEE Transactions on image processing*, Vol. 13, No. 1, pp. 87-99, 2004.