International Journal of Electrical, Electronics and Computer Engineering **3**(2): 85-90(2014)

ISSN No. (Online): 2277-2626

Robust Approach to Recognize and Localize Text from Natural Scene Images

Khushbu C. Saner*

**Student M.E. Department of Computer Science & Engineering,*
Matoshri College of Engineering & Research Center, Nashik, (MS) India,

(Corresponding author: Khushbu C. Saner)
(Received 05 August, 2014 Accepted 11 October, 2014)

ABSTRACT: Text detection and localization in natural scene pictures is used for content-based image analysis. The experimented and designed approach for text detection is robust and localizes texts in natural scene pictures. This system is basically divided into three region 1) Pre-processing 2) Connected Component Analysis 3) Optical Character Recognition. From pre-processing step, get the binary image. To filter out the non-text components, a conditional random field (CRF) model considering unary part properties and binary discourse part relationships with supervised parameter learning is projected .After that recognized text is localized in original image and then text part area units are classified into text lines. After performing the line partition the character recognition will be done using Optical character recognition to recognize the character. Results are evaluated on the natural image dataset. The experimented extended approach yields higher precision and recall performance compared with progressive ways.

Keywords: Index Terms: Conditional random field (CRF), connected component analysis (CCA), text detection, text localization, Text Recognition, Optical Character Recognition (OCR).

I. INTRODUCTION

With the increasing use of digital image capturing devices, such as digital cameras, mobile phones and PDAs, content-based image analysis techniques are receiving more attention in coming years. Among all the contents in images, text information has fostered great interests, since it can be easily understood by both human and computer, and finds wide applications such as translation, mobile text recognition, and content-based web image search. In this system text recognition and detection is done by using three method preprocessing, connected component and optical character recognition. Basically literature survey is divided into two parts Region based and Connected Component. Table 1 gives the brief about the region based approach

Table 1: Region-based approach.

Sr	Researcher Name	Year	Technique	Description	Problem
1	V. Wu, R. Manmatha and E. cRiseman	1997	Finding text in Images	k-means clustering and morphological operators are used to group text pixels into text regions	Non text region cannot be removed
2	Y. Zhong H. J. Zhang and A. K.Jain,	2000	Automatic caption localization in compressed video	Discrete cosine transforms (DCT) and wavelet decomposition, to extract features	the approach will not work properly color information is not utilized

3	H. P. Li, D. Doermann, and O. Kia	2000	Automatic text detection and tracking in digital video	Proposed an algorithm for detecting texts in video by using first and second order moments of wavelet decomposition responses as local region features classified by a neural network Classifier.	Poor OCR performance Semantic indexing need extensive training
4	M. Lyu, J. Q. Song, and M. Cai,	2002	Comprehensive method for multilingual video text detection, localization, and extraction,	Detect candidate text edges of various scales with a Sobel operator. A local thresholding procedure used to used to filter out non text edges, and the text regions are then grouped into text lines by recursive profile projection analysis Only focuses on Chinese and English language.	1) it cannot detect motion texts due to the assumption of stationary text 2)no horizontally aligned texts cannot be localized
5	R. Lienhart and A.Wernicke,	2002	Localizing and segmenting text in images and videos,	Computes the gradient map with color derivative operators.	Localized text lines are then scaled to a fixed height of 100 pixels

Table 2 will discusses the earlier techniques for connected component based approach

Table 2: Connected component (CC)-based Approach.

Sr. No	Researcher Name	Year	Technique	Description	Problem
1	D.-Q. Zhang And S.-F. Chang,	2004	Learning to detect scene text using a higher-order MRF with belief propagation,	Presented Markov random field (MRF) method for exploring the neighboring information of components. After building up a component adjacency graph, a MRF model integrating a first-order component term and a higher order contextual term is used for labeling components as text or non-text.	Not included the text regions missed in the automatic segmentation process The region detection miss is mainly due to the small text size
2	K. H. Zhu, F. H. Qi, R. J. Jiang, L. Xu, M. Kimachi, Y. Wu, and T.Aizawa,	2005	Using Adaboost to detect and segment characters from natural senes, first use a onlinear local binarization algorithm to segment candidate CCs. Several types of component features, including geometric	First use a nonlinear local binarization algorithm to segment candidate CCs. Several types of component features, including geometric, edge contrast, shape regularity, stroke statistics and spatial coherence features, are then defined to train an Adaboost classifier for fast coarse-to-fine pruning of non-text components.	Minimizes classification error not number of false negatives because of edge contrast features

3	H. Takahashi,	2005	Region graph based text ex-traction from outdoor images,	extracts candidate text Components using a Canny edge detector in color images. A region adjacency graph is then built up on the extracted CCs and some heuristic rules based on local component features and adjacent component relationships are Designed to prune non-text components.	Cannot segment text component accurately
4	Y.X. Liu, S. Goto, and T. Ikenaga,	2006	A contour-based robust algorithm for text detection in color images	Extracts candidate CCs based on edge contour features and removes non-text Components by wavelet feature analysis. Within each text component region, a GMM is used for binarization by fitting the Gray-level distributions of the foreground and background pixel clusters.	Minimizes classification error not number of false negatives.
5	X.B. Liu, H. Fu, and Y.D. Jia,	2008	Gaussian mixture modeling and learning of neighboring characters for multilingual text extraction in images	This employs a GMM to fit third-order neighboring information of components Using a specific training criterion: maximum minimum similarity (MMS). Their experiments show good performance on their multilingual image datasets.	How to discriminate character region from non character region base on relation between neighboring character

In Table 3 will give the hybrid approach which includes merging of above approachs.

Table 3: Connected component (CC)-based and region based.

Sr.No.	Researcher Name	Year	Technique	Description	Problem
1	Yi-Feng Pan, Xinwen Hou, and Cheng-Lin Liu,	2011	A Hybrid Approach to Detect and Localize Texts in Natural Scene Images	Presented a Markov random field (MRF) method for exploring the neighboring information of components. After building up a component adjacency graph, a MRF model integrating a first-order component term And a higher order contextual term is used for labeling components as text or non-text.	Not included the text regions missed in the automatic segmentation process The region detection miss is mainly due to the small text size

II. PHASES OF ROBUST APPROACH

Fig. 1 show the phages of system. Basically system is divided into three phases namely preprocessing, connected component and optical character recognition. In pre-processing step will get the binary image as output. This image will pass to connected component with help of condition random field features will get the text component and after getting the text component by applying the line partition separate out the words .Then using OCR we recognize the character.

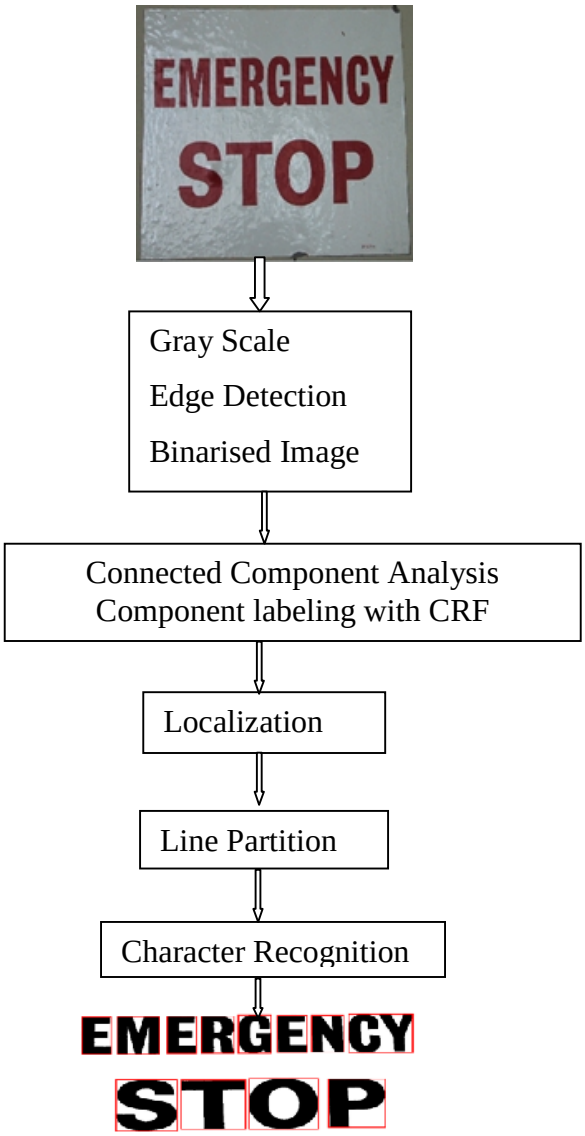

Fig. 1. Flow chart of system.

Algorithm Used for implementation:

(i) Connected Component Analysis -

-Reads a grey scale image and perform thresholding with value = 43 and outputs a binary image.

-Input will take binary image and produces screen print out with labels. And then scan from left to right and top to bottom

-If value is non-zero and if all neighbors are zeros, get a new label

-If all non-zeros neighbors have the same label, take the label from a neighbor.
-If non-zeros neighbors have different labels, get the min label and update equivalent table.
-As finished with above process scan from bottom to top.whn he center pixel is not zero.
-If all neighbours are zero then do nothing
-If all non-zero neighbors have the same label, get label from a neighbor
-If neighbors have different labels, get min from neighbors and also look up equivalent table to get the same label with less value.

(ii) Character Recognition (OCR) -
-After finding out the connected component we apply the localization. For localization Krushkal algorithm is used.
-For recognition of character an image of every character must be converted to appropriate character code
-Sometimes it will create several character codes for uncertain images. For example for recognition of the images I character produces '-','1', and'1' codes and final character will be selected.
-Display the character.
(iii)Otsu Binarization –
Otsu's method [1] is used to automatically perform clustering-based image thresholding, or, the reduction of a gray level image to a binary image.
Compute histogram and probabilities of each intensity level

(a) Set up initial $\omega_i(0)$ and $\mu_i(0)$
(b) Step through all possible thresholds $t = 1 \ldots$ maximum intensity
1. Update ω_i and μ_i
2. Compute $\sigma_b^2(t)$
(c) Desired threshold corresponds to the maximum $\sigma_b^2(t)$
(d) You can compute two maxima (and two corresponding thresholds). $\sigma_{b1}^2(t)$ is the greater max and $\sigma_{b2}^2(t)$ is the greater or equal maximum
(e) Desired threshold= $\frac{\text{threshold}_1 + \text{threshold}_2}{2}$

III. RESULT ANALYSIS

Efficiency of the robust system is a key factor for success of text localization. Here efficiency is obtained with the help of how much data is localized and detected from the image. The effectiveness of robust system is measured mainly by three parameters such as precision and recall. A high precision means less percentage of irrelevant images in the retrieval i.e few false alarms. A high recall means less percentage of failure of relevant images to be retrieved.

Precision = {relevant data ∩ retrieve data }/ retrieved data
Recall ={relevant data ∩ retrieve data }/ relevant data

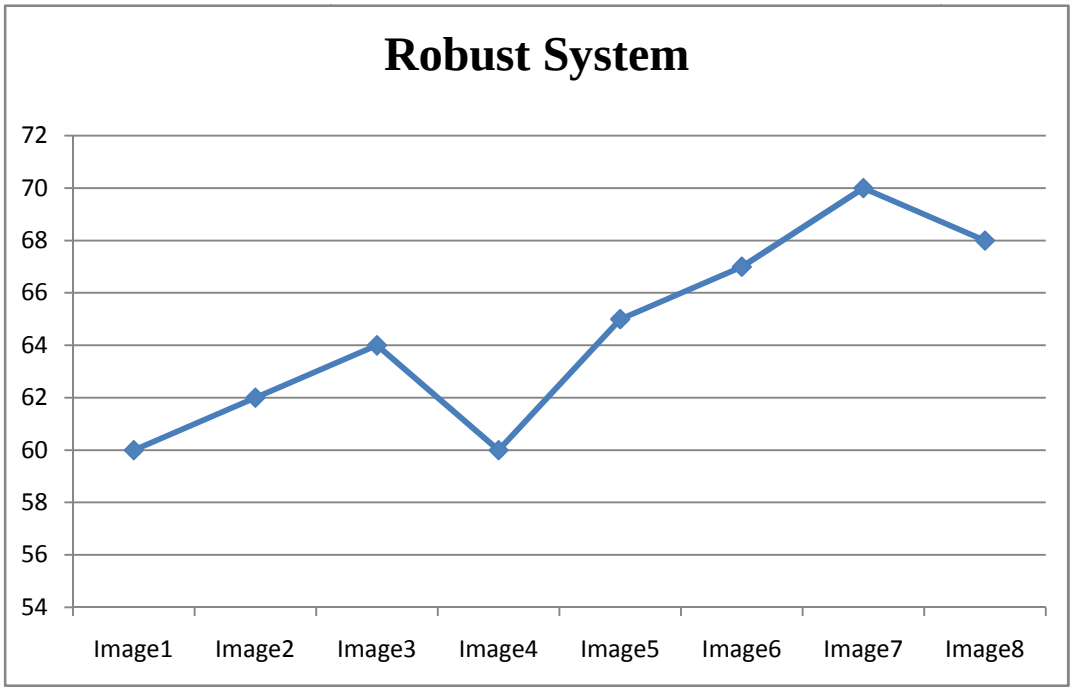

Fig. 2. Precision.

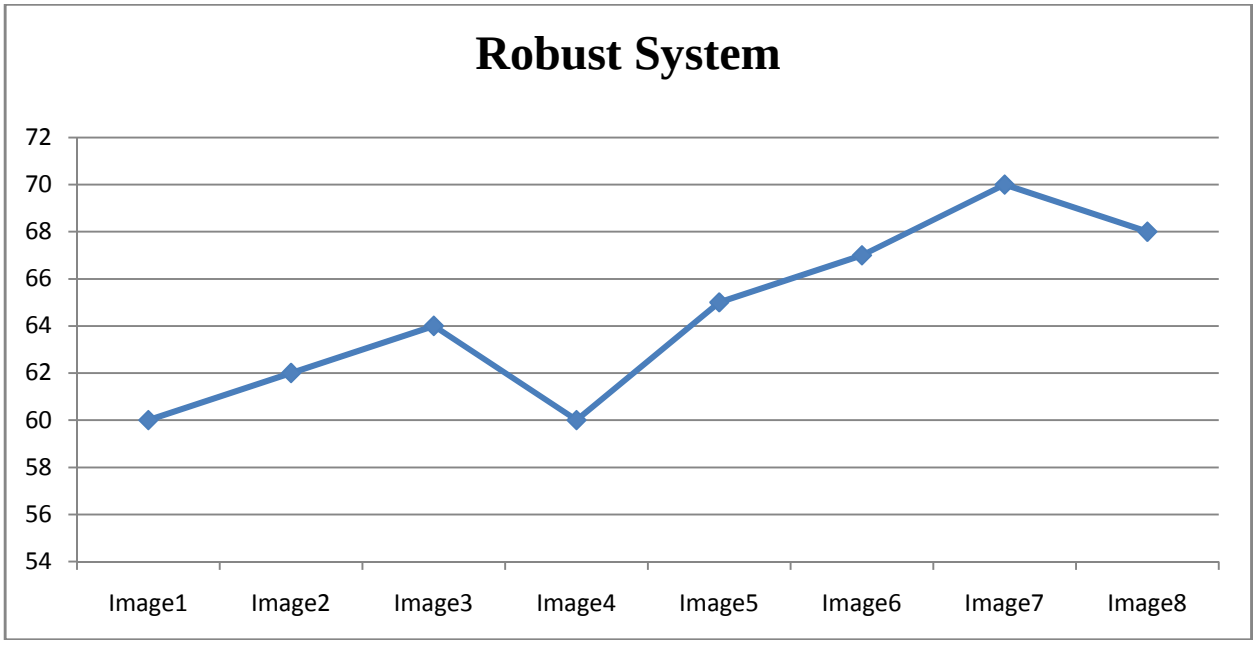

Fig. 3. Recall.

Table 4. Result.

Method	**Precision (%)**	**Recall (%)**	**Speed(s)**
Robust System	**69%**	**66%**	**4.42**

IV. CONCLUSION

A hybrid approach is used to localize scene texts by integrating region information into a robust CC-based method. But, it has not considered the text regions which are detected in automatic segmentation processes. Hence to deal with problem designed robust approach.

In robust approach main focus is on the localization and text detection part .Found that our method gives maximum precision rate, recall rate and performance metric. That is our system is better result. But processing speed of our method is slower. Hence there is need to improve the speed of detection and localization. This system can also be enhanced by adding the text detection part.

REFERENCES

[1]. Ping Sung Liao, Tse Sheng Chen and Pau Choo Chung," A Fast Algorithm for Multilevel Thresholding", *Journal of Information science and Engineering*, **17**, 713-727 (2001).

[2]. Yi-Feng Pan, Xinwen Hou, and Cheng-Lin Liu," A Hybrid Approach to Detect and Localize Texts in Natural Scene Images" *IEEE Transaction On Image Processing*, Vol. **20**, NO. 3, MARCH 2011

[3]. S.L. Feng, R. Manmatha, and A. McCallum, Exploring the use of conditional random field models and HMMs for historical handwritten document recognition, in *Proc. 2nd Int. Conf. Document Image Analysis for Libraries (DIAL06), Washington*, DC, 2006, pp. 3037.

[4]. X.L. Chen, J. Yang, J. Zhang, and A.Waibel, Automatic detection and recognition of signs from natural scenes, *IEEE Trans. Image Process*, vol. **13**, no. 1, pp. 8799, Jan. 2004.

[5]. J. Gllavata, R. Ewerth, and B. Freisleben, Text detection in images based on unsupervised classification of high-frequency wavelet coefficients, in *Proc. 17th Int. Conf. Pattern Recognition (ICPR04), Cambridge*, U.K., 2004, pp. 425428.

[6]. J. Lafferty, A. McCallum, and F. Pereira, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, in Proc. 18th Int. Conf. Machine Learning (ICML01), San Francisco, CA, 2001, pp. 282289.

[7]. H. P. Li, D. Doer-mann, and O. Kia, Automatic text detection and tracking in digital video, *IEEE Trans. Image Process.*, vol. **9**, pp. 147156, Jan. 2000.