IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 04, Issue 07 (July. 2014), ||V4|| PP 19-20

A Hybrid Approach To Detect And Recognize Text In Images

[1]Miss.Poonam B. Kadam, [2]Prof. Latika R. Desai
[1](Dept of Computer Engineering, [2](Dept of Computer Engineering

Abstract: **-** Text detection and recognition in images is a research area which attempts to develop a computer system with the ability to automatically read text from images or videos. This problem is challenging due to the complex background and large variations of these components features like color, size, shape, orientation or texture. Here new proposed framework is explored which can automatically detect and recognize aligned text from urban scene images e.g.shop name,landmark,street.The proposed framework evaluated on new generated dataset.It consist of a three main step 1) image partition perform to segment text based on color information.2) character grouping to detect text character in every text string depend on character size differences, distance between neighboring characters.3) the detected text recognition using neural network. This proposed method efficiently and accurately detects and recognizes the text with a low false positive.

Keywords: *- Text detection, preprocessing, segmentation, character recognition.*

I. INTRODUCTION

Many facilities are involved in human reading and similarly, there has been a large amount of research on computational methods for text recognition. The problem of developing a model for reading is broad, surrounding many facets of information. Also there are many typefaces, and they may be encountered at different legible sizes. Text is frequently printed on a surface and viewers generally doesn't want that the surface be fronto-parallel to their eyes in order to read it. Additionally, changing lighting conditions and designers use many colors for background and text. All of these factors leads to make the robust reading problem very challenging. Text understanding system consist of four stages: text detection, text localization, text extraction, text recognition. These stages are used interchangeably. Text detection consists of determination of the occurrence of text in images. Text localization is the process of determining the location of text. In text extraction stage the text components in images are segmented from background. After, the extracted text images can be converted into plain text using OCR technology. Through Text detection and recognition in images is coupling of text-based searching technologies and optical character recognition (OCR).Text in images or video is a powerful source of information. Caption embedded with news videos provide information about name of related location, people, date and time, purpose. Caption provide an summary or abstract of the image.

II. LITERATURE REVIEW

There have been a different methods dealing with text detection and recognition in images [4, 17]. Comprehensive surveys can be found in [8].Some approaches to text detection classified into three categories: texture-based methods, region based methods and hybrid methods. Texture-based methods [3,16] involves texture properties of text such as style, orientation and the construction of gray-level co-occurrence matrix. These methods are computation demanding as all locations and scales are exhaustively scanned. Moreover, these algorithms mostly concentrate to detect horizontal texts. Region based methods [7,18] use the properties of the color or gray scale or alignment in a text region or their differences in properties of the background. First extract candidate text regions through segmentation or clustering and then remove non-text regions.

The third category, hybrid methods [9]is a fusion of region-based and texture-based methods. Different document or web, e-mail images, in which text characters are normalized and proper resolutions, natural scene images, embed text can be in size, shapes and orientations into complex background. It is impossible to recognize text in images directly through OCR software because of complex background. Thus, we need to detect image regions containing text and their corresponding orientations. This is compatible with the detection procedure described text extraction algorithms survey[4,6].

Yi Method work in two stages for text detection first, the connected component in image partition based on gradient and color feature, then through text line grouping and adjacent character grouping [1]done which gives set of candidate or text line in image obtained. Then, Haar features are extracted from gradient maps and stroke by the block patterns presented in [2].

III. SYSTEM DESIGN

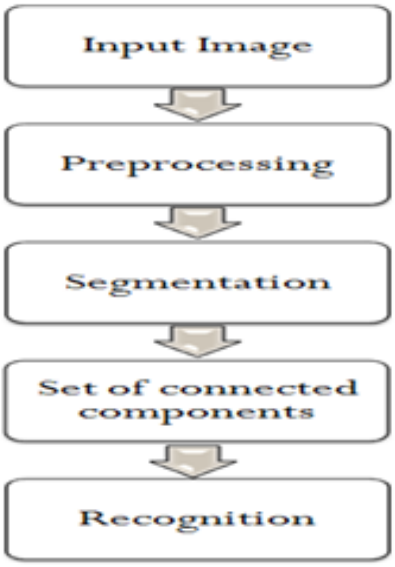

Fig 1. System Architecture

The proposed framework used to detect and recognize scheme efficient in urban scene context. This is an unified framework for text detection and recognition. The framework consist of three stages 1) image partition perform to segment text based on color information.2) character grouping to detect text characters in same text string depend on character size differences, distance between neighboring characters.3) the text recognition using neural network of detected text. After study of previous methods we found that combination color based partition and adjacent character grouping achieves better results in text detection. Therefore here by using effective color segmentation approach for text detection which help to get accurate result for recognition. Brief description of our proposed work.

3.1 Module I: Preprocessing

First, preprocessing and segmentation of input urban scene image, involves removing low frequency background noise, removing reflections, it is used for efficient extraction of text region information, a text region detector is designed to estimate the text confidence and the corresponding scale, based on which text candidate components are segmented and analyzed accurately. This stage significantly increase the reliability of an optical inspection. Here new hybrid color image segmentation approach is used. In most of images, color uniformity acts as stronger indicator to distinguish text characters of text string from background. Here k-means algorithm used.

K-Means Algorithm

The whole process of clustering is depicted in Figure 11. Given a set of observations (x1, x2,..., xn), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets ($k \leq n$) S = S1, S2,..., Sk so as to minimize the within-cluster sum of squares (WCSS):

$$argmins \sum_{i=1}^{k} \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (1)$$

where μi is the mean of points in Si.

The initial step is choosing k cluster centroid positions, sometimes called as seeds randomly from the image, the clusters are named as m1,m2,……,mk. Then, the algorithmruns iteratively, which mainly consists of two steps.

- Assignment step:

Assign each observation to the cluster whose mean yields the least within cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean distance, this is intuitively the nearest mean:

$$S_i^t = \{x_p : \|x_p - m_i^t\|^2 \leq \|x_p - m_j^t\|^2 \forall j, 1 \leq j \leq k\} \quad (2)$$

where each xp is assigned to exactly one St , even if it could be is assigned to two or more of them.

- Update step:

Calculate the new means as the centroids of the observations in the new clusters:

$$m_i^{t+1} = \frac{1}{|S_i^t|} \sum_{x_j \in S_i^t} x_j \quad (3)$$

In second step, calculate the new means as centroid of the clusters. The iteration done until assignments no longer change, which is a convergence to a local optimum. The k-means clustering is that we must specify the number of clusters in advance.

After the color quantization, the image is divided into different clusters, and we treat the connected components with the same cluster labels as a region. Each region is represented by its average color components. The output of the stage is the image partitioned into thousands of tiny regions quantized by a small number of colors.

3.2. Module II: Segmentation

The first step in a process of character recognition is a detection of a text area. After detecting the image text area segment it using horizontal projection. After text is segmented then characters are extracted from horizontal segments. Extracted characters are normalized by checking light related parameters like brightness etc. Then, recognized using by NN algorithm.

Edge Detection

Let us define input image as a rectangular area with horizontal and vertical edges. The high density of vertical and horizontal edges on a small area is in different cases caused by contrast characters, but not in every case. This process can cause sometimes detect a wrong area that does not correspond to actual text in image. Because of this, we detect several candidates using this algorithm, and then select the best one by a further heuristic analysis. The input image edge can detected vertically and horizontally.

Here segmentation using a horizontal projection is done. After cropping text area is deskewed, segment it by detecting spaces in its horizontal projection The adaptive thresholding is used to separate dark foreground from light background with non-uniform illumination before segmentation. After the thresholding horizontal projection perform f (x,y) to find horizontal boundaries between segmented characters. These boundaries correspond to peaks in the graph of the horizontal projection.

3.3. Module III: Character Recognition

After pre-processing and segmentation step feature set is extracted this is further used for training and recognition step. Feature extraction stage where features of the characters that are crucial for classifying them at recognition stage are extracted.

After this Neural network used as learning mechanism .In this comparison of input character pattern with stored character training set and find their matching probability. In the used method, various characters are taught to the network in a supervised manner. A labeled input character also have several variant patterns of the same character are taught to the network under the same label. Using it the network learns various possible variations of a single pattern and becomes adaptive in nature. This process also know as Character Matching which compare the segments against characters in a database, it check matching score of each database template to input character segment.

IV. RESULT

We evaluated proposed method on the most challenging street scene text dataset ,namely SVT (street view, landmarks, boards).The images in used dataset similar to traditional OCR setting which not contain low contrast or low resolution images. The motivation of this approach is to detect significant and accurate text in image and recognize it. The combination of color and texture based segmentation and after that using neural network for highly accurate character recognition. In recognition, the classifiers are evaluated with several metrics like accuracy refers to the percentage of the testing samples which were correctly recognized. Accuracy refers to the percentage of the characters of the testing samples which were correctly recognized. We have tested many random sample urban scene images which are collected as test set. We show performance of our method in GUI mention below.

4.1. Dataset and Evaluation

We used the street view text(SVT)images used for robust word recognition in our system. Our street view text dataset contain images taken from Google. We have tested 100 sample images. Most of the images are from urban signage. Our work focus on detecting words and recognizing respectively.

In our experiment color based segmentation and neural network are performed to partion and recognition of detected text. After testing 100 sample referred images from outdoor (street) indoor(mall and office).Our dataset the resulting precision, recall and f-measure are 0.83,0.93,0.25 respectively. Based on definition of precision and recall are: Precision (P) is the proportion of the predicted positive cases(tp) that were correct,as calculated using the equation:

$$precision=|tp| / |tp + fp|$$

The recall or true positive rate (tp) is the proportion of positive:

$$recall= |tp| / |tp + fn|$$

where false positive(fp) and false negative(fn).The f measures is combination of the precision and recall

$$f = 2.precision.recall / (precision + recall)$$

Sample Input Image from street scene text dataset.

Step-I Preprocessing and Segmentation

As shown in Figure we get cropped words separately after color segmentation. The Figure 2.shows different color segments in image depending that we get partitioned image for text detection.

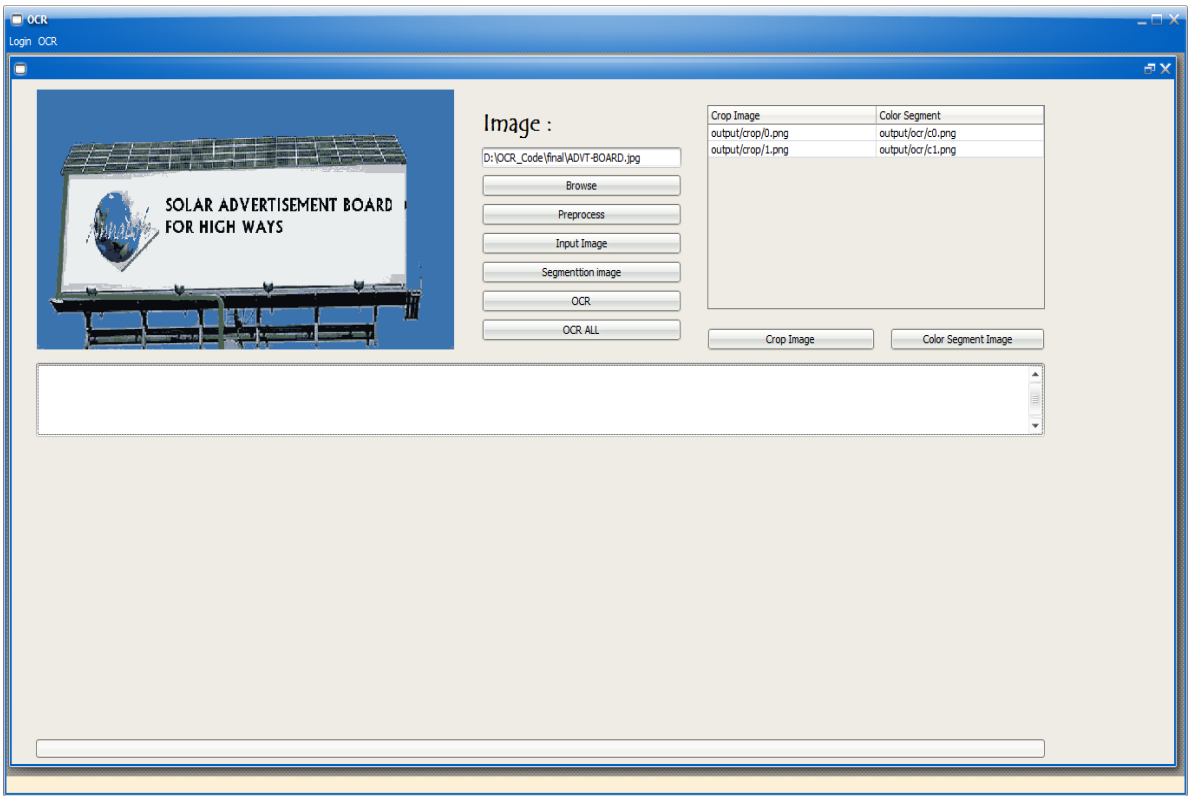

Fig 2.After Preprocessing and Segmentation

Step-II Recognition

All adjacent character groups i.e cropped words are recognized accurately as shown in Figure .Also we can separately recognize each word as shown in Figure3,4.

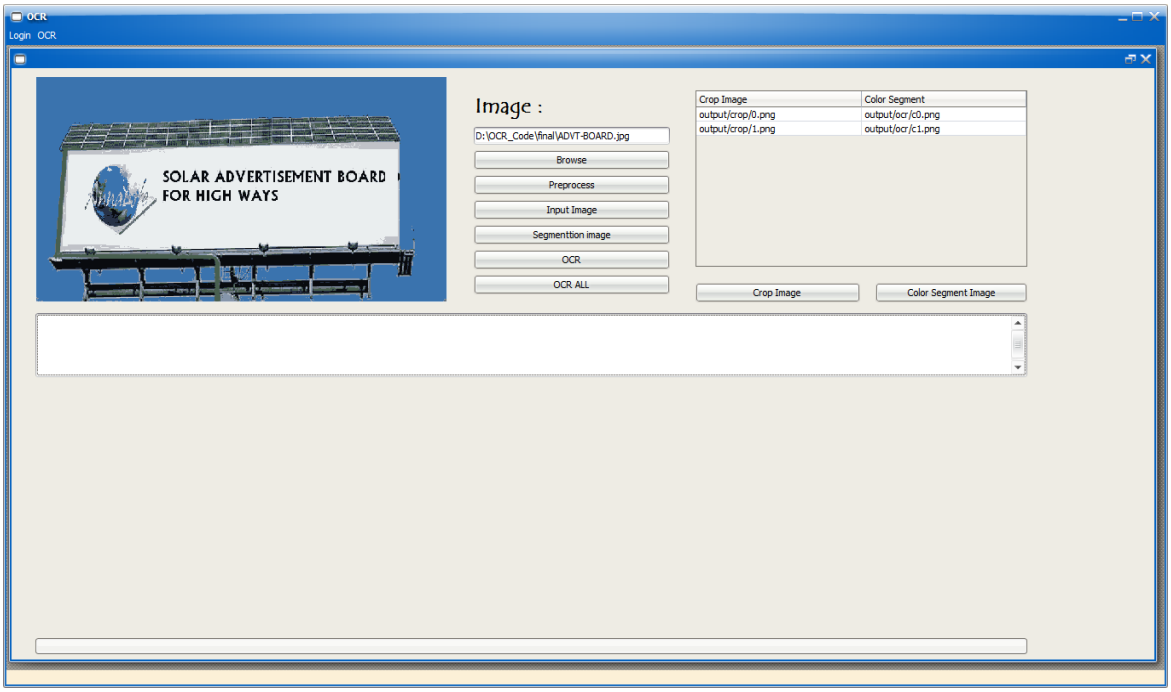

Fig 3. Character Recognition

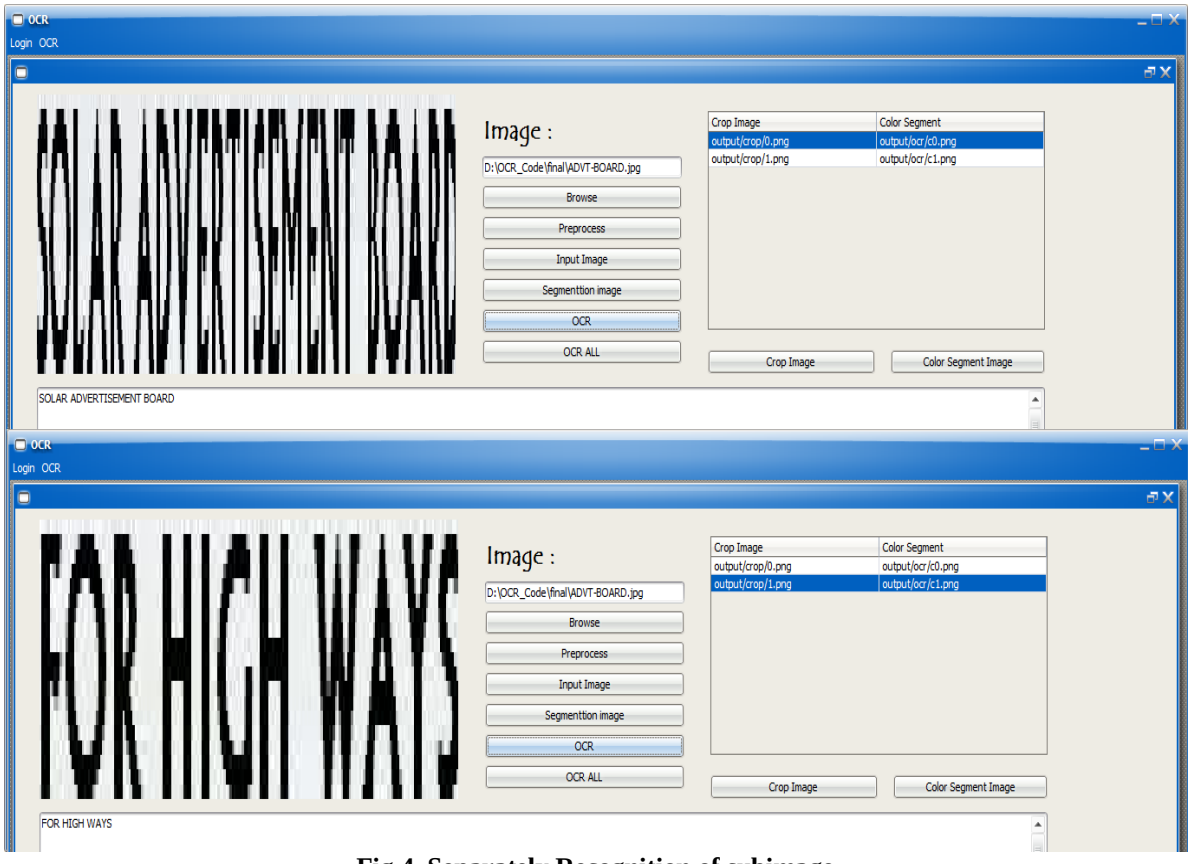

Fig 4. Separately Recognition of subimage

4.2. Result Table

Sr.No	Input Images	Character Recognition
1.		CENTURY
2.		GOOD EATS
3.		KOUTONS
4.		ESSEX
5.		HUAWEI
6.		ESTATE AGENTS SAXONS
7.		Save 4

Sr.No	Input Images	Character Recognition
8		HSBC
9		ALDO
10		KIDS EXCHANGE
11		BISON SHOP
12		PANDORA

Table 1. Sample Images with Result Table

V. CONCLUSION

This is unified approach on learning based methods for text detection from image having complex background and normalized text recognition. We have proposed the detection model eliminates the need for layout rules by learning spatial context parameters. Also, eliminate the peephole view of the sliding window approach and unify the entire process under a probabilistic model. Finally, robust reader needs rough text detection windows and completes both word and character segmentation in a recognition-driven process. To deal with the problems involved in training a computer to read from detection, to segmentation and recognition. To read text information embedded in those images, we propose a new framework which is more concentrating on how to give detected text having less false positive input to OCR give efficient and accurate recognized text. Here we choose texture color based image segmentation method for text detection and fast, accurate recognition perform using neural network framework.

VI. FUTURE ENHANCEMENT

There are different challenges occurs in text extraction from image because of size,style, orientation, shadow which makes recognition difficult. To improve the efficiency and transplant these algorithms into a system prepared for way finding of visually impaired people, keyword search engine application.

REFERENCES

[1] C. Yi and Y. Tian,Text string detection from natural scenes by structure-based partition and grouping, vol. 19, no. 12,2011.

[2] Chen and A. L. Yuille,Detecting and reading text in natural scenes in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2004, pp. 366.

[3] N. Zhong,B. Epshtein, E. Ofek, and Y. Wexler, Detecting text in nature scenes with stroke width transform, in Proc.IEEE Conf. Comput. Vis. Pattern Recognit., 2010, pp. 2963.

[4] K. C. Jung, K. I. Kim, and A. K. Jain,Text information extraction in images and video: A survey, Pattern Recognit., vol. 5, pp. 977.May 2004.

[5] Datong Chen, Jean-Marc Odobez,Herve Bourlard,Text detection and recognition in images and video frames Pattern Recognition, vol.37 ,2004,pp 595.

[6] Asif Shahab, Faisal Shafait, Andreas Dengel,ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images, Int. Jour.on Document Analysis and Recognition,2011.

[7] Yassin M. Y. Hasan and Lina J. Karam,Morphological Text Extraction from Images, IEEE Transactions on Image Processing, 9 (11) (2000) 1978-1983.
[8] Y. Pan, X. Hou, and C. Liu, A hybrid approach to detect and localize texts in natural scene images, IEEE Trans. on Image Processing, vol. 20, no. 3, pp.800, Mar. 2011.
[9] K. I. Kim, K. Jung, and J. H. Kim.Texture based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm. IEEE Trans. PAMI, 2003.
[10] L. Neumann and J. Matas.A method for text localization and recognition in real-world images. In Proc. of ACCV, 2010.
[11] Chucai Yi ,Yingli Tain Localizing text in scene images by boundary clustering,stroke segmentation,and string fragment classi_cation IEEE trans.on Image processing Vol 21 No.9 Sep 2012.
[12] Anand Mishra, Karteek Alahari,C.V.JawaharTop down and Bottom Up cues for text recognition IEEE conference of Image processing,2012.
[13] R.Minetto, N.Thome,M.cord,Text detection and Recognition in Urban Scene IEEE 2011.
[14] Ning-Yu An ,Chi-Man PunColor image segmentation using adaptive color quantization and multiresolution texture characterization. Springer-Verlag London Limited 25 May 2012.
[15] Fakhraddin Mamedov, Jamal Fathi Abu Hasna.Character Recognition Using Neural Networks, Near East University, North Cyprus, Turkey via Mersin-10, KKT C2005.
[16] R. Jiang, F. Qi, L. Xu, and G.Wu, A learning-based method to detect and segment text from scene images, J. Zhejiang Univ., vol. 8, pp.568-574, Apr. 2007.
[17] Pranob, K Charles, V.Harish,A Review on the Various Techniques used for Optical Character Recognition, Vol. 2, Issue 1,Jan-Feb 2012.
[18] Neural Network Approach To Mathematical Expression Recognition System, IJERT Dec 2012.
[19] Anil.K.Jain TextFeature extraction methods for character recognition-A Survey, Pattern Recognition, vol. 29, no. 4, pp. 641-662, 2006.