National Conference on Emerging Trends in Computer, Electrical & Electronics (ETCEE-2015)
International Journal of Advance Engineering and Research Development (IJAERD)
e-ISSN: 2348 - 4470 , print-ISSN:2348-6406,Impact Factor:3.134

Survey of Different Techniques for Text Detection from Natural Images

Priyanka Muchhadiya[#1], Prof. Toral P. Zinzuwadia[*2]

[#1]*Post Graduate Fellow, ECD, AITS, Rajkot, Gujarat, India.*
[*2]*Assistant Professor, ECD, AITS, Rajkot, Gujarat ,India*

***Abstract*— Texture analysis can provide very useful and vital information for content-based image analysis. Text recognition and analysis includes many applications such as: license plate recognition, sign detection as well translation, helping tourists and blind persons to understanding environment, drawing attention of a driver, content-based image search and so on. Locating text in case of variation in style, colour, as well as complex image background makes text reading from images more challenging. in this paper the various techniques available for detecting and recognizing text are explained, finally a hybrid approach using segmentation explained which can improve the qualitative texture analysis among other techniques.**

***Keywords*—Text detection, Text localization, Text recognition.**

I. INTRODUCTION

As we know that Text detection in natural scene image has attracted researcher's attention for many years, also with the rapid increasing use of digital image capturing devices, such as digital or analog cameras, mobile phones and PDAs, content-based and region based image analysis techniques are receiving more intensive attention in recent years. Among all the contents in images, textual data has inspired great interests, since it can be easily very understood by both humans and computer, and finds wide applications such as license plate recognition, sign detection and translation, mobile text recognition, helpful to blinder person content-based web image search etc.

In the last few years, many significant achievements have been made by researchers in the field of text localization[1]. However, most of the previous work present methods for text localization in image or video, but accurate text localization in natural scene image is still a challenge due to imaging conditions like viewing angle, blur, lighting, Resolution, Aliasing effects and considered scenes such as variations of text font, its size, colour and alignment orientation, complex background, illumination changes. On the other hand, most of the previous works, consider the texts in English But detection of Farsi or Arabic or Hindi, or even Chinese text in natural scene image has attracted a little. In this paper, we propose a method for localization different texts in natural scene images.

Texture analysis system consists of mainly four stages: text detection, text localization, text extraction, text recognition. We can use these stages text detection, localization, and extraction where Text detection consists of determination of the occurrence of text in image, Text localization is the process of determining the location of text, shown in figure 1.

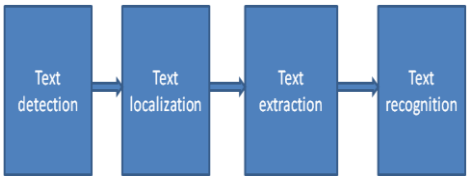

Figure 1 Basic block of texture analysis

Usually extraction of different texts can be done using segmentation. So 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. Now with the help of text extraction and detection from natural images, which coupling of text-based searching technologies and optical character recognition (OCR), is now recognized as a key component which is present in the images.

But as we know that, text characters [5] contained in images can be multicolour or gray-scale value with variable size, low resolution, and embedded in noisy backgrounds. Many experiments done on text recognition by applying conventional. OCR technology directly it leads to decrease rates of recognition. Therefore, for efficient detection and segmentation of text characters from the background is necessary to fill the gap between images and the text input of a OCR system[8][9]. Previous methods classified into top-down methods and bottom-up methods. In Top down approach algorithms first detect text in images and then segment each of them into text and background. In bottom-up approach after segment images into regions and then group character regions into words. The recognition performance therefore relies on the segmentation.

As I explained earlier for detecting text from natural images have many methods. Here we consider only region based component based method.

Now text detection and localization methods[13] can be categorized into two parts: region-based and connected component (CC)-based both are based on segmentation finally will describe the method for text extraction and recognition.

A region-based segmentation method for text detection consists of following stages:

1) Text detection to estimate text existing confidence in local image regions by classification,

2) Text localization to cluster local text regions into text blocks,

3) Text verification to remove non-text regions for further processing.

And a connected component (CC) based segmentation method consist of following stage

1) CC extraction to segment candidate text components from images;

2) CC analysis to filter out non-text components using heuristic rules or classifiers;

3) Post-processing to group text components into text blocks (e.g., words and lines).

These paper methods are categorized according to which technique will be used for detection and localization. That is CC based method and region based method. These methods are complementary to each other. If we merge this technique we get the robust output[14] called hybrid approach. Remaining section arrange as follow II Techniques/Method for text detection and localization III Edge Detection Operators IV Conclusion.

II. TEXT DETECTION TECHNIQUES

For scene texts localization[17][18][19], the process of our method consists of three stages: 1) pre-processing, designing a text region detector to generate the text confidence map, based on which text components can be segmented by local binarization, 2) CC analysis, presenting a CRF model[12] to formulate component analysis into component labelling problem, which is solved by minimum classification error (MCE) learning and graph cuts inference algorithm and 3) text line grouping, where component minimum spanning tree is built with a learned distance metric and inter-line edges are cut off with an energy minimization model. The flow chart of the proposed method is shown in Fig. 2

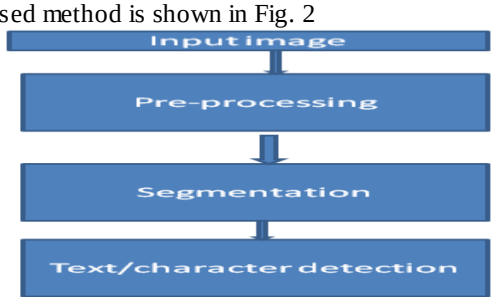

Figure 2: General procedure for text detection

Pre processing

A. Pre-processing

Here, the input image background is removed using different algorithms and image is gray scaled and then binarized (In black and white) and store in matrix of binary values. In pre-processing stage noise in image removed by using following steps.

Binarisation

- Load Input image can be JPG or BMP
- Retrieve and calculate the properties of image like width, height and n channels.
- Get the pointer to access image data
- For each height and for each width of the image, convert image to gray scale by calculating average of r,g,b channels of the image convert to gray scale manually.
- Convert Input RGB colour image to a gray-scale intensity image[20] Y as follows

$$\mathbf{Y = 0.299R + 0.587G + 0.114B} \quad (1)$$

- Convert this greyscale image Y in to binary image (bitmap) pixel values are obtained using the characteristic function as shown below.

$$B(x, y) = 1 \text{ if } g(x, y) < T$$
$$= 0 \text{ if } g(x, y) > T \quad (2)$$

T= Intensity mean (Threshold)

Remove Line: Horizontal and Vertical line removal Travers the image vertically and horizontally for black pixels according to column and row. Then if the total number of pixels compare to image height and width are less than 85% then turn all pixels into white.

Gap Removal: After Horizontal and Vertical lines are removed and then occurred discontinuity or gap will occur in image. To remove it by using 4– Connected Component algorithm. It will detect white pixel if that pixel have at least one black pixel then convert that white pixel to black. By using it we can remove all gaps which are occurred within image it helps in character recognition [6][7] of discontinuous character efficiently.

B. Segmentation

This stage try to segment the input image using different segmentation techniques, region based and connected component.

➢ **Region-based segmentation :**

Region-based segmentation methods[2] are based on observations that text regions have distinct characteristics from non-text regions such as distinctive gradient distribution, texture and structure. These methods generally consist of two stages: text detection and text localization. For text detection, features of local regions are extracted to determine if they contain texts. Then specific grouping or clustering approaches are employed to localize text regions accurately.

➢ **Connected component based segmentation:**

Connected Component (CC)-based methods[2][3] are based on observations that texts can be seen as sets of separate connected components, each of which has distinct intensity, colour distributions and enclosed contours. These methods generally contain three stages: 1) CC extraction to segment CCs from images, 2) CC analysis to determine whether or not they are text components by heuristic rules or classifiers and 3) group text components into text regions such as words, lines etc. Although some existing methods have reported promising results, there still remain several problems difficult to be solved. For CC-based methods, text components are hard to be segmented accurately without prior information of text position and scale. Furthermore, designing fast and reliable.

CC analysis method is also difficult since there are too many text-like components in images. On the other hand, the performance of region-based methods is sensitive to the text orientation and cluster number. Most of these methods can only localize texts containing many characters in horizontal alignment.

C. **Character or Text Recognition/detection**

It is used to recognize each alphabet (letter)[9][10] after the text in image has been segmented. The classifier is used to recognize each character. Character Matching Comparing the segments against characters in a database, it check matching score of each database template to input character segment.

III. EDGE DETECTION OPERATORS

A. Sobel edge detection

The Sobel operator[11] is more sensitive to diagonal edges than vertical and horizontal edges. The Sobel 3 x 3 templates are normally given as

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

X - direction Y - direction

B. Roberts edge detection

The Roberts operator[11] has a smaller effective area than the other mask, making it more susceptible to noise. The Roberts 3 x 3 templates are normally given as

$$H_r = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad H_c = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

C. Prewitt edge detection

The Prewitt operator[11] is more sensitive to vertical and horizontal edges than diagonal edges. The Prewitt 3 x 3 templates are normally given as

$$H_r = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad H_c = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

The Frei-Chen mask

$$H_r = \begin{bmatrix} 0 & 0 & -1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & 0 & -1 \end{bmatrix} \qquad H_c = \begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix}$$

In many applications, edge width is not a concern. In others, such as machine vision, it is a great concern. The gradient operators[16] discussed above produce a large response across an area where an edge is present. This is especially true for slowly ramping edges. Ideally, an edge detector should indicate any edges at the centre of an edge. This is referred to as localization. If an edge detector creates an image map with edges several pixels wide, it is difficult to locate the centres of the edges. It becomes necessary to employ a process called thinning to reduce the edge width to one pixel. Second order derivative edge detectors provide better edge localization.

IV. CONCLUSION

In this paper we revised the existing method for detection and localize the text. In region based method the text will be detected and localize by texture analysis. In CC-based method directly segment candidate text components by edge Detection.

REFERENCES

[1] Y. X. Liu, S. Goto, and T. Ikenaga, "A contour-based robust algorithm for text detection in color images", IEICE Trans. Inf. Syst., vol. E89-D, no. 3, pp. 1221 1230, 2006

[2] 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. 87–99, Jan. 2004

[3] M. Cai, J. Song, and M. R. Lyu, "A new approach for video text detection," in Proc. Int. Conf. Image Process.,Rochester, NY, Sep. 2002, pp.117–120C.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate en-ergy minimization via graph cuts.IEEE Trans. Pattern Anal.Mach. Intell., 23(11):1222–1239, 2001

[5] Liang J., Doermann D., and Li H., "Camera-based analysis of text and documents: A survey," Int. J.document Anal. Recogn., 2005; 7: 84–104..

[6] Wu V., Manmatha R., and Riseman E. M., "Text Finder: An Automatic System to Detect and Recognize Text in Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009; 21: 1224-1229.
[7] C. Yi and Y. Tian, "Text string detection from natural scenes by structure-based partition and grouping",IJFST vol.19, no. 12,2011.
[8] Elie Bursztein, Matthieu artin, Stanford University
[9] Text- based CAPTCHA Strengths and Weaknesses ACM(CSS 2011). Pranob, K Charles, V.Harish, "A Review on the Various Techniques used for Optical Character Recognition" IJJST Vol. 2, Issue 1,Jan-Feb 2012.
[12] Yi-Feng Pan, Yuanping Zhu, Jun Sun, Satoshi Naoi "Improving Scene Text Detection by Scale-adaptive Segmentation and Weighted CRF Verification".
[13] Xiaoqian Liu, Ke Lu, Weiqiang Wang "Effectively localize Text in Natural Scene Images" ICPR 1197-1200, Nov. 2012.
[14] Xu-Cheng Yin, *Member, IEEE,* Xuwang Yin, Kaizhu Huang, Hong-Wei Hao "Robust Text Detection in Natural Scene Images" Vol. 36, No. 5, 970-983, May 2014.
[15] Jae-Hyun Seok, Jin Hyung Kim "Scene Text Recognition with a Hough Forest Implicit Shape Model" 599-603, ICDAR 2013.
[16] B.H.Shekar, Smitha M.L., P. Shivakumara "Discrete Wavelet Transform and Gradient Difference based approach for text localization in videos" 280-284, ICSIP 2014.
[17] Mohanabharathi.R, Surender.K, Selvi.C "Detecting and Localizing Color Text in Natural Scene Images Using Region Based & Connected Component Method", IJMER, Vol.3, Issue.1, Jan-Feb. 2013 pp-331-335.
[18] Maryam Darab a, Mohammad Rahmati "A Hybrid Approach to Localize Farsi Text in Natural Scene Images", INNS-WC 2012, 171 – 184.
[19] Khushbu C. Saner, Mrs. Neeta. A. Deshpande " REVIEW: ROBUST APPROACH TO DETECT AND LOCALIZE TEXT FROM NATURAL SCENE IMAGES", IJSER June 2013, Volume 4, Issue 6, 2841-2851.
[20] Paraskevas Diamantatos, Ergina Kavallieratou, Pilar Gomez-Gil "Binarization: a Tool for Text Localization", IEEE 649-654, ICFHR.2014.
[10] Miss.Poonam B.Kadam, Mrs. Latika R. Desai "A Hybrid Approach to Detect and Recognize Texts in Images" IJRCCT vol 2 issue7,2013.
[11] Rafeal.C.Gonzalez, Richard E. Woods, "Digital image processing" Pearson Education, Second Edition.