Scientific Journal of Impact Factor(SJIF): 3.134

e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 1,Issue 11, November -2014

SURVEY ON TEXT DETECTION AND TEXT RECOGNITION METHODOLOGY FOR NATURAL IMAGES

Miss. Beena Dhamecha[1] ,Prof. Vaishali J. Kalariya[2]

Department of Computer Engineering, R.K.University, Rajkot, Gujrat, India[1]. Email Id:dhamechabeena16@gmail.com
Department of Computer Engineering, R.K.University, Rajkot, Gujrat, India[2]

Abstract –*Text in Natural Images contains a valuable information about scene, convey the information about what is actually depicted in the images. There are many applications like image searching and indexing, navigation, understanding of image, identify vehicles number plate and human computer interaction etc. Manually assignment of text from scene images is very time consuming andcostly. Hence automation of text extraction is challenging task because of text might be in different style, different font, illumination condition, poor quality, surface deformation , complex background etc. Different techniques have been proposed to solve this problem. This paper gives description of work done for automatically text detection and recognition in scene images having complex background.*

Keywords-*text detection; text recognition; complex background; localization; extraction*

I. INTRODUCTION

Text in images carry high-level semantic information of scene. Images are increasing on webs and in databases. It is a pressing task to develop effective methods to manage and retrieve these resources by their content. It is a difficult task to detect and segment text from scene/captured images due to main reasons like: different types of text patterns like size, font style, orientations, colors, background outlier similar to the text characters. After text detection and segmentation, text recognition system is applied to convert image into readable text, but it performs poorly when there is a text on the complex background. Text recognition is important for a lot of applications like automatic sign reading, navigation, language translation, license plate reading, content based image search etc. So it is necessary to understand scene text than ever.

With the rapid growth in digital technologies and gadgets which are made with megapixel cameras and other devices are like PDA, mobiles etc. are responsible for increasing the attention for information retrieval and it leads to a new research task.Text, in the images contain valuable information and provide a cues about images. So it is very important for a human as well as computer to understand the scenes.

II. LITERATURE REVIEW

2.1 Text Detection

Number of methods has been proposed for text detection in the past. Normally Text detection can be classified into following categories: Region-based method, Edge-based, texture-based,and connected-component based method. Whereas Text Recognition can be classified into two categories such as Traditional Optical character recognition (OCR) and Object recognition based method.

In region-based method, scan the images at multiple scales and uses the text/non-text classifier to find the potential text areas. Generally a feature vector extracted from local area is fed into a classifier. Because of text region have different properties from non-text ones, it can detect and localize text accurately even when are noisy. For region-based methods, the speed is very slow, and it is sensitive to text alignment orientation.

Coates et al. [1] proposed to learn features automatically from unlabeled data using unsupervised features learning and then train liner SVM to classify whether a sliding window is text or non-text ones.

In connected-component based method, it directly segments the candidate text components by edge detection, color clustering to get the CCs. The non-text components are then removed using heuristic rules or classifiers.In this number of segmented candidate components is small, so that computational cost is low and the located candidate text components

are directly used for text recognitions. CC-based methods cannot segment text component properly without prior knowledge of text position and scale.

Epshtein et al. [2] used the CCs in a SWT images to form text lines. Shivakumara et al. [3] proposed to extract CCs by performing K-means clustering in the fourier - Laplacian domain, and use text straightness and edge density to discard false positives. Chen et al. [4] proposed edge-enhanced MSER as basic text candidates and geometric filtering and SWT is used to discard the non-text ones.

Edge-based method focus on the 'high contrast between the text and the background' and edges og the text boundary are identified and merged. Liu et al. [5] extract statistical features from Sobel edge in four directions and use K-means classifier to classify text or non-text cluster. This method is robust for complex background .but still it fails to detect low contrast and small font size texts. It is also expensive. Wong et al. [6] compute maximum gradient difference to identify the line segments then extended to neighboring top and bottom rows to form candidate text regions. It has low false positive rate, but it uses many classifiers and sensitive to threshold values.

In texture-based methods, it considers text as a special textures. It apply Fast Fourier Transform, DCT, wavelet decomposition, and Gabor filter for feature extraction. Ye et al. [7] calculate the wavelet energy features at different scale and perform the thresholding to find out the candidate text pixels then it is merged into text lines.

2.2 Text Recognition

In Traditional Optical Character Recognition (OCR) based method, different binarization methods have been applied to get the binary images, that directly fed to the off-the-self OCR[8]. Text in scene images differs from the scanned document in terms of size, font, illumination condition, resolutions etc.The loss of information in binarization is not recoverable. So the result is poor and do not correctly recognize the text.

Other method Object Recognition based method, it directly extract features from original images and uses various different classifier to recognize the text [8]. In this method, do not do the binarization and segmentation but uses multi-scale sliding window technique to get the candidate character detection result. A special structure information is not used in the sliding window technique, so it will produce many false positives. Because of this, it is depended on post-processing methods like pictorial structure or CRF model.

In Otsu's method, it is based on histogram and used a global thresholding[9]. Using k-means clustering,text detection and binarization method is worked for Korean sign board. But in complex background and lighting, it is difficult for finding best value for k. Different methods have been suggested for text extraction.

Cai et al. [10] proposed a text detection method which is character features such as edge density, edge strength. First apply a coloe edge detection algorithm in YUV colorand filter out non-text edges by using a low threshold. Then, By applying a local thresholding techniques to keep low-contrast text and simplify the background. At last, to localize text regions projection profiles are analyzed.

Kim[11] proposed a method in which LCQ is performed for each color separately. Each color is assumed as a text color without knowing whether it is real text color or not.to reduce processing time, imag is coverted into a 256-color image before color quantization done. To find candidate text line, the candidate components which are extracted for each color are combined when show text region features. Disadvantage of this method is high processing time.

Jain and Yu[12] by bit dropping first perform a color reduction and color quantization, then after multi-valued images decomposition algorithm is applied to decompose the input image into multiple background and foreground images. Then to localize text candidate ,connected component analysis combined with projection profile features are performed on each of the images.it extract only horizontal text for large sizes.

Neumann and Matas[13] ,first detect characters as a MSERs and the apply text recognition using segmentation which is obtained by the MSER detector.

Wang et al. [14] to train the text detection and character recognition moduler, it used a CNN method, and built an end-to-end system with non-maximal suppression (NMS). And search with help of lexicon.Speed is relatively slow.

REFERENCES

[1] A. Coates, B. Carpenter, C. Case, S.Satheesh, B.Suresh, T. Wang, D.Wu, A.Ng, Text detection and characterrecognition in scene images with unsupervised feature learning, in International Conference on Document Analysis and recognition (ICDAR), IEEE, 2011, pp.440-445.

[2] B. Epshtein, E. Ofek, Y. Wexler, Detecting text in natural scenes with stroke width transform, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE,2010,pp,2963-2970.

[3] P.Shivakumara, T.Phan, C.Tan, A Laplacian approach to multi-oriented text detection in video, IEEE Trans. Pattern Anal. Mach. Intel. 33 (2) (2011) 412-419.

[4] H.Chen, S.S. Tsai, G.Schroth, D.M. Chen, R.Grzeszczuk, B.Girod, Robust text detection in natural images with edge-enhanced maximally stable extremal regions, in: 18th IEEE International Conference on Image Processing (ICIP), IEEE, 2011, pp. 2609-2612.

[5] C. Liu, C. Wang, and R. Dai, "Text Detection in images Based on Unsupervised Classification of Edge-Based features", Proc. Eighth Int'l Conf. Document Analysis and Recognition, pp.117-120,2002.

[6] E.K. Wong and M. Chen, "A New Robust Algorithm for Video Text Extraction," Pattern Recognition, vol. 36, pp. 1397-1406, 2003.

[7]Q. Ye, Q. Hung, W. Gao, and D.Zhao, "Fast and robust Text Detection in Images and Video frames," Image and vision Computing, vol.23, pp. 565-576, 2005.

[8] Cunzhao shi,Chunheng Wang, Baihua Xiao, Yang Zhang, Song Gao,Zhong Zhang "Scene text Recognition using Part based tree structured Character Detection,"pattern recognition in 2013 Elsevier Ltd.

[9] N. Otsu, A threshold selection method from gray-level histograms. Automatica, 11:285-296, 1975.

[10] M. Cai, J. Song and M.R. Lyu. A new approach for Video Text detection. In Proc. Of International Conference On Image Processing, Rochester, New York, USA, pp. 117-120, 2002.

[11] P.-K. Kim." Automatic Text Location in Complex Color Images Using Local Color Quantization.IEEE TENCON, vol. 1, pp. 629-632, 1999.

[12] A.K. Jain and B.Yu. Automatic Text Location in Images and Video Frames. In Proc. Of International Conference of Pattern Recognition (ICPR), Brisbane, pp. 1497-1499,1998.

[13] L. Neumann, J.Matas, A method for text localization and recognition in real world images, in: ACCV, Springer, Queenstown, New Zealand, 2010, pp. 770-783.

[14] T.Wang, D.J. Wu, A. Coates, A.Y. Ng, "End to end text recognition with convolutional neural networks," in: 21st International Conference on Pattern Recognition (ICPR), IEEE, 2012, pp. 3304-3308.

