Text Extraction from Image Using MSER Approach

V.Kalai selvan
M.Tech (Multimedia Technology)[1]
[1]Department of Computer Science and Engineering
K.S.R College of Engineering, Tamilnadu, India
[1]kalaiselvanflower@gmail.com

M.Prakash
M.Tech, (Ph.D)[2]
[2]Asst.Prof Department of Computer Science and Engineering
K.S.R College of Engineering, Tamilnadu, India
[2] mmsprakash@gmail.com

Abstract—The automated understanding of textual information in images is an important problem to solve for the Computer Vision and Document Analysis for extracting that information for processing. This needs to generate required word regions and the remaining to be filter out the nontext area. For this, we extract the connected components (CCs) in images by using the maximally stable extremal region algorithm. Whereas in the existing system the region based method is considered. These extracted CCs are partitioned into clusters so that we can generate candidate regions Instead of using heuristic rules for clustering we train an AdaBoost classifier which determines the adjacency relationship and cluster those CCs by using their pair wise relations. Then we normalize candidate word regions and determine whether each region contains text or not. Adaboost classifier is based on multilayer perceptrons and we can control recall and precision rates with a single free parameter we develop text/nontext classifier for normalized images. Finally we obtain the extracted text by matching the trained set of templates.

Key words: Connected Components, Maximally Stable Extremal Region, binarization, Geometric normalization.

I.INTRODUCTION

Learning to read is a long road for humans the majority of the reading activity is carried out over written words on paper, the texts may also appear written on objects around us: text on street signs, a motto painted on a wall, and even text produced by arranging small stones on the sand. Environment text is an important source of information in our daily life. Due to increase in use of portable devices, equipped with integrated built-in digital cameras Reading text in natural scenes leads to new applications such as automatic translation, license plate readings, and visually impaired persons, which shows that the field has gained increasing attention of the researchers in the last decades, which also brought many methods as solution there are lot of works done for this problem by surveying these methods we can conclude with two common method for cc extraction the Region based method which is also known as sliding window approach which involves lot of comparison it classify the given region contains text or not. The other one is the connected components extraction which computes the calculation on selection area only. Here it focuses the process to extract text like cc , to filter out non text regions, to infer text blocks from CCs.

A. Related work

The problem of extracting the text from image is commonly said to be text information extraction (TIE). Which involves the steps of (i) text detection, (ii) text localization, (iii) text tracking, (iv) text extraction and enhancement, and (v) recognition using(OCR) . Smith and Kanade [5] defined a scene-change based on the difference between two consecutive frames and then used this scene-change information for text detection. They achieved an accuracy of 90% in scene-change detection. Lim et al. [6] made a simple assumption that text usually has a higher intensity than the background. They counted the number of pixels that are lighter than a predefined threshold value and exhibited a significant color difference relative to their neighborhood, and regarded a frame with a large number of such pixels as a text frame. This method is extremely simple and fast. However, problems can occur with color-reversed text. Zhong et al. [8] and Antani et al. [9] performed text localization on compressed images, which resulted in a faster performance. Therefore, their text localizers could also be used for text detection. The text detection stage is closely related to the text localization and text tracking stages Lee and Kankanhalli [7] applied a CC-based method to the detection and recognition of text on cargo containers, which can have uneven lighting conditions and characters with different sizes and shapes. Edge information is used for a coarse search prior to the CC generation. The difference between adjacent pixels is

used to determine the boundaries of potential characters after quantizing an input image. Local threshold values are then selected for each text candidate, based on the pixels on the boundaries. These potential characters are used to generate CCs with the same gray-level. Thereafter, several heuristics are used to filter out non-text components based on the aspect ratio, histogram of contrast, measurement. Despite their claims that the method could be effectively used in other domains, experimental results were only presented for cargo container images. Ohya et al.[10] Cluster-based templates are used along with geometrical information, such as size, area, and alignment. They are constructed using a K-means clustering algorithm from actual text images for color image.

B. Proposed work

Our approach consist of three stages which is shown in the figure 1 CC-based approaches is used for generating the candidate region here we have focused the problem of region based method this will leads to simple text/nontext classification the problems to connect text blocks from CCs and to determine whether a given patch is a part of a text region will be addressed based on machine learning techniques. The Adaboost classifier determines adjacency relationship between CCs for the grouping of the text components. The CC extraction algorithm is the maximally stable extremal region (MSER) is able to detect almost all characters even when the image is in low quality. The MSER algorithm detects a large number of non-characters, most of the character candidates need to be removed before further processing. It is invariant to scale affine to the intensity change.

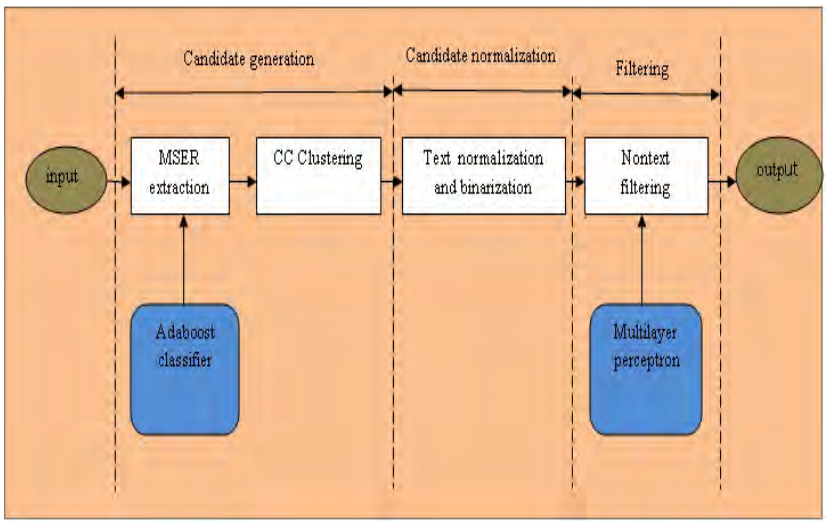

Figure 1 Architecture Diagram

C. Organization of the paper

The Section II deals about our candidate generation method. In Section III, we present our normalization method. The normalization in our paper is both geometric and binarization . Section IV shows about Text/Non text classification. Finally, conclude the paper in Section V.

II. CANDIDATE GENERATION

For candidate generation, the connected components are extracted in images which are then partitioned into clusters, since the Adaboost classifier is based on an adjacency relationship

CONNECTED COMPONENT EXTRACTION

We use the MSER algorithm for extracting the CCs .since it is stable over a range of thresholds it can be used for local binarization. It can find the candidate components that are either darker or brighter than their surroundings intensity values as shown in the below figure a and b

Figure 2(a) Figure 2(b)

By using the AdaBoost classifier that tells us whether the given set of connected components is adjacent or not. And the result of the classifier the cc clustering is done which depends on the threshold values of the CCs adjacency.

III. CANDIDATE NORMALIZATION

After CC clustering, we have a set of clusters. From that we will normalize corresponding regions for the reliable text/nontext classification it is of two steps as follows

A. Geometric Normalization

For the given set of clusters we first localize its corresponding region. We approximate the shape of text boxes whose left and right sides are parallel to y-axis. This approximation reduce the difficulties in estimating text boxes for this we only have to find a skew and four boundary supporting points.

B. Binarization

By using the geometrically normalized images binary images can be built. we perform the binarization separately by estimating text and background colors. The MSER results also can be used as binarization result however it may miss some character components and/or yield noisy regions due to the blur and also we have to store the point information of all CCs for the MSER-based binarization. We consider the average color of CCs as the text color and the average color of an entire block as the background color.

IV. TEXT/NONTEXT CLASSIFICATION

To get final results, we develop a text/nontext filter which rejects nontext blocks among normalized images. Since we have small number of regions to be filtered simple classifier is enough on the other hand, the variable aspect ratio for this we need to split the normalized images into patches of letters and develop a character/non-character filter.

Finally these separated characters can be compared with the trained data blocks and the matched letter can be used as text thus the each letter blocks can be concatenated to give as word and line.

V. CONCLUSION

In this paper, we have presented a novel text detection algorithm based on machine learning techniques. We have developed two classifiers: one is designed to generate candidates and the other is for the filtering of nontext candidates. And for grouping the text we have introduced a filter for filtering the nontext part by separating the each character and finally each letter blocks can be concatenated to give as word and line.

REFERENCES

[1] Y.-F. Pan, X. W. Hou, and C.-L. Liu, “Text localization in natural scene images based on conditional random field,” in Proc. 10th Int. Conf. Document Analysis and Recognition (ICDAR’09), Barcelona, Spain,2009, pp. 6–10.

[2] G. Schroth, S. TsaiD. Chen, H. Chen ,R. Grzeszczuk, and B. Girod,“Robust text detection in natural images with edge-enhanced maximally stable extremal regions,” in Proc. IEEE Int. Conference. Image Processing,(2011), pp. 2609–2612.

[3] O. Chum, J. Matas, T. Pajdla and U. Martin,, “Robust wide baseline stereo from maximally stable extremal regions,” in Proc. Brit. Mach.Vision. Conference., (2002), pp. 384–393.

[4] E. Ofek ,B. Epshtein, , and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in Proc. IEEE Conf. Computer Vision. Pattern Recognition., Jun. (2010), pp. 2963–2970.

[5] T. Kanade , and M.A. Smith, Video Skimming for Quick Browsing Based on Audio and Image Characterization, Technical Report CMU-CS-95-186, Carnegie Mellon University, (1995).

[6] S.W. Lee , S. H. Choi, and Y. K. Lim, Text Extraction in MPEG Compressed Video for Content-based Indexing, Proceeding . of International Conference on Pattern Recognition, (2000) pp. 409-412.

[7] Kankanhalli, and C.M. Lee Automatic Extraction of Characters in Complex Images, International Journal of Pattern Recognition Artificial Intelligence, (1995), pp 67-82.

[8] Anil K. Jain, Yu Zhong, Hongjiang Zhang, and Automatic Caption Localization in Compressed Video, IEEE Transactions on Pattern Analysis and Machine Intelligence, (2000) ,pp385-392.

[9] D. Crandall, U. Gargi, S. Antani T. Gandhi, and R. Kasturi, Extraction of Text in Video, Technical Report of Department of Computer Science and Engineering, Penn. State University, (1999).

[10] S. Akamatsu , A. Shio J. Ohya, , and, Recognizing Characters in Scene Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, (1994) ,pp214-224.