An Efficient Text Detection in Natural Scene Images Using MSER Based Approach

Dr.R.Mariappan

Professor, Department of ECE, S V College of Engineering, Tirupati, A.P, India

ABSTRACT: In this paper, the performance of text detection and recognition in colour images is analyzed and compared to solve the technical challenges. Maximally stable Extremal Regions (MSERs) are extracted as character candidates by a fast and effective pruning algorithm. Text candidates are formed by grouping the character candidates using single link clustering algorithm. Here clustering threshold and distance weights are known automatically by self training distance metric learning algorithm. Character classifier is used to estimatethe probability of text candidates corresponding to non text ones. Text classifier is to identify texts by eliminating the text candidates having high non text probabilities. The simulation result reveals the better performance of the proposed work.

KEYWORDS: scene text detection, Maximally Stable Extremal Regions, link clustering, the distance learning algorithm.

I.INTRODUCTION

Multimedia technology has increased tremendously in recent years. In multimedia technology image is one of the major part and the image can have different contents in it such as human, scenes, text etc. All the contents of the images are important, but text is found to be one of the most important features to understand the image contents. Text in images can be used for the purpose of indexing. The text information can be extracted by two stages: text detection and text recognition. External region is the region that is extracted from an image using the text detection and text recognition stage system retrieves the text information from the external regions [1]. Retrieving the content from images is challenging due to the image quality and background noise. There are different kinds of images that have text as its part with background, such as document images, scene images, and born digital images. Digital cameras and camera phones enables the acquisition of images and video data's containing scene text like street signs, but these devices also introduce new imaging conditions such as noise, angle of viewing, and lower resolution, etc. Considering above scene text properties and problems, natural scene text detection and recognition is more difficult task in comparison With text in born digital and document image.

Human beings are very good at deriving the information from the images, because of our instinctive visual and mental abilities. About most of the information received by humans is in the pictorial form. The analyses of pictures that employ perspective, including the radiation invisible to human eye are considered. Textual part images or a video frame of a scene does not have a prior knowledge of environment, objects supporting text, lighting, acquisition parameters and text. It can be easily viewed as text in real
world conditions without any constrains. Text extraction in natural scene images is challenging because of the different conditions under which the image is captured. Natural scene text detection aiming at extracts the text from daily images is the main concern of this text. There are many information sources for text information extraction in images like colour, texture.

II.EXISTING WORKS

The Maximum Stable Extremal Regions methods are giving the better performance in real time projects. But current MSER based methods are still having some drawbacks. For example, they will suffer from detecting of insufficient text candidates construction and repeating components algorithms. In this paper we will review the MSER based methods focusing on these two problems. The

other scene text detection methods can be referred to some survey papers [2]. The algorithm to detect the video text for low and high contrast images [3]. The low and high contrasts are classified by calculating the edge difference between sobel and canny edge detectors. After the process of calculating edge and texture features high contrast and low contrast thresholds are used to extract text objects from low and high contrast images separately. The inner distance based shape filter and an intensity histogram based filters are used to eliminate the false positives and extract the text blocks whose intensities are similar to those of their components coming

from the same object and the adjoining areas of the object [4]. Ujjwal bhattacharya and bityut b Propose a work of automatic detection of character components from video frames of class room lectures [8]. There exists a few works on detecting the text from the white boards from video frames. The work of text localization is does not exist on black, blue, green and etc. And then the features along with perception network had not reported earlier.

The video frames used in the present study had been collected from video lectures available online and thus the results of this experimentation are reflecting the robustness of the proposed method. But this method will work properly detect the text on white board. Hyung Koo, and Duck Hoon Kim [9] presents a novel scene text detection algorithm on the machine learning technique. Here this technique is uses two classifiers. One classifier is used to generate the character candidates. And another classifier is used to filter the non text candidates. So the symbols and numerals are not identified by this method.

Carlos Merino Gracia, Majid Mirmehdi [10] presents a new reading system based on tracking the non text regions. It works on a real time, automatically detects and recognises the new text regions and removes the old ones. Here the technique is only focussed on larger text, and is not suited to deal with smaller texts. Observing the patterns of movement and context in the surrounding is resolvable one. Abdullah Amirknani Sharharki, and Amir Ebrahimi Ghahnavieh, [11] presents a new segmentation method for segmenting the lines in hand written Persian texts. The results are based on eliminating the additional signs between lines and morphological operations. So it will indicate the proper performance.

III.ROBUST SCENE TEXT ETECTION

The traditional MSER based methods can be incorporated by several key improvements. So we propose a novel MSER based scene text detection method. The structure of the proposed system, as well as the sample result of each stage is presented in Figure 1.

Before the process of extracting the character candidates we need to do pre processing. The image pre processing is also called image restoration. Pre processing involves the correction of distortion, degradation and noise introduced, during the image processing. Pre processing enhances the input image. In order to achieve high recognition rate, prior to character recognition, it is essential to eliminate the noise and imperfections introduced in the image. There are various techniques available for pre processing. The important technique is binarization. In binarization, gray scale images are converted into binary image with the help of thresholding. And reducing or increasing the number of pixels in the image can improve the visual ability of the image.
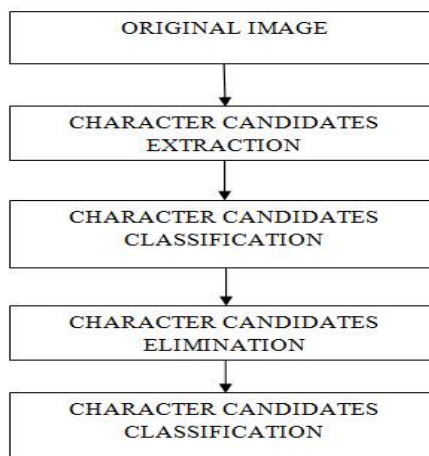
Fig 1: Flow chart of the proposed system.

The Flow chart shows the flow chare of the proposed system, and result of each step with the example. Text candidates are labelled by blue bounding rectangles; character candidates identified as characters are coloured green, and others red.

A. CHARACTER CANDIDATE'S EXTRACTION

Using the MSER algorithm the character candidates are eliminated and the repeating components are removed by the proposed MSER pruning algorithm by minimizing regularized variation.

B. TEXT CANDIDATE'S CONSTRUCTION

Here the self training distance metric learning algorithm is used to identify the distance and weighs of the text. By using proposed metric learning algorithm clustering threshold are learned simultaneously. By the single link clustering algorithm character candidates are grouped into text candidates.

C. TEXT CANDIDATE'S ELIMINATION

The text candidates corresponding to non texts are eliminated by using the character classifier probabilities are removed with text candidates.

D. TEXT CANDIDATE'S CLASSIFICATION

Text classifier identifies the text candidates corresponding to true texts. To test whether the given text candidates are corresponding to true text or not by using AdaBoost classifier.

IV. CHARACTER CANDIDATES EXTRACTION

A. PRUNING ALGORITHM

The problem of over fitting noisy data is avoided by common frame work called pruning. The basic idea is to incorporate simple theories to avoid complex rules with low coverage that contain irrelevant literals that have only been added to exclude noisy examples. The problem of false classification for the image features owing to the methodologies adopted in the classification procedure suffered by most of the image processing techniques. Enhancement and restoration of images through filtering techniques and morphological approaches [5], report miss classification occurs due to the misrepresentation of the information content. This problem is more common in severe noisy environments.

The most common fact is that the fact probability of erroneous classification of an image feature into a noise feature and vice versa, gets increases with the degree of noise levels.

The MSER algorithm applied over a text segmentation process is affected by the repeating components. The pruning algorithm is designed by using the MSER structures. In real world if the characters contain or cannot be contained by other characters, is used to remove the children own to be a character, and vice versa. It is very safe and prevents all the characters after elimination when the MSER tree is purned by applying this type of parent children elimination again and again.

Fig 2: Character corresponds in MSER trees.

(a) A MSER tree whose child corresponds to characters.
(b) A MSER tree whose parent corresponds to characters.

To eliminate the non text MSER candidates the pair wise relations may not enough and some amount of the complicated higher order properties of text is exploitable by pruning. The regularized variation scheme is used to estimate the accuracy. The MSER tree having different type of situations one child and multiple children. There will be two types of algorithms based on the operation of parent children elimination. They are given by tree accumulation algorithm and linear reduction algorithm. The line segments are removed by linear reduction algorithm. Further repeated characters are removed by the tree accumulation algorithm.

B. LINEAR REDUCTION

If the MSER having only one child, the linear reduction algorithm is used. This algorithm chooses minimum variation and it discards the other regions from parent and child. This process is applied over the whole tree. The MSER tree will returns the root of the proposed tree whose linear segments are reduced.

C. EXTRACTION USING CONNECTED COMPONENT METHOD

The connected components are extracted using the MSER algorithm. Since it is stable over a range of thresholds. This is used for binarization. If the character candidates are either darker or brighter, then surroundings the MSER algorithm will find the candidate components. These results are shown in figure 4. Here the classifier is used to determine whether the connected components are adjacent or not, and the result of the classifier the connected component clustering is done. This depends on threshold values of the connectedcomponent adjacency.

Fig 4: Surroundings intensity values.

V. TEXT CANDIDATES CONSTRUCTION

A. TEXT CANDIDATE'S CONSTRUCTION ALGORITHM

The text candidates are constructed by clustering character candidates using single link clustering algorithm. The clusters produced by the single link clustering algorithm are extended, and it is particularly suitable for the text candidate's construction task. It is belongs to hierarchical clustering. The data is considered as singleton cluster and it merged until it reaches single remaining cluster. In this two closest members are merged which have the smallest distance. The single link clustering algorithm is suitable algorithm for the task construction of character candidates. The single link clustering algorithm for merging smallest distance is processed in [7]. Whenever the distance between the neighbourhood clusters exceeds the given threshold value, the clustering process is terminated. The resulting clusters of cluster forest or single link clustering algorithm from a hierarchical cluster tree if the termination threshold is specified. In this single link clustering algorithm, every data points represent the top level clusters and a character candidate in the final forest tree corresponding to the next candidates. The main issue of this course is to determining the distance function and threshold for the single link clustering algorithm.

B. DISTANCE METRIC LEARNING ALGORITHM

Most of the clustering algorithms are confide on proper distance metric on the input space. One of the task of supervised clustering is learning a distance metric and it satisfies the constrains or labels in the supervised data given in the clustering algorithm. The clustering algorithm is based on good distance metric. The distance metric is to be learned it satisfying the constrains and labels in the processed data by clustering algorithm. The metric learning is to learn the distance function pairs in by minimizing the distance between points and maximizing the distance between point's pairs in M, where C specifies pairs of points in different cluster and M
in same cluster. By the single link clustering algorithm

$$D(u,v;w) > \text{ for all} (u,v) \in C$$
$$D(u,v;w) \leq \text{ for all} (u,v) \in M$$

From these equations, it is clear that C and M can be corresponds as the negative and positive sample set of a classification problem such that, the threshold and feature weights can be learned by minimizing the classification error. Small cluster forms a cluster and final cluster forms the binary cluster tree. In the single link clustering algorithm the clusters are formed by merging smaller clusters and the final cluster will forms a binary cluster tree, in this no singleton clusters have rightly two direct sub clusters. This type of single link clustering algorithm is helps to design a new learning algorithm. This is also used to learn the threshold values simultaneously.

VI. TEXT CANDIDATES CLASSIFICATION

For final results, we would develop a character or non character filter. That filter will rejects non character blocks among normalized images. Since we have only small number of regions to be filtered the classifier is simple enough. The aspect ratio for these normalized images which we need to split them into patches of letters and develops a character or non character filter. Finally the separated characters can be compared with trained data blocks. The matched letter can be used as text and each letter blocks can be concatenated to give as word and line. By the next candidates construction algorithm our experiments proves only smallest percentage of the text candidates corresponded to true texts. It is hard to train an effective text classifier using an un balanced database, as most of the non character candidates need to be removed before training the classifier. The character classifier is used to remove the posterior probabilities of the text candidates corresponding to the non character candidates and remove character candidates either high non character probabilities.

VII. RESULTS AND DISCUSSIONS

In this section we compare our technique with several methods on variety of algorithms. Here we present our algorithm flow of MSER pruning, distance metric learning algorithm and character candidate's elimination process of the proposed method. The proposed method will indicates the character regions by using bounding boxes. The results are shown in Figure5.
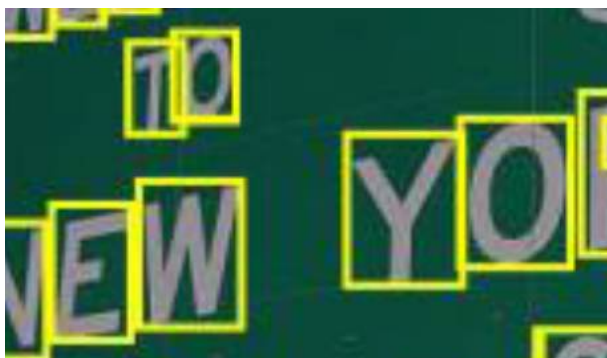
Fig 5: Result and discussions of proposed Methodology

VIII. CONCLUSION

The text detection in natural scene images is challenging for complex background. There are many methods available for detecting the text and recognition from natural scene images. Here we present a new MSER based technique. In this paper the pruning algorithm will used to detect the character candidates if the image is even in a low quality. Then distance metric learning algorithm that used to calculate distance weights and threshold values. The text candidates are constructed into character candidates by using single link clustering algorithm. Finally the character classifier is used to eliminate the non text candidates. By integrating the above
techniques we can build a text detection and recognition method.

REFERENCES

[1] Xiaobing Wang, Yonghang Song, Yuanlin Zang (2013) "Natural Scene Text Detection in Multi Channel Connected Component Segmentation", 12th International conference on Document Analysis and Recognition, Vol 117, No 2, pp182-194.
[2] K. Jung, K. Kim, and A. Jain, (2004) "Text Information Extraction in Images and Video: A Survey" Intrenational Conference on Pattern Recognition, Vol 37, No. 5, pp 977-997.
[3] Shivakumara, p., Huang, W., Phan, T.Tan, C.L. (2010) "Accurate video text detection through classification of low and high contrast images", Pattern Recognition Vol. 12, No. 3, pp 738- 750.
[4] Liu, Z., Sarkar, S, (2008) "Robust Outdoor Text Detection Using the Intensity and Shape Features", International Journal of Pattern Recognition, Vol. 14, No. 5, pp 675-693.
[5] X. Yin, X-C. Yin, H.W. Hao, and K. Iqbal, (2012) "Effective Text Localization in Natural Scene Images with MSER, Geometry Based Grouping and Adaboost," in Proceedings of International Conference of Pattern Recognition, Vol. 21, No. 9, pp. 4256– 4268.
[6] J. Dina, M. Vanitha, S. Pradeepkumar, B. Udhaya, and Saravana A, (2014) "Robust Text Detection and Voice Conversion", in International Journal of Science and Engineering Communications. Vol. 03, No. 3, pp.1062-1068.
[7] Xu Cheng Yin,Xu Wang Yin., Kaizhu Huang, and Hong Wei hao, (2014) "Robust Text Detection in Natural Scene Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 5, pp 143-154.
[8] Ujjwal Bhattacharya, Bityut B, (2014) "Automatic Detection of Handwritten Texts from Video Frames of Lecturer", IEEE Conference on Frontiers Handwriting Recognition. Vol. 36, No. 5, pp137-144.
[9] Hyung Koo, and Duck Hoon Kim, (2013) "Scene Text Detection Via Connected Component Clustering and Non text Filtering", IEEE Journal on Image Processing. Vol. 22, No. 6, pp1057- 7149.
[10] Carlos Merino Gracia, Majid Mirmehdi, (2014) "Real Time Text Tracking in Natural Scenes", IETE Technical Review. Vol. 8 , No. 6, pp. 670– 681.
[11] Abdollah Amirknani Sharharki, and Amir Ebrahimi Ghahnavieh,(2014) "A Morphological Approach To Persian Handwritten Text Line Segmantation", IEEE Conference Pattern Recognition. Vol. 44, No. 4, pp. 917-928.