IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Detection of Text by Enhancing Stroke Width Transform and Maximally Stable Extremal Regions

Asst.Prof A. Vishnu Vardhan, Padarthi Maheeja , Manukonda SaiSudhanvi, Nuthalapati Preethi ,

Mannava Veena

Asst. Prof , Student , Student, Student , Student
Department of Computer Science & Engineering ,
Vasireddy Venkatadri Institute of Technology, Guntur , India

Abstract: Text Detection is one of the main goals of image processing. Various mechanisms exist to detect the text in images. The existing ones majorly detect text from scanned type and are able to identify the text in camera captured images only to some extent. These methods can't handle the features such as different size, color, font, blur and distorted images. So, to eliminate these kind of advantages, Stroke Width Transform (SWT) and Maximally Stable Extremal Regions (MSER) are the two algorithms used for this purpose. The main aim is to enhance results and improve performance by applying the two algorithms together. The output obtained by using them together is better when compared to the output individually. The flaws in the existing algorithms are eliminated by improving performance.

Keywords: MSER, Stroke Width Transform, Text detection

I. INTRODUCTION

In order to perform text detection, there is a need to know about the different types of images with respect to color and graphics, they are natural and synthetic images. Natural images are those that are captured through camera or mobile phones. Text always has different gray values, sizes and backgrounds due to which it is difficult to be detected. Our contribution to this method is a new system for automatic text detection for the purpose of inpainting, which requires no user interaction. In earlier stages, text recognition was mainly focused on analyzing scanned documents, but with increase in usage of mobile phones, tablets, PC's user applications related to image processing have higher demand. Text detection in natural images is still a challenging one. There are various applications such as support to visually impaired people, robotic navigation and scene understanding. Text information can be easily understood by human as well as computer; hence, it is of great interest and finds wide applications such as reading posters, mobile text recognition, sign detection, product cover, etc. . Text detection can be broadly classified into two main groups texture (also called region) based and CC based methods.

Texture-based methods consider the embedded text as a particular texture pattern that is distinguishable from other parts of the image and its background by scanning the image at a number of scales. This method retains various features of an image with respect to different regions. The detection of text is based on either supervised or unsupervised classifier. In [1], Color clustering feature is used to identify character regions. To verify whether the exracted region is character or not, we use SVM classifier. Continuous Adaptive shift algorithm is used to get text. In [2] , the advantage of the desirable characteristic of gray-scale invariance of local binary patterns (LBP), a modified LBP operator is designed to extract the features of the characters. Then the classifier for is made by a polynomial neural network (PNN) to get the character regions. At last a post-processing procedure including verification and fusion is used to produce text regions. [3] used the feature of Wavelet transform to obtain character regions. High frequency wavelet coefficients are considered to perfectly

identify text regions. K-Means and projection analysis are used to recognize text. The test set used here is video frames of MPEG-7 video set.

CC based methods say that text regions share similar properties such as color and distinct geometric features with a close spatial relationship. These methods are relatively new when compared to the texture based methods. In [4], thresholding mechanism is used to locate stable regions and then character regions are extracted. after recognizing the MSERs, a proficiently pruned thorough search algorithm is utilized to sift through the nesting or duplicate regions At that point the morphological features and Single-link algorithm are used to group the character regions into text areas. In [5], the Stroke Width Transform (SWT) is applied to the image. Then we detect the connected components with similar stroke widths as the candidate character regions. The text lines are recognized with the features of shape and distance. In [6], Efficient pruned exhaustive search algorithm is used to filter duplicate regions and identify character regions. To group character regions into text, we use Single link algorithm.

II. PROPOSED ALGORITHM

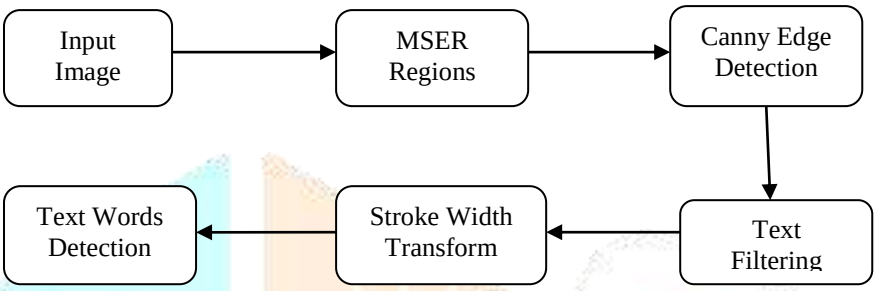

Fig.1: Text Detection flowchart

The entire process carried out in the algorithm is shown in the above figure (Fig.1). In the proposed algorithm, a CC-based approach is followed that uses two algorithms MSER and Stroke Width Transform (SWT). First, MSER regions are identified from the input image. MSER regions are those that have similar threshold value maintained through variable no.of iterations. These are also referred as stable regions. Then, canny edge detection is performed for proper identification of small letters. Later, text filtering is done to extract perfect fine edges. K-means mechanism is used in the text filtering phase. Finally, Stroke width transform is performed to extract the text in image. In this method, the intensity values for each pixel are calculated considering the stroke of nearly constant width.

Fig.2: Text detection process (a) MSER regions detected in an image, (b) Text candidates obtained after filtering and (c) Final detected text in natural image.

A. MSER

Maximally Stable Extremal Regions (MSER) is a method for detection of blobs in images. This algorithm extracts stable regions which are referred as MSER. The stable regions are identified based on the threshold values. Thresholding is done multiple times to identify stable regions. If we perform ‘n’ iterations, the value in the nth iteration and (n-1)th iteration should be same to treat it as a stable region.‘Extremal’ indicates that all pixels inside the stable regions have either higher or lower intensity than all the pixels surrounding them. This method operates well on clear images and on the images with separated boundaries. The chance of detecting the text regions in blurred images is very low with this method.

Fig.3: MSER regions in image.

B. Canny Edge Detection

To reduce the amount of data to be processed, we use a technique called canny edge detection that extracts structural information from various kinds of vision objects. This mechanism mainly focuses to decrease the memory usage by storing the edge detected image instead of the original image. It helps to remove the noise from the images. Noise refers to the non-text parts. This method detects wide range of edges in images. These edges act as boundary for stable regions and helps in removing outside information. In this process, the outline of image is produced as output.

(a) Gray-Scale Image (b) Canny edges

(c)Fine edge obtained after removing extra edges grown

Fig.4: Edge Enhancement process

C. Text Filtering

After identification of edges, non-text parts are separated by approximating aspect ratio to 1. As text letters are of same size, we use a threshold value that helps us to extract only text areas from the images.

D. Stroke Width Transform

A Stroke in the image is a continuous band of nearly constant width. The Stroke width Transform (SWT) is a local operator which calculates for each pixel the width of most likely stroke containing the pixel. First, all pixels are initialized with ∞ as their stroke width. Then, we calculate the edge map of the image by using the Canny edge detector. We consider the edges as possible stroke boundaries, and we wish to find the width of such stroke. We now have a map of the most likely stroke-widths for each pixel in the original image. The next step is to group these pixels into letter candidate. This will be done by first grouping pixels with similar stroke width, and then applying several rules to distinguish the letter candidates. Since single letters won't appear in images, we will group closely positioned letter candidates into regions of text. Optical Character Recognition (OCR) is used to convert the hand-written, printed text into machine encoded text and returns it as output.

Fig.5: Stroke width image

III. RESULTS

In this paper, we proposed a text detection algorithm that works well for both uniform and non-uniform background. The algorithm is applied on natural scenes, images captured with the help of camera or mobile phones.

IV. CONCLUSION

The main focus here is on text detection that takes place with the help of Stroke width transform and MSER algorithms. In this algorithm, MSER regions are extracted from original image which are then subjected to canny edge detection process. Later, text filtering is performed to obtain edge based text parts perfectly. Finally, Stroke width transform is done to separate text and non-text parts. This algorithm will not work effectively for blurred images.

REFERENCES

[1] Kim, K.I., Jung, K., Jin, H.K.: Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm. IEEE Trans.Pattern Anal. Mach. Intell. 25(12), 1631–1639 (2003)

[2] Ye, J., Huang, L.L., Hao, X.: Neural network based text detection in videos using local binary patterns. In: Chinese Conference on Pattern Recognition, CCPR 2009, pp. 1–5 (2009)

[3] Gllavata, J., Ewerth, R., Freisleben, B.: Text detection in images based on unsupervised classification of high-frequency wavelet coefficients. Proc. Int. Conf. Pattern Recogn. 1(3), 425–428 (2004)

[4] Huizhong Chen, Sam S. Tsai1, Georg Schroth, David M. Chen, Radek Grzeszczuk and Bernd Girod: Robust Text Detection in Natural Images with Edge-Enhanced MSER.

[5]Boris Epshtein Eyal Ofek Yonatan Wexler Microsoft Corporation: Detecting Text in Natural Scenes with Stroke Width Transform

[6] Yin, X.C., Yin, X., Huang, K., et al.: Robust text detection in natural scene images. IEEE Trans. Pattern Anal. Mach. Intell.