Analysis and Comparison of various Methods for Text Detection from Images using MSER Algorithm

Dr. Dilip Sharma
Ujjain Engineering College, Ujjain (Mp)
Email: drdilipsharma72@gmail.com

Amit Kumar Pandey
Ujjain Engineering College, Ujjain (Mp)
Email: amitrwa@gmail.com

Abstract — In this paper analysis and comparison of various methods for text detection is carried by using canny edge detection algorithm and MSER based method along with the image enhancement which results in the improved performance in terms of text detection. In addition, we improve current MSERs by developing a contrast enhancement mechanism that enhances region stability of text patterns to remove the blurring caused during the capture of image Lucy Richardson de blurring Algorithm is used.

Keywords- *MSER, CC,*

*****

I. INTRODUCTION

In present daily life text plays an important role in daily life because of its rich information that is why automatic text detection in natural images has many applications [1-4]. But detecting the text from natural image is always a challenging problem. Due to the presence of variation in the background and as the size of the text also not fixed in case of natural images it is very difficult to identify the text accurately. Through tremendous efforts have recently been devoted in this research but still reading texts in unconstrained environment is still challenging and remain a problem [4-6]. Today text detection finds many applications in various fields, including visual impairment assistance, tourist assistance, content based image retrieval and unmanned ground vehicle navigation. Today most of the images are taken from the camera and other handhold devices which is not fixed and sometimes due to movement of the object the problem of blurring is observed which makes it even more difficult to detect the text from natural images [7-9]. Here in this thesis idea is proposed to detect and recognize the text contains in the image as the main problem in computer vision is to separate the text from the background components [9-12]. There are many methods which are still used to detect the text from the natural scene such as text detection using edge analysis, robust text detection, Real time text tracking, but none of them is promising [13].

II. DIFFERENT METHODS FOR TEXT DETECTION

2.1 Texture based method: **-** Surface based techniques look at nearby composition highlights inside little districts of a picture. The content present in the pictures displays some unmistakable textural properties, which might be utilized to recognize it from the foundation. [3]

Gabor channels, Wavelets, Fast Fourier change, and so forth are generally used to remove the textural properties of a content district in a picture. In the event that the composition 35 elements are steady with the attributes of the content, all pixels in the locale are set apart as content.[8]

2.2 Region based technique: - Area based strategies use properties of the shading or dim scale in a content locale or their disparities with the relating properties of the foundation. Area based strategies can be further separated into two classes

1. Connected segment (CC) based
2. Edge-based

These techniques are otherwise called base up methodologies, because of the way they work; i.e. by first recognizing rudimentary (little) sub-structures, for example, CCs or edges, and after that blending these sub-structures progressively into bigger structures, until all the content territories are identified [14].

2.2.1 Connected part (CC)

In CC-based strategies, the fundamental components are made utilizing the 31 likeness of neighbor pixels in grayscale or shading levels, while the edge construct techniques center in light of the high differentiation between the content and the foundation, distinguishing first the edges brought on by the content shapes and afterward gathering them, if conceivable CC-based techniques utilize a base up methodology by gathering little segments into progressively bigger parts until every one of the areas are recognized in the picture. A geometrical investigation is expected to combine

the content segments utilizing their spatial course of action in order to sift through non-content segments and check the limits of the content districts. In the CC approach, little districts speaking to the content and non-content are to be distinguished. With this in perspective, shading lessening by bit dropping and shading bunching quantization is endeavored and a short time later a multi-esteem picture deterioration calculations utilized to decay the information picture into various closer view and foundation pictures [Jain and Yu 1998].

2.2.2 Edge based methodology-

In the edge based methodology is was endeavored to get high complexity edges for the continuous content hues y utilizing the red casing of the RGB shading space (Agnihotri and Dimitrova 1999). By method for the convolution procedure with various veils, first the picture is improved, and after that the edges are identified. This edge picture is further prepared by gathering the neighboring edge pixels to single associated part structures [7].

III. ANALYSIS OF EDGE BASED METHODOLOGY

3.1 Edge detection: - Edge identification is an operation in PC vision framework which recognizes the sharp change in the picture pixel. By recognizing the edges present in the picture we can extraordinarily diminish the measure of information to be handled .There are a few diverse edge identification calculation exists yet here we are concentrating for the most part on the calculation created by john F. Vigilant in 1986 [16]. In spite of one of the most seasoned technique for edge identification it is one of the standard edge recognition strategies and still utilized by the specialists.

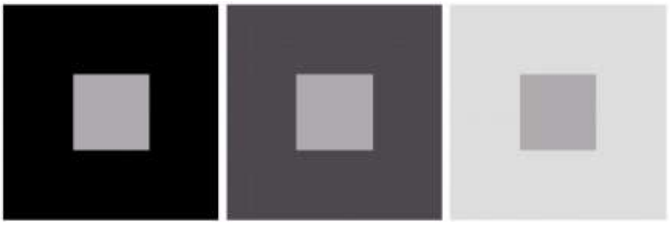
3.2 CANNY EDGE DETECTION ALGORITHM:

The vigilant edge recognition is the most generally utilized edge identification calculation to find sharp force changes which is utilized to distinguish object limit in any image. In shrewd edge location technique the calculation characterize the pixel as an edge if the angle greatness of the Pixel is bigger than those of the pixel at both its sides in the direction of maximum power change

The calculation keeps running in 5 steps:

1. Smoothing: Blurring of the picture to expel clamor.
2. Discovering angles: The edges ought to be checked where the inclinations of the picture has huge extents.
3. Non-most extreme concealment: Only neighborhood maxima ought to be set apart as edges.
4. Twofold thresholding : Potential edges are controlled by thresholding.
5. Edge following by hysteresis: Final edges are controlled by stifling all edges that are not associated with an extremely certain edge.

3.3 Image Enhancement-It is a process in which the quality of image (poor illumination, coarse quantization) is enhanced .In case of image enhancement the quality of the image need to be improved without the availability of the reference image. The idea behind the image enhancement is to produce certain changes in the image which make the vision system to easily understand the idea behind the image.



(a) (b) (c)

Fig.1: three different backgrounds with same grayscale

3.4 Contrast stretching

Low-contrast images can result from poor illumination, lack of dynamic range in the image sensor, or even wrong setting of a lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of the gray levels in the image being processed [10].

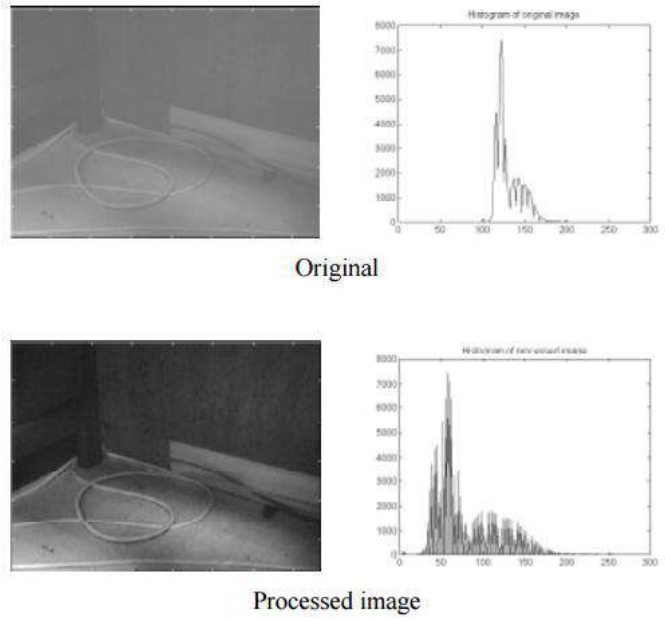


Fig. (2). Image and its histogram before and after contrast enhancement

3.5 Smoothing filter

Smoothing channels are utilized for obscuring and for clamor diminishment. Obscuring is utilized as a part of preprocessing steps, for example, expulsion of little points of interest from a Picture before item extraction and spanning of little crevices in lines or bends. Commotion diminishment can finish by obscuring with a straight channel furthermore by nonlinear sifting [12].

3.6 Maximally Stable Extremal Regions

MSER regions are connected areas characterized by almost uniform intensity, surrounded by contrasting background. They are constructed through a process of trying multiple thresholds.

The selected regions are those that maintain unchanged shapes over a large set of thresholds. For color images MSER algorithm replaced thresholding of the intensity function with Agglomerative clustering, which is based on the color gradients [3].

FIGURE(3) : EXAMPLES OF MSER REGION

3.7 MSER algorithm

MSER is a technique for blob location in pictures. The MSER calculation separates from a picture various co-variation locales, called MSERs: a MSER is a stable associated part of some dark level arrangements of the picture. • MSER depends on taking areas which stay almost the same through extensive variety of limits. – All the pixels underneath a given edge are white and every one of those above or equivalent is dark. – If we are demonstrated a grouping ofthresholded images with casing t relating to limit t, we would see initial a dark picture, then white spots comparing to nearby power minima will seem then become bigger.

These white spots will in the long run converge, until the entire picture is white. The arrangement of every associated segment in the succession is the arrangement of all extremal locales. Optionally, circular edges are appended to MSERs by fitting ovals to the districts. Those areas descriptors are kept as elements. The word extremal alludes to the property that all pixels inside the MSER have either higher (brilliant extremal districts) or lower (dim extremal locales) power than every one of the pixels on its external limit.

3.8 Methodology:

Text Recognition Phase

Step 1: Load Image

In this step firstly load the input image in which we have to detect text. Before preceding towards next step first of all the algorithm crop that portion of image that contains text, Further the text can be rotated in plane, if required.

Step 2: Noise Removal and De-blurring Image

Because of defects in the imaging and catching procedure, be that as it may, the recorded picture constantly speaks to a degraded adaptation of the first scene. The corruption results in picture blur, affecting identification and extraction of the helpful data in the pictures. It can be brought about by relative movement between the camera and the first scene, by an out of center of optical framework, environmental turbulences and deviations in the optical framework.

Lucy Richardson (LR) calculation is an iterative nonlinear restoration method.The L-R calculation emerges from most extreme probability plan in which picture is displayed with toxic substance measurements. Its execution within the sight of commotion is observed to be better than that of other deconvolution calculations.

We can use other deblurring methods also like wiener filtering.

Step 3: Contrast Adjustment and Conversion RGB image to Binary Image

Picture upgrade strategies are utilized to enhance a picture, where "enhance" is now and again characterized dispassionately (e.g., build the sign to-commotion proportion), and once in a while subjectively (e.g., make certain elements less demanding to see by altering the hues or intensities)

Further in this progression RGB Image is changed over into dim scale Image

Step 4: Edge Enhancement

In this progression, canny edge identification calculation is utilized for picture edge discovery. The calculation keeps running in 5 separate strides: Smoothing: Blurring of the picture to evacuate clamor. Discovering slopes: The edges ought to be checked where the inclinations of the picture has extensive extents.

Non-most extreme concealment: Only nearby maxima ought to be set apart as edges. Twofold thresholding: Potential edges are controlled by thresholding. Edge following by hysteresis:

Final edges are dictated by smothering all edges that are not associated with an exceptionally certain (solid) edge.
To adapt to obscured pictures the propose calculation utilized the properties of Canny edges.

Step 5: MSER region detection

As the power complexity of content to its experience is regularly critical and a uniform force or shading inside each letter can be expected, MSER is a characteristic decision for content recognition. While MSER has been distinguished as one of the best area identifiers because of its vigor against perspective point, scale, and lighting transforms, it is delicate to picture obscure. Along these lines, little letters can't be recognized or recognized in the event of movement or defocus obscure by applying plain MSER to pictures of constrained determination.

3.9 Text Extraction Phase

Step 1 and 2: Geometric Filtering and Character Connecting

With the extraction of edge-improved MSER, we get a paired picture where the forefront CCs are considered as letter hopefuls. As in most best in class content identification frameworks, we play out an arrangement of basic and adaptable geometric minds every CC to sift through non-content items. As a matter of first importance, substantial and little protests are rejected.
At that point, subsequent to most letters have angle proportion being near 1, we dismiss CCs with extensive and little viewpoint proportion. A moderate limit on the angle proportion is chosen to ensure that some extended letters, for example, "i" and "l" are not disposed of.

Step 3 & 4: Text line formation and Word separation

Content lines are imperative signs for the presence of content, as content quite often show up as straight lines or slight bends. To detect these lines, we first pair wise bunch the letter competitors utilizing the accompanying principles. The following phase of the calculation finds lines of content inside the distinguished competitor districts. This permits the aggregate number of CCs to be lessened, evacuating non-character CCs and thus enhancing the odds for higher exactness.

IV. COMPARISON

Connected component based method fails in some natural scene images which have very poor contrast text and strong illumination.

Table No. 1

Methods	Accuracy	Advantage/Disadvantage
Texture based Method	88.52%	Inefficient when background in the image is more complex like trees, vehicles.
Edge-based method,	94.66%	Works on complex backgro und. Fails for small slanted/curved text.
Morphology operators, Histogram Projection (X and Y histogram)	84.66%	Fail in case of touching characters and over-lapping lines.
Maximum Color Difference (MCD), Boundary Growing Method (BGM),	89.67%	Insensitive to contrast

Texture based techniques usually give better results in complex backgrounds than region based techniques but have computationally very heavy hence not suitable for retrieval systems for hefty databases. Therefore, there is a need to improve the detection results of region-based techniques to be used for retrieval and indexing of large multimedia data.

V. CONCLUSION

This paper presents review on existing methods for text detection, and recognition with their feature. Also this paper summarizes the key ideas, advantages, disadvantages and applications of text detection technique. Detecting and recognizing text from natural scene image is more difficult task than all other types of images. It has various affecting factors like light effects, orientation, font styles, blur, etc. Even though there are many algorithms, no single unified approach can fits for all the applications. So there is lot of scope to work with the text detection, extraction, segmentation and recognition from natural scene images. Also there is scope for detecting text from various.

REFERENCES

[1] Gao, Jiang, and Yang, Jie. An adaptive algorithm for text detection from natural scenes. In Proceedings of the, *"IEEE Conference on Computer Vision and Pattern Recognition"* (December 2001)

[2] Robust Text Detection in Natural Scene Images Xu-Cheng Yin, Member, IEEE, Xuwang Yin, Kaizhu Huang, and Hong-Wei Hao, *"IEEE Transactions on Systems",* june,(2008)

[3] Gao, Jiang, Yang, Jie, Zhang, Ying, andWaibel, Alex. Text detection and translation from natural scenes.Tech. Rep. CMU-CS-01-139, "*Carnegie Mellon University, School of Computer Science, Carnegie Mellon University, Pittsburgh"*, PA 15213, 2001.

[4] Text-Attentional Convolutional Neural Network for Scene Text Detection Tong He, Weilin Huang, Member, "*IEEE"*, Yu Qiao, Senior Member, IEEE, and Jian Yao, Senior Member, IEEE, (2007)

[5] B. Shiva Kumar Reddy, Lakshmi Boppana and Ashok Agarwal. "BER Analysis of CVSD Vocoder for WiMAX using GNU Radio". IEEE Region 10 Symposium, 2014 IEEE.

[6] Real-time text tracking in natural scenesbyCarlos Merino-Gracia1,2, Majid Mirmehdi2 Neurochemistry and Neuroimaging Laboratory, "University of La Laguna" , La Laguna, Spain, (2009)

[7] Gao, Jiang, and Yang, Jie. An adaptive algorithm for text detection from natural scenes. In Proceedings of the , "IEEE Conference on Computer Vision and Pattern Recognition" (December2001)

[8] Gao, Jiang, Yang, Jie, Zhang, Ying, andWaibel, Alex. Text detection and translation from natural scenes.Tech. Rep. CMU-CS-01-139, "Carnegie Mellon University, School of Computer Science, Carnegie Mellon University, Pittsburgh", PA15213,2001

[9] Zhang, Dong-Qing, and Chang, Shih-Fu. Learning to detect scene text using a higher-order MRF with belief propagation.In IEEE Workshop on Learning in Computer Vision and Pattern Recognition vol. 06, pp. 101–108 (2004).

[10] Y.-F. Pan, X. Hou, and C.-L.Liu.Text localization in natural scene images based on conditional random _eld. In ICDAR, pages 6-10, IEEE Computer Society, (2009).

[11] Jain, A.K., and Bhattacharjee, S. Text segmentation using Gabor filters for automatic document processing. Machine Vision Applications 5 pages 169– 184, (1992).

[12] Wu, Victor, Manmatha, R., and Riseman, Edward M. Finding text in images. In "Proc. Intl. Conf. on Digital Libraries" (1997).

[13] Wu, Victor, Manmatha, R., and Riseman, Edward M. Finding text in images. In "Proc. Intl. Conf. on Digital Libraries" (1997).

[14] Thillou, C´eline, Ferreira, Silvio, and Gosselin, Bernard. An "embedded application for degraded text recognition.Eurasip Journal on Applied Signal Processing" 13 (2005).

[15] Garcia, C., and Apostolidis, X. Text detection and segmentation in complex color images. In "Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing" vol. 4, pp. 2326–2330, (June 2000).

[16] Sergei Azernikov. Sweeping solids on manifolds. In "Symposium on Solid and Physical Modeling", pages 249–255,2008.

[17] John Canny. A computational approach to edge detection.Pattern Analysis and Machine Intelligence, "IEEE Transactions on, PAMI"-8(6):679–698, Nov. (1986).

[18] F. Mai, Y. Hung, H. Zhong, and W. Sze."A hierarchical approach for fast and robust ellipse extraction. Pattern Recognition", 41(8):2512–2524, August (2008).

[19] David Ger´onimo, Antonio L´opez, and Angel D. Sappa Computer Vision Center, "University at Aut`onoma de Barcelona Edifici O, 08193 Bellaterra, Barcelona", Spain,june (2007)

[20] Piotr Doll ´ ar, Christian Wojek, Bernt Schiele, and Pietro Perona, Submission to "IEEE Transaction on pattern analysis and machine intelligence", vol .1, page no. 1-19, (2012)

[21] X. Lin. Reliable OCR solution for digital content re-mastering. In Society of "Photo-Optical Instrumentation Engineers (SPIE) Conference Series", (Dec. 2001).