International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

Detection, Localization of Text in Images by Mser and Enhanced Swt

S.Shiyamala, S.Suganya

Abstract--The number digital images and digital videos has raised highly. The text shows in the images is major thing for fully understanding the images. So it is very signifigant to recognize the text. First step of text recognition system is known as text detection. But detecting text from images is very challenging and it is receiving high amount of concentration. Text detection is improved by pre-procssing techniques like noise removal, contrast changes and etc. Recently MSER (Maximally Stable Edge Region) and SWT (Stroke Width Transformation) methods are used seperately or combined with each other for detecting and localizing text. This paper analyzes and compares technical challenges and performance of above methods and take in new and fast method for text detection and localization. The proposed method comprises an enhanced technique based on SWT which combined with MSER. The proposed method detect wording which is located in differnt directions, different angle and combined letters. Performance of exisisting and proposed methods on road signal images depicted in experimental results.

Keywords : MSER, SWT ,Canny edge detection

I. INTRODUCTION

Detecting and localizing text in real world images are signifigant fragment computer vision applications , such as such as onscreen assist for visually weakened, robotic and automatic navigation. The complexity of the problem comes out of many factors which are size, light and language. In natural scenes, numerous objects, such as buildings,bikes, cars or parts of them have similar shape and appearances to text. For example corner of a building can easily be recognized as I or wheels of vehicle as O. These challenges make the role of text detection and localization techniques very precious to discriminate text from non-text.

Identifying the text regions are divided into two steps. One is text detection and another one is text localization.Text detection is used to distinguish the presence of text in an image. The results of detection are grouped together to form one or more text instances. This is usually draw boundary line just about the text area.

Scene text detection and lacalization can roughly be categorized into three groups: sliding window, Texture-based and Hybrid methods. Sliding window method engages a window to look for for likely texts in the copy and then employ mechanism learning skill to make out texts. Region based method divided into two groups which are connected component methods and edge based methods.

Manuscript published on 30 August 2019.
***Correspondence Author(s)**
Mrs. S.Shiyamala, pursued M.SC.,M.Phil., in computer science from Bharathiar university, Coimbatore.
S.Suganya, Associate professor, Department of computer science, Rathnavel Subramaniam College of arts and science, Sulur,Coimbatore.

Connected component based methods extract character candidates from images by connected component analysis followed by grouping character candidates interested in text; additional verify may be carry out to eliminate false positives. Edges are a reliable feature of text regardless of color/intensity, layout, orientations, etc. Edge based method is focused on far above the ground contrast between the text and the backdrop. Texture-based methods use the watching that text in images has distinct textural properties that distinguish them from the background. Mathematical morphology is a topological and geometrical based method for image analysis. Morphological feature extraction techniques have been efficiently applied to character recognition and document analysis.

Fig .1 Text detected and localized of an traffic signal image

Fig .1 Shows, text detection and lacalization system of a traffic signal image. The basic idea of text detection and localization is to improve the extraction part. At the present text detection system mostly uses MSER technique or SWT or combination of techniques. Exisisting system also used the above techniques. But it fails to handle transparent text, text is too small, blurred text, curved text and text with strong highlights and the qualitative results of existing system [1] show that there is still for improvement. The resulting of proposed system is able to detect text not considering of its size, direction, fonttype and language. This work differentiates from the preceding ones by enhancing the SWT technique whose output enables fast and dependable detection of text. The resulting of proposed system is clever to detect text apart from of its range, direction, fonttype and verbal communication.

II. LITRATURE SURVEY

Component-based method[1] is first extracted text from image. In modern duration, component-based method has developed into the mainstream of text detection. The Maximally Stable Extremal Regions (MSER) is tough to the affine difference of the picture, and can efficiently mine the text section in image. Texture-based method[2] concerns for text has a individual kind of quality and utilize its properties. The properties may be limited intensities, sort out responses and wavelet coefficients to differentiate among wording and non-wording regions of an image.

Hybrid system [3] obtains benefits of both texture-based methods and component-based ones. A hybrid and multi-scale text detection algorithms that can improved handle "tough text" such as multi-size, multi-colour and multi-orientation. Some learning be statement in remoteness based category. The view of remoteness is basic in picture study. Models for picture demonstration usually engage least amount remoteness among pixels and boundaries of mechanism. It is frequently dimensions are functioning flanked by a tip and a set somewhat than linking points. Stroke width has been taking on as a fine text descriptor in text investigation. But stroke width transform fail to handle transparent text, text is too small, blurred text, curved text and text with strong highlights. So still SWT needs to be improved, that makes the text extraction system more robust and accurate.

III. PROPOSED WORK

This segment consists of a novel way of identifying and localizing letter in normal scene imagery. This method mostly stand on MSER and stroke width. Canny edge detection is used just before spot an edge map. It is combined with MSER method which is used to detect maximally stable edge regions. Finally enhanced SWT algorithm is applied. Flowchart of text detection and localization system is shown in Fig.2. Each block in flowchart is fully explained with a sample output as a practiced example.

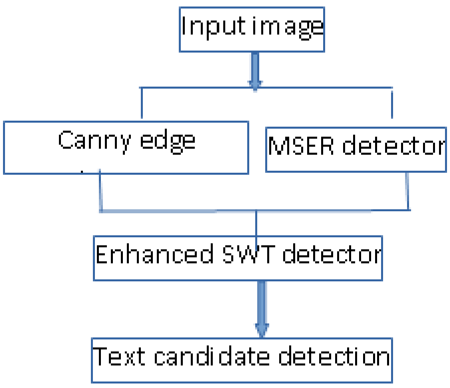

Fig. 2 Flowchart of text detection and localization system

IV. CANNY EDGE DETECTION

Canny edge detection becomes aware of boundaries with noise reduction at the same time.

1. Smooth the picture by means of a Gaussian filter

$$g(m,n) = G_\sigma(m,n) * f(m,n)$$

where

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{m^2+n^2}{2\sigma^2}\right)$$

2. Compute gradient of $g(m,n)$ using any of the gradient operators (Roberts, Sobel, Prewitt, etc) to get:

$$M(n,n) = \sqrt{g_m^2(m,n) + g_n^2(m,n)}$$

and

$$\theta(m,n) = tan^{-1}[g_n(m,n)/g_m(m,n)]$$

3. Threshold M:

$$M_T(m,n) = \begin{cases} M(m,n) & \text{if } M(m,n) > T \\ 0 & \text{otherwise} \end{cases}$$

4. Repress non-maxima pixels in the edges in M_T obtained above to thin the edge edge. $M_T(m,n)$ non-zero is greater than its two neighbours along the gradient direction $\theta(m,n)$. If so, keep $M_T(m,n)$ unchanged, otherwise, set it to 0.
5. Threshold the previous result by two different thresholds τ_1 and τ_2 (where $\tau_1 < \tau_2$) to obtain two binary images T_1 and T_2. Note that T_2 with greater τ_2 has fewer noise and smaller amount false limits but superior space between edge segments, when compared to T_1 with smaller τ_1.
6. Tie edge section in T_2 to form unbroken edges. To do so, outline each part in T_2 to its last part .Search its neighbours in T_1 to discover every boundary segment in T_1 to flyover the gap awaiting attainment a further edge part in T_2.

V. MSER TEXT DETECTION

The fundamental code of MSER is towards obtaining a threshold worth from 0 to 255. Then obtains the corresponding greyscale images. In all the gray scale images obtained, a few of the connected area modify little or even keep on the alike. These regions are called the Maximally Stable Extreme Regions. Its algebraic description is:

Q(i) = | Q i + Δ − Q i + Δ | / | Q i | (1)

Q(i) indicate the linked region of threshold i, signify the small alteration of grey value. q (i) is the varying rate of section Q(i) by threshold i. While q(i) is local minimal, the Q i is the Maximally Stable Extreme Regions.

Retrieval Number: J96120881019/19©BEIESP
DOI: 10.35940/ijitee.J9612.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

VI. ENHANCED SWT

The result of canny edge detection and MSER are merged together. The merged result is given as input for enhanced SWT. It has three main stepladders: the SWT, assemblage pixels keen on pixel based on their stroke width, and merging pixel hooked on regions of text.

Computes per pixel, width of the most likely stroke containing pixel [4].Then work out the least number of four row segments fly-by-night through every tip of element, p0, in four orientations: horizontal, vertical, diagonal 45 and diagonal 135 degrees. The enhanced SWT algorithm calculates as following.

1.Initially set SWT = ∞

2.Ray r=p + n · dp, n>0

3.If dq=-dp ± Π/6 then SWT= |p-q| where θ =0, θ =45,θ =90,and θ =135 and dp++ else discard the ray.

4. Calculate smallest of four line segments passing through each point by using Euclidian distance.

5.If SWT ratio <= 3 then group neighbouring pixels.

6.If two letters are contain related stroke width, that can be assemblage.

7.At last draw rectangle box for identified words.

VII. RESULT AND DISCUSSION

In order to get better text detection, it is essential to improve the SWT. A set of qualitative metrics such as, precision rate, recall rate and F- measure are used here to compare the results. Table 1 demonstrate that the result of quality metrics for each techniques which are used to discover the letters from images. Table 1 depicts that Enhanced SWT give best result.

The Fig 3. depicts the result of different algorithms implemented in a road signal image. based on the table values. According to the chart, enhanced SWT shows the highest performance. Fig 4.shows the outcome result of enhanced SWT on road signal images.

Table 1: Quality metrics and techniques used to discover the wording from natural images.

Techniques / Measure	Precision rate	Recall rate	F-measure
Enhanced SWT	0.81	0.73	0.61
SWT	0.72	0.62	0.57

Fig .3 Performance Comparison between Techniques in different metrics

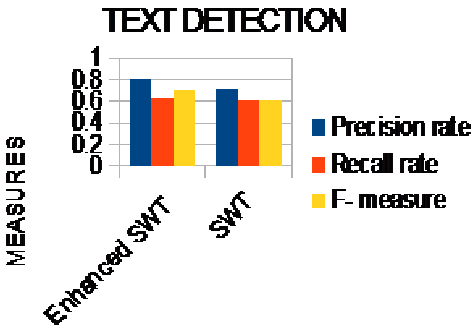

Fig 4. Outcome of enhanced SWT

Fig 4. Depicts that the outcome of enhanced SWT.

VIII. CONCLUSION

In this paper, it is been concluded that the improved SWT algorithm detects the text regions in natural images and recognize the text from the same detected textual area, even it detects the curvy texts and handwritten text regions also. An improved SWT method of components can improve the detection of characters. This way, better results for circular text can be detected which cannot be done by normal SWT. This also allows identifying curvy letters better.

Also there is scope for detecting text from various languages, which have different characteristics than English .In many applications, text or graphics with the predefined thickness should be extracted. In such applications, the proposed method seems to work well. Therefore, the proposed way has the capability to take a broad view.

REFERENCES

1. Gomez, L., Karatzas, D. (2014). A fast hierarchical method for multi-script and arbitrary oriented scene text extraction, 2014, 19(4) 1-15.
2. Wu, H., Zou, B., Zhao, YQ. (2016). Natural scene text detection by multi-scale adaptive colour clustering and non-text filtering, Neurocomputing, 214. 1011–1025.
3. Tian, Z, Huang, W, He, T, et al. (2016). Detecting Text in Natural Image with Connectionist Text Proposal Network, Computer Vision – ECCV 2016. Springer International Publishing, 2016.
4. Epshtein, B., E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform in **IEEE Conference on Computer** Vision and Pattern Recognition (CVPR). 2010.
5. Chucai, Y. and T. YingLi, Text String Detection From Natural Scenes by Structure-Based Partition and Grouping. Image Processing, IEEE Transactions on, 2011.20(9): pp. 2594-2605.

Retrieval Number: J96120881019/19©BEIESP
DOI: 10.35940/ijitee.J9612.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

AUTHORS PROFILE

Mrs. S.Shiyamala pursued M.SC.,M.Phil., in computer science from Bharathiar university, Coimbatore. She is currently doing her P.hd in Rathnavel subramaniam college, Sulur. She is a life member of Computer Soceity of India since 2008. She has published a review paper in text extraction. Her main research work focuses on Image processing.

S.Suganya, received her Ph. D from Anna University , Chennai. She is currently working as Associate professor, Department of computer science, Rathnavel Subramaniam College of arts and science, Sulur,Coimbatore. Her research interest includes Mobile Computing, Machine Learning, Deep Learning and Image processing. She has published 15 technical papers in various international journals. She has produced three Ph.Ds and 8 scholars pursuing Ph.D under her guidance.

Retrieval Number: J96120881019/19©BEIESP
DOI: 10.35940/ijitee.J9612.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)