ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

NCAIT-2017

JSS Academy of Technical Education, Bengaluru

Vol. 6, Special Issue 5, May 2017

Text Detection from Images using Sharpening and Blurring Technique

Ajay B N[1], N Swetha[2]

Assistant Professor, Department of CSE, SJBIT, Bangalore, India[1]

PG Student, Department of CSE, SJBIT, Bangalore, India[2]

Abstract: Text plays a pivotal role in the images. Text in images accommodate some useful raw facts about the scene and surroundings. Therefore, it is important to detect and recognize them. Some of the limitations to detect texts are intensity, fonts and sometimes even because of complex background. In order to overcome these limitations, first sharpen and blur the image using Gaussian blur and low pass filters, then all the covariant regions are detected. Using all the covariant regions, convolute the image with the Gaussian kernel and continuously down sample the image by half. After convoluting, create a feature map and apply morphological gradient to the resultant image in order to cluster the text regions. After doing this, the non-text regions are eliminated and the text regions are detected. The proposed approach is experimented on ICDAR 2015 dataset and the results show that the proposed approach is robust and it outperforms for various kinds of images.

Keywords: Gaussian kernel, text detection, morphological gradient.

I. INTRODUCTION

Text detection in images is mainly concerned with identifying the position of the texts in images. Text in images are useful in order to understand the image and also for information grabbing. But it is not very easy to detect them because of the fonts that are used in the text, languages, and sometimes even because of the colors used and also due to the degraded quality of the text. In images, text characters are usually close to each other. Therefore, by using the proposed strategy we can easily identify the text from images. There are many text detection methods like the edge connected, texture based, component connected etc. However each of these methods have their own limitations. The main difficulty arises when there is a complex background or when an unknown color is used in the text.

(a) Original Image

(b) Text Detected Image

(c) Original Image

(d) Text Detected Image

Figure 1: Results of the proposed approach (indicated by yellow lines) on plain and complex background

As stated in [1], the text can be detected from a complex background using multi-degree of sharpening and blurring process. This technique can be used for low quality images. In this, features like spatial layout, color, and distance are calculated in order to construct the text candidates. OCR (Optical Character Recognition) can detect text only against a plain background and not on complex background. The most recent technique i.e., MSER (Maximally Stable Extremely Regions) usually detect text which has high color contrast against a complex background [3].

In earlier approaches [11], the character candidates were extracted using Text Confidence map and the text candidate detection, but these two methods fail for complex background. Therefore MSER technique and edge detection technique were merged together to detect the text. These two techniques detect maximum number of characters. After this, Dynamic Time Warping (DTW) using the HOG & SIFT features are used to filter out the non-text regions.

In sliding window approach, the windows are slided with different scales and also through different positions in

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

NCAIT-2017

JSS Academy of Technical Education, Bengaluru

Vol. 6, Special Issue 5, May 2017

order to extract the texts[4-7]. The main advantage of using this approach is that almost all the true regions are detected and the problem with this is that they generate many candidate regions, which needs to be detected in the further processes. This leads to high number of computation. The canny detector method uses edges to detect text. This method depends on edge detection to connect components therefore it is sensitive to blur images. Peijun Tang [8], proposed a novel approach to detect colorful texts. In [9], a text line detection method is used in order to detect text from camera captured document images. This method uses MSER algorithm and then extracts CC's by using bottom up clustering method. This method can detect random oriented text lines. In [10], a text detection method is introduced which is based on sparse representation and morphological component analysis. This method outperforms for natural scene images. It does not have any limitation on texture of the text size or color.

Figure 1 shows that the proposed approach can detect text from both plain and complex background. The proposed approach is experimented on ICDAR 2015 dataset and the results show that the proposed approach is robust and it outperforms for various kinds of images.

II. METHODOLOGY

The proposed approach contains the following phases: Sharpening and blurring, covariant region detection, Text candidate detection, Non-text candidate elimination, text candidate classification. Fig 2 shows the block diagram of the proposed approach and Fig 3 shows different phases of the proposed approach.

2.1 Sharpening and Blurring Process

Sharpening and blurring is done in order to destroy distortion and also to improve the image quality for text detection. Blurring makes the background simpler and it allows text components to be extracted from complex background in a simpler manner .The input image is first sharpened using the Gaussian filter in order to avoid unnecessary noise in the image. The resultant image is shown in Fig 3(a). Then the image is subtracted from the input image as shown below:

$$I(x,y)= \{S(x,y) - T(x,y)\}$$

Where T(x,y) is the input image and S(x,y) is the intricacy of the input image with the Gaussian filter.

2.2 Covariant region detection

After being sharpened and blurred, an image is obtained. Usually images will have different intensity function which is useful for text detection. The regions which have same/uniform intensity are characterized as covariant regions. These regions are obtained by trying out with multiple thresholds. Whenever the regions do not change their shape over a multiple thresholds then they are considered as covariant regions. Then compute the connected regions for each threshold.

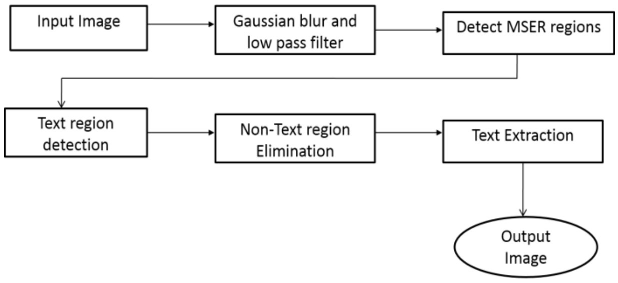

Fig 2: Block Diagram of proposed approach

2.3 Text candidate detection

The image is successively smoothened and sub sampled and then it is decimated at four levels. Decimation is a process of resizing the image resolution from its original resolution to a lower resolution. The image is decimated in the following sequence: 1/2, 1/4, 1/8, 1/16 and so on. Here we are decimating only at four levels.

These images are convoluted using the direction filter at different orientations as shown in Fig 3(b). Then each pixel is weighted to verify it as a text candidate or non-text candidate. If the pixel is highlighted in all the edges then it is treated as a text candidate or else it is treated as non-text. Then all the pixels are merged. The resultant image is the combination of all the pixels which is known as Feature Map. The image after being feature mapped is shown in Fig 3(c). The next step is the construction of text candidate and elimination of non-text regions. In order to construct the text candidate, the isolation of text regions is done. One of the property of text is that the characters appear close to each other to form a cluster. This can be done using morphological gradient method. This method enhances the regions which lie close to each other in order to cluster the text characters together and thus eliminates the pixels which are far away from each other as shown in Fig 3(d). During enhancing the regions, some of the noise blobs gets added. Therefore we need to eliminate them from the images.

III. EXPERIMENTS AND RESULTS

The experiments are carried out on the standard dataset i.e., ICDAR .This dataset is mainly used to check the robustness of the method therefore, it is called as robust reading database. This dataset contains different kinds of images like blue image, image with uneven lightning, mirror reflecting image and so on. ICDAR 2015 contains 229 training set and 233 test set images.

Here we compare our method and the top scoring methods like Zhang's method [4], Qin's method [12], Tian's method [13] and Neumann's method [14]. Precision, recall and f-measure are used to evaluate these methods.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

NCAIT-2017

JSS Academy of Technical Education, Bengaluru

Vol. 6, Special Issue 5, May 2017

Table 1: Performance of different text detection methods on ICDAR 2015 dataset

Approach	Year	Precision	Recall	f-measure
Our Method	-	0.80	**0.81**	0.72
Qin et.al [12]	2016	**0.88**	0.78	**0.83**
Neumann et.al [14]	2015	0.81	0.72	0.77
Tian et.al [13]	2015	0.85	0.75	0.80
Zhang et.al [4]	2015	**0.88**	0.74	0.80

$$\text{Precision Rate} = \frac{\text{Correctly detected words}}{\text{Correctly detected words} + \text{False positives}} * 100$$

$$\text{Recall Rate} = \frac{\text{Correctly detected words}}{\text{Correctly detected words} + \text{False negatives}} * 100$$

Precision rate is the ratio of correctly detected words to the sum of correctly detected words plus false positives. False positivesare the regions in the image which are actually not characters of a text, but still have been detected by the proposed approach as text regions. Recall rate is the ratio of correctly detected words to the sum of correctly detected words plus false negatives. False Negativesare the regions in the image which are actually text characters, but still have not been detected by the proposed approach.

From table 1, we can see that the proposed method gets excellent performance in terms of recall rate. Compared to other methods, the performance improvement in terms of recall rate is 3% higher for ICDAR 2015 dataset. Fig 4 shows the successful detection of text from the proposed method on the ICDAR 2015 dataset and Fig 5 shows unsuccessful detection of text from the proposed method on the same dataset.

(A) Sharpened and blurred image

(B) Convoluted Image

(C) Feature Map

(D) Text Detected Image

Fig 3: Various phases that are carried out to detect text

The unsuccessful detection of text occurs because of the image quality or if text like structure is present in the image. The proposed method overcomes several challenging scene images like single character, low quality image, low contrast image, mirror image etc., some of which cannot be successfully detected by other approaches.

IV CONCLUSION AND FUTURE WORK

In this paper, we have proposed a novel approach to detect text from images. In order to improve the quality of image, image is sharpened and blurred using Gaussian blur and low pass filters. Then all the covariant regions are extracted from the preprocessed image. After extracting the covariant regions, the image is convoluted using the Gaussian kernel and the image is down sampled by half. A feature Map is created and in order to cluster the text regions, morphological gradient is applied to the image. After doing this, non-text regions are eliminated and the text regions are detected.

Experiments are carried out on ICDAR 2015 dataset and the results show that the proposed approach outperforms for various kinds of images. Here are some of the future directions that can be made for the proposed approach: (1) for sharpening and blurring, Gaussian blur and low pass filters are used, instead other filters can also be used. (2) Some of the non-text regions are also detected in few of the images, which can be avoided.

Fig 4: Successful samples of the proposed method on ICDAR 2015 database

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

NCAIT-2017

JSS Academy of Technical Education, Bengaluru

Vol. 6, Special Issue 5, May 2017

Fig 5: Unsuccessful samples of the proposed method on ICDAR 2015 database

REFERENCES

[1] Juhua Liu, Hai Su, Yaohua Yi, Wenbin, "Robust text detection via multi-degree of sharpening and blurring", 2016
[2] Youbao Tang, Xiangqian Wu, "Scene Text Detection via Edge Cue and Multi-Features", International Conference on Frontiers in handwriting recognition, 2016
[3] Roussem Turki, Mohamed Ben Ralima, Ade1 M. Alimi, "Scene Text Detection Images With Pyramid Image and MSER Enhanced", IEEE, 2015
[4] Z. Zhang, W. Shen, C. Yao, and X. Bai. Symmetry-based text linedetection in natural scenes. In proceedings of CVPR, 2015.
[5] K. Wang, B. Babenko, and S. Belongie. End-to-end scene textrecognition. In proceedings of ICCV, 2011.
[6] A. Mishra, K. Alahari, and C. Jawahar. Top-down and bottom-up cuesfor scene text recognition. In proceedings of CVPR, 2012.
[7] M. Jaderberg, A. Vedaldi, and A. Zisserman. Deep features for text spotting. In proceedings of ECCV, 2014.
[8] Peijun Tang, Yule Yuan, Jin Fang, Yong Zhao, "A Novel Similar Background ComponentsConnection Algorithm for Colorful Text Detection in Natural Images", IEEE, 2015
[9] Hyung Il Koo," Text-Line Detection in Camera-Captured DocumentImages Using the State Estimationof Connected Components",IEEE Transactions on Image Processing, 2016
[10] Shuping Liu, Yantuan Xian, Huafeng Li, and Zhengtao Yu,"Text Detection in Natural Scene Images Using Morphological Component Analysis and Laplacian Dictionary", IEEE/CAA Journal Of Automatica Sinica, 2015
[11] X.Chen,A.Yuille,Detectingand reading text in natural scenes, ProceedingsoftheIEEEComputerSocietyConferenceonComputer Vision andPatternRecognition,2004
[12] S. Qin and R. Manduchi. A Fast and Robust Text Spotter. Inproceedings of WACV, 2016.
[13] S. Tian, Y. Pan, C. Huang, S. Lu, K. Yu, and C.L. Tan. Text Flow: AUnified Text Detection System in Natural Scene Images. In Proceedings of ICCV, 2015.
[14] L. Neumann and J. Matas. Scene text localization and recognition withoriented stroke detection. In proceedings of ICCV, 2013.
[15] T. He, H. W, Y. Qiao, and J. Yao. Text-Attentional Convolutional Neural Networks for Scene Text Detection. IEEE Transactions on Image Processing, 2016.