ISSN: 2321-7782 (Online)
Volume 3, Issue 8, August 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Character Recognition in Natural Scene Images using Novel Technology

Renuka G. Sitafale[1]
M.E 2nd Sem Department Computer Science & Engg.
Anuradha Engg. College, Chikhli
Maharashtra, INDIA

Prof. P.P.Kulkarni[2]
Department of Computer Science & Engg.
Anuradha Engg. College, Chikhli
Maharashtra, INDIA

Abstract: ***With the increasing use of digital media, video and images in daily day to day life, text retrieval in this medium has become research area. Nowadays, text localization and text recognition is one of the research areas which is useful for recognizing the text. Character Recognition in natural images has many applications as Getting License number from license plate, Getting book title from front cover of book, sign detection. In this paper text recognition and detection is done by using pre-processing, text localization and optical character recognition. This paper provides an improved technology to recognize text extracted from video frames and images provided by user.***

Keywords: Localization, Segmentation, Template matching, Edge based text.

I. INTRODUCTION

As this is the era of digitalization, the use of internet, digital media or digital media capturing devices are widely increases for example mobile phones, cameras etc people are capturing images, video and upload it to the internet websites like mypics, youtube, instagram etc. Text retrieval in this medium has become research area now a day's[2].

A lot of research work has been done on text localization and recognition. For this the technology used is classified into three classes as 1) texture based, 2) connected component based and 3) edge based. From the camera images or video text can be easily retrived and compared to another semantic content.

The recognized texts is relevant to the video or image it has many applications like content based web search, logo detection in CCTV video feeds, sign detection, licence plate reading.

Extracted text can widely use as main component for indexing or parsing of video or images. But some text may not be recognized or localized properly due to some constraints like contrast, color and stationary location[4]. Optical character recognition is an important technology which is used for recognizing the text localized in the image. Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data.

There are some approaches of the OCR as

A. Binary input image: Input to the optical character recognition is binary input which we get from output of threshold image as in the total black and white form or we can use output image of clustering or sobel edge detection algorithm to recognize text in the caption.

B. Thinning: Thinning can be applied to binary input image for easy process and time saving process in the optical character recognition .As thick characters can be complex in process and cannot easily process due to color contrast and background hence we need to thin it[13].

C. Scaling: As all characters should be in random in size, shape and position. All has to be in one size it can be achieved by scaling it. Dirty function is use to scale the characters.

D. Template matching: By using technique of template matching we can find small part of template image. Template match process can be used to navigate mobile robot and to detect edges.

E. Training: Training is important constraint in the optical character recognition. If some character detected is not present in the stored database then it has to be train to recognize it.

F. Image retrieval includes several processes such as image preprocessing, edge detection, clustering, localization, blurring and recognition of text from the video or image provided by user[6].

Retrieval of text from video or file involves several steps as:

- Frame Extraction: Extract the frames from the video provided by user.
- Text localization: Find out the area where text is localized.
- Text segmentation: Segment the vertical line and horizontal character on binary input image.
- Text recognition: Text recognition is carried out by thinning, training the text.

One of the use, as visually impaired person can't see it will useful for them to access text and clustered with text to speech algorithm and make them to read cover of book, labels on door, medicine labels etc. hence caption localization and detection become important research now a days. Our main moto of the proposed system is to detect text in internet video, low quality images downloaded from internet or capturing devices[3].

II. Literature Review

- Bharatratna P, Gaikwad Ramesh R. Manza Ganesh R. Manza worked on template matching, tracking and feature classification algorithm implements to optical character recognition stated that they got text recognition in uppercase letter at 92%.
- Jie Xi worked on detection of text and recognition for retrieving clue text in several superimposed text appear in news videos. In is proposed algorithm there is opening procedure of morphological on the blurred edge area map. Recognition rate for localization in their method is 94.7% while precision rate of recognizing is 67.5%.
- Palaiahnakote Shivakumara had taken segmented text and worked on non significant elimination of edged to get text line boundary at 93%.
- Rainer Lien hart research on the text localization and Text segmentation in videos and images proposed hat every line text with sub pixel had tracked with the efficient rate and recognition at 69%.

III. Proposed Technology

Proposed technology of text recognition works in the fashion that it involves several processes from extracting frames to the recognized output in the binary image.

A. Video Frame Extraction

As video contains combination of images so we have to get these images for the further processing on the frames or images which we are extracting from the video. Only we have to give text containing video the frame extractor. Output of frame extractor is frames in the form of .png. We probably take 10 frames for 30 second video file.

B. Pre-processing

Title In the preprocessing phase several contents are there for processing on images is grayscale conversion, thresholding, sobel edge detection, clustering. If extracted frame from video is color image because as described in the introduction it has some drawback about complex background hence for easy process it has to convert in grayscale. Grayscale conversion carried out by retrieving the RGB contents from the image taking average of it and assign to the new pixel. Thresholding is completely black and white image can be get by putting any value between 0-255 e.g. 126 to compare it with value of grayscale pixel. Once we get threshold image sobel edge detection is to be apply on it to detect the edges in the images as text is combination of strokes of edges. Sobel edge detection algorithm finds the magnitude of edges in x and y direction. Magnitude of x and y direction is combined with the direction of pixel because most of the character has opposite edge pairs going in opposite direction. Sobel edge filter 3x3 matrix template for finding the gradient magnitude[11].

Gradient magnitude in x direction is termed as Gx and Gy fo y direction.

Gx=X1*X2

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad X1=$$

X1 is the values of default sobel filer and X2 are the values of the edge pixel.

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

Gy=Y1*Y2

Y1=

Sobel gradient= sqrt (Gx2 +Gy2).

Edge direction =tan-1 (Gy/Gx).

In this way we can get the values of edge gradient magnitude and direction. Clustering: As in the sobel edge detection we have found the edge values in this section we are clustering the edge pixels got by sobel operator having same properties this can be achieved by finding the frequency counts the x direction in one time to the edge pixels. Those areas having approximate frequency counts are grouped by two lines i.e. red and green lines. The edges found in between these two lines are referred as localized text which is used for further processing for recognition. 3. Text recognition: There are several processing in text recognition are as follows: Vertical segmentation, Horizontal character segmentation, Thinning, Scaling, Template matching, Training.

Horizontal character and vertical line segmentation: Horizontal character and vertical line segment finds the black pixel in image provided by user to start the scan algorithm and by combining it with horizontal a separate rectangular box is assign to each character or text in the image for the thinning purpose[8]. Thinning: As in the above segment we get the segment of character but if it is variable in size it can not recognized easily by OCR hence for getting accuracy it has to thin so in this segment we are thinning the character to recognize it well.

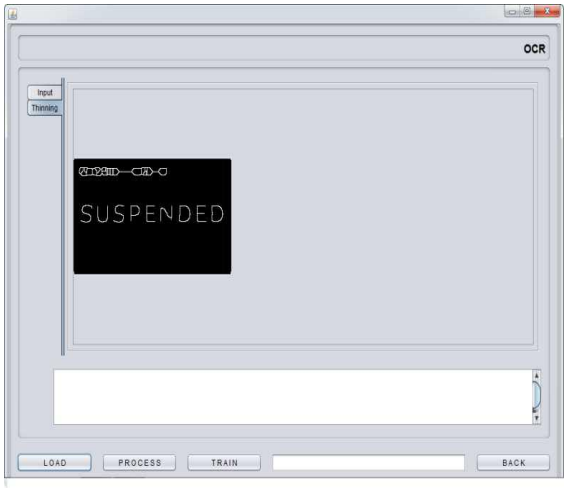

Fig 1: Thinning character

In above fig we can see we have thin the character thinning works in three stages as pixel which we are deleting should not be connected to more than two pixels. Several template standards should be containing in it. Scaling: Scaling is needed to be all character will be in one size for better result. Template Matching: Whatever the characters which we have thinned and scaled has to be match in the stored database to recognize what kind of text is present in the image provided by user. Template matching takes the binary image of text and searches it with the stored database if not present in database then it has to trains in training phase by considering shape of text.

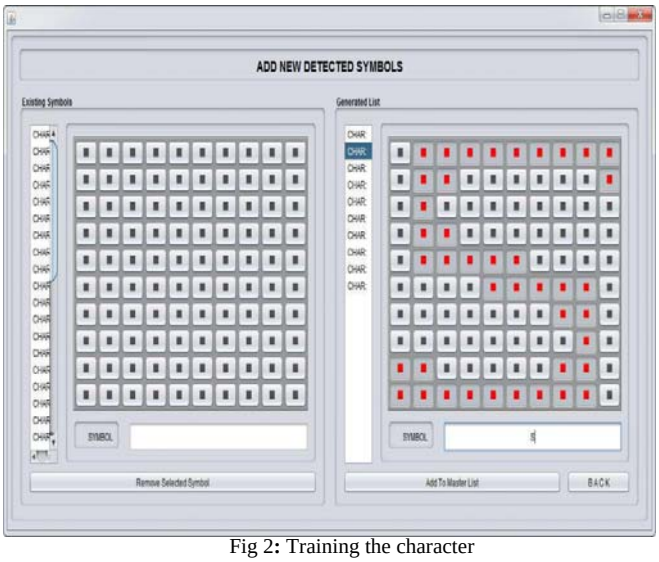

Fig 2**:** Training the character

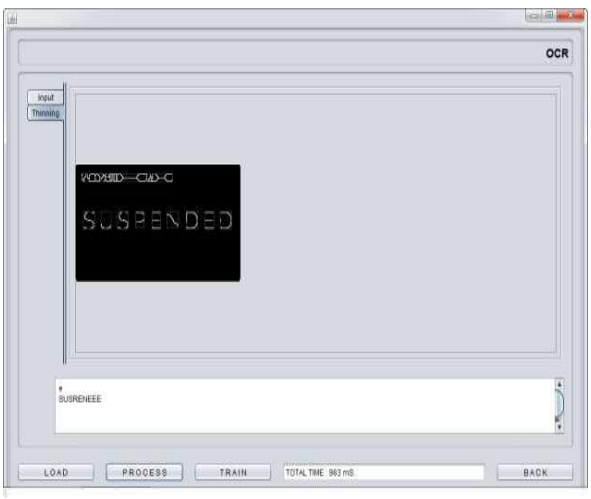

Fig 3: final output

IV. Experimental Results

Experimental carried out on random images andvideos captured by camera in natural scene image to recognize the text.Precision rate= No. of accurate detected textNo. of accurate detected text+ False positiveFalse positive rate= No. of inaccurate textNo. of accurate detected textRecall rate= No. of accurate detected textNo. of accurate detected text+ No. of missed text.

TABLE I
Recognition Table

Text	Precision Rate	Recall rate
Uppercase	92	93
Lowercase	92.5	92.6
N umbers	91	93

V. Conclusion

This paper presented character recognition in natural scene images using the novel technology with the text localization and text recognition as main concept .In this paper we have approached a new novel technology for recognizing text in images. The system is based on segmentation, OCR approaches like Template matching and training the text. The characters and numbers are detected and recognized well very rare cases shows inaccuracy.

References

1. Bo Bai, Fei Yin, Cheng-Lin Liu “A Fast Stroke-Based Method for Text Detection in Vide” IEEE 10th IAPR International Workshop on Document Analysis Systems,march 2012,pp 69-73.
2. Khaled Alsabti,Sanjay Ranka,Vineet Singh “An Efficient K-Means Clustering Algorithm”
3. Minhua Li, Meng Bai “Text Segmentation from Image with Complex Background Based on Markov Random Fields” IEEE 2012 International Conference on Computer Science and Electronics Engineering,pp. 486-489.
4. Bharatratna P. Gaikwad ,Ramesh R. Manza, Ganesh R. Manza “Video Scene Segmentation to Separate Script” 3rd IEEE International Advance Computing Conference (IACC).March 2013,p.p.1269-1275.
5. Xian-Sheng Hua, Liu Wenyin,” Automatic Performance Evaluation for Video Text Detection,” 6th International Conference on Document Analysis and Recognition (ICDAR2001), September 10-13, (2001), pp545-550.
6. Xi Jie, Xian-Sheng Hua, Xiang-Rong Chen, Liu Wenyin, HongJiang Zhang “A Video Text Detection and Recognition System”, IEEE International conference 2001,
7. P. Shivakumara, W. Huang and C. L. Tan” Efficient Video Text Detection Using Edge Features”, 8th Workshop on Document Analysis Systems(IAPR) , pp 307-314(2008).
8. Chih-Chang Yu, and Hsu-Yung Cheng, Yen-Wen Chung “Detecting and retrieving texts from electronic marquees in natural scenes”, IEEE 17th International Symposium on Consumer Electronics (ISCE), pp.203-204.
9. Huang Xian-Xiang,Zhang Jin-Yu, Chen Yan“Edge Detection of Images Based on Improved Sobel Operator and Genetic Algorithms”_,IEEE international conference 2009. 7. N. A. Otsu, “A threshold selection method from gray-level histograms,” IEEE Transactions on Systems,, vol. 9, no. 1. pp. 62-66.
10. Zafarifar, J. Cao “Instantaneously responsive subtitle localization and classification for TV applications,” IEEE Transaction on Consumer Electronics, vol.57, pp.274-282, Feb. 2011.
11. N.Nathiya, K.Pradeepa “Optical Character Recognition for Scene Text Detection, Mining and Recognition”, IEEE International Conference on Computational Intelligence and Computing Research,2013.
12. Y.F.Pan,X.Hou and C.L.Liu,“A hybrid approach to detect and localize texts in natural scene images” IEEE Trans. on Image Proc, vol. 2, pp. 801–811.
13. David Crandall, Sameer Antani, RangacharKasturi “Extraction of special effects caption text events from digital video” International journal on document analysis and recognition, Volume: 5, pp:138- 157, 2003.

