Ote-Ocr Based Text Recognition and Extraction from Video Frames

Shashank Shetty[1] , Arun S Devadiga[2], *Department of Computer Science and Engineering*
[1]*Centre for Development of Advanced Computing (CDAC), Pune.*
[2]*NMAM Institute of Technology, Nitte, Karnataka*
[1]Shashankshetty06@gmail.com, [2]arun121cool@gmail.com

[1]S.Sibi Chakkaravarthy, 2K.A.Varun Kumar, *Department of Computer Science and Engineering*
[1]*Centre for Development of Advanced Computing (CDAC), Pune.*
[2]*Vel Tech Dr.RR & Dr.SR Technical University, Chennai.*
[1]Sbsibi.cse@gmail.com,varun.[2]kumar300@gmail.com

***Abstract* -The goal of this paper is to provide a new method to detect and recognize the text from the video frames. The task performed is divided into three step approach that combines the text detection and text recognition from the video frame. The video frame creation involves in dividing the video into an individual frames. The individual frame is grabbed and passed to the rest two phases. The text detection is a two-step approach, which involves text localization phase and the text verification phase. The text recognition involves in text verification phase and the optical character recognition phase. The final outcome of this paper is the detection of the text from the video frames in a word file. Experimental results demonstrating the proposed approach was also included, which shows the accuracy level of Optical character recognition(OCR) in terms of text extraction.**

Keywords: OCR; localization; Text Extraction; Text Verification; Text recognition; video frame.

I.INTRODUCTION

After the segmentations of the connected regions, the templates are compared with the segments. This is done to match the connected regions with the alphabets, numbers etc, to obtain the text. The text is got as the output from the O.C.R. The text got by comparison[8] is written to the text file likenotepad, wordpad etc. Hence the non-editable text, which was got from the text detection phase[9], is converted into the editable text by passing it to O.C.R.Fig.5 denotes the basic of OCR clearly which was citated in the section " appendix". OCR mainly deals with recognizing offline optical characters[6][7].Input sequence may be video or images which were scanned document or printed images.Major processing elements of OCR be denoted as [1]

Scanning[1]
Pre- processing[1]
Segmentation[1]
Feature extraction[1]
Recognition[1]

II.BASICS OF OCR

The algorithm mentioned above shows the various steps involved in obtaining a text using the Optical Character recognition (O.C.R)[4]. The block of text is obtained using the text detection process. These blocks of text is sent to the O.C.R, where these block of text is segmented into a single character and the templates are generated.

into the editable text by passing it to O.C.R.Fig.5 denotes the basic of OCR clearly which was citated in the section " appendix".

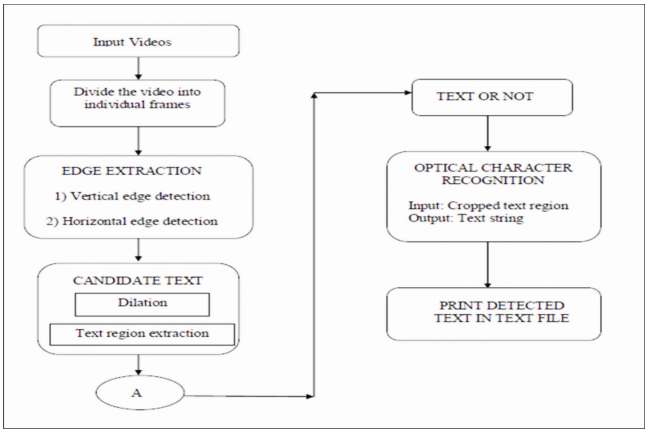

Fig.1.Optical Character Recognition

III.METHODOLOGY

978 -1- 4799 - 3837 - 7/14/$31.00© 2014 IEEE

The Algorithm mentioned above explains the various steps involved in the proposed system. Firstly, the Video is given as the input to the proposed system. The Input video is divided into individual frames and each individual frames are passed through the rest of the two phases and the individual frame represents the RGB image. The RGB image to gray scale conversion is done. The edge detection(i.e., horizontal and vertical edge detection) is done to the gray scale image using the Sobel and canny masks.Using edges as the prominent feature of our system gives the opportunity to detect characters with different fonts and colours since every character present strong edges, despite its font or colour, in order to be readable.Canny edge detector is applied to grayscale images. Canny uses Sobel[12][13][4][5] masks in order to find the edge magnitude of the image, in gray scale, and then uses non-Maxima suppression and hysteresis threshold. With these two post-processing operations Canny edge detector[4][5] manage to remove non-maxima pixels, preserving the connectivity of the contours.

Later, dilation on the resulted image is done to find the text like region. Dilation by a cross-shaped element is performed to connect the character contours of every text line. The common edge between the vertical and the horizontal edges is extracted and it is dilated again to get the accurate text like regions. The groove filling mechanism is applied to fill the gap between the non-connected pixels.

The co-ordinates of the dilated regions is sent to find whether the text like regions extracted is text or not. Once, the regions extracted are verified as the text, and then this detected text is passed to the optical character recognition (O.C.R). In the O.C.R, the block of detected text is segmented into characters and then the non-editable text is converted into editable text. The text is saved in the text file(i.e., notepad or WordPad).

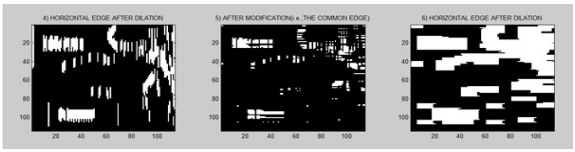

Fig.3a. Edge Detection before and after dilation

Pseudo code– OTE

```
Begin
Function OTE
        Input:Video file(.avi)
Output:Image/frames
        Step:Pre-processing
Convert RGB-→grey;
Sobel(horizontal,videoframes);
Canny(vertical,videoframes);
Plot(octagon);
Plot(rectangle);
Dilate(videoframes);
        for i=1:m
        for j=1:n
        FindText(min(n),max(m),min(m),max(n));
        End
        End
Joincharparts(FindText);
End
```

Pseudocodefortext verification is mentioned in section "appendix"

IV. Experiment And Results

Based on optical character recognition, our proposed method is used to recognize the text characters from the video streams. A video sequence of 12 frames is used as the data set model ,OTE is applied in order to retrieve the text from the video frames. The text recognized in the video frames was "Rhino is rare".Fig 3 denotes the output sequence of the OTE. Firstly the character recognized by OTE is considered to be the segmented frames of the video sequence. Various filters are used (which was denoted in the fig 4)in order to plot the diagonal mapping on the text characters.Full implementation is done through MATLAB [10][11]

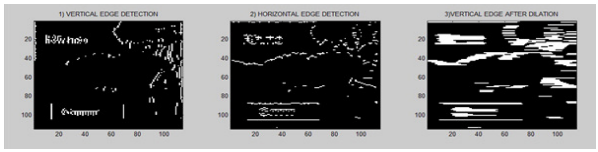

Fig.3b. Edge Detection before and after dilation

Various frames are consider, pre-processed with basic filters and processed using OTE methodology; its results are denoted at the final frame of fig 3a & b

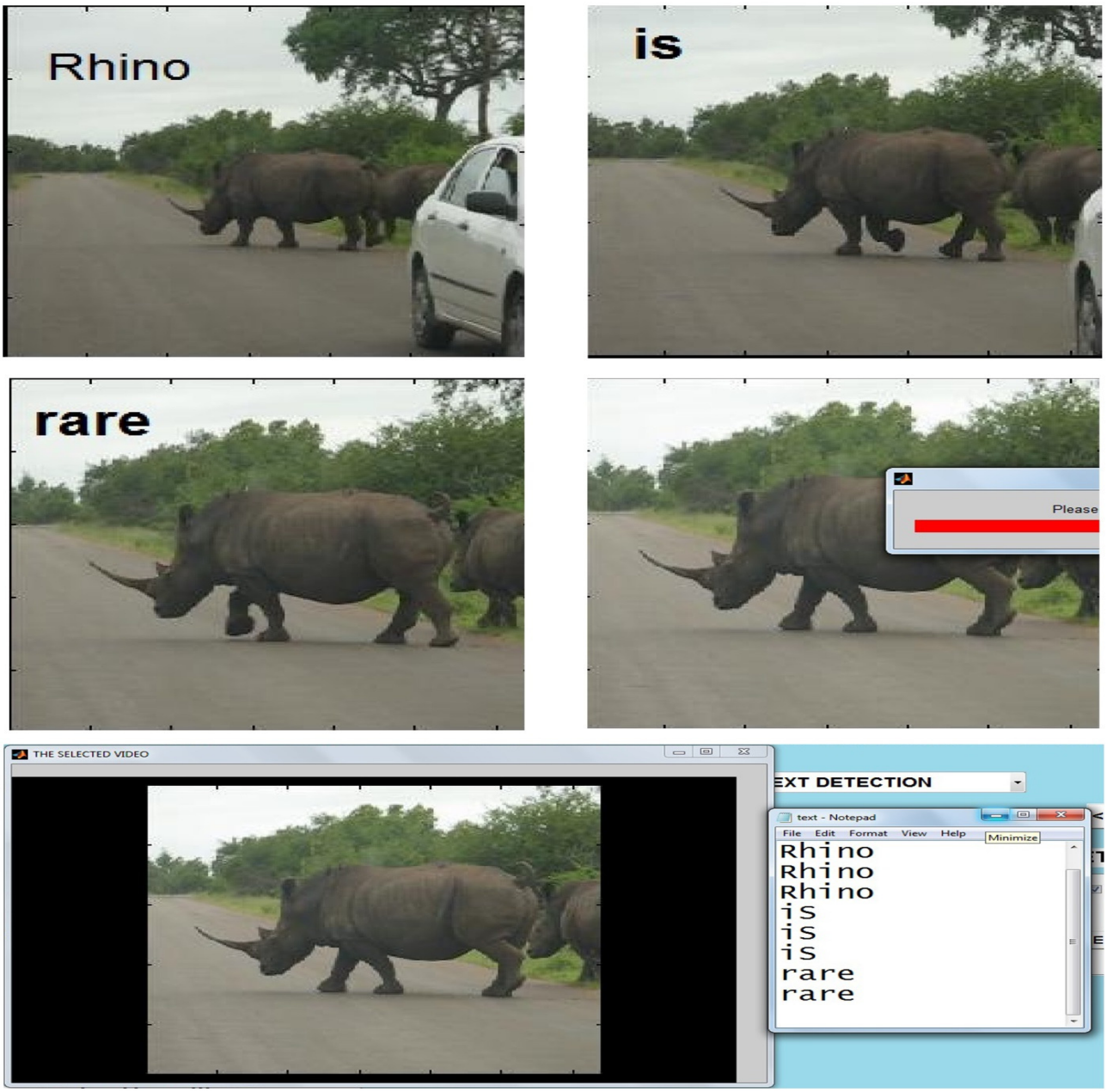

Fig.4.Ouput for the video sequence – OTE method

4. Ouput for the video sequence – OTE me acted from the video frames. In frame 1 the word ”Rhino” is extracted. In frame 2 the word “is” is extracted and in frame 3 the word “rare” is extracted,in frame four, all the extracted words from previous frames are processed and at final frame the words are recognized and displayed.

CONCLUSION

As the conclusion of this paper we experimentally performed some basic test in order to recognize the characters.Optical character recognition is playing a vital role in the field of image processing research and used in various applications.OCR processes mainly with segmentation and classification.The proposed method is evaluated experimentally with the video sequence and its results are shown clearly in fig 3 & fig 4.

TABLE.1.RESULT ANALYSIS

Number of video frames	12
Detection Percentage	97.08%
Extracted, Recognized Percentage	91.60%

Results shows that our proposed method has better performance in recognizing the characters in the video frames. Our proposed methodology has reduced noise with an accuracy of 97,08% .In future we extend this work with more samples and increased accuracy to recognize the text and its patterns in the video sequences.

V.REFERENCE

[1] PriyantoHidayatullah, NurjannahSyakrani, Ida Suhartini, WildanMuhlis;” Optical Character Recognition Improvement for License Plate Recognition in Indonesia”,IEEE symposium,UKSim-AMSS 6th European Modelling Symposium,2012.

[2] D. Chen, K. Shearer, H. Bourlard, Text enhancement with asymmetric filter for video OCR, in: Proceedings of the 11th International Conference on Image Analysis and Processing, 2001, pp. 192–198.Lukas Neumann and Jiri Matas, A method for text localization and recognition in real-world images.Center for Machine Perception, CzechTechnical University in Prague, Czech Republic, November 2010. 8-12.

[3] J.T. Tou and R.C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley PublishingCompany, Inc., Reading, Massachusetts, 1974M. Szmurlo, Masters Thesis, Oslo, May 1995.

[4] M. Swain, H. Ballard, Color indexing, Int. J. Comput. Vision7 (1991) 11–32.B. Manjunath, W. Ma, Texture features for browsing andretrieval of image data, IEEE Trans. Pattern Anal. Mach.Intell. 18 (8) (1996) 837–842.

[5] F. Mokhtarian, S. Abbasi, J. Kittler, Robust and eAcient shapeindexing through curvature scale space, in: British MachineVision Conference, 1996, pp. 9–12.

[6] Tollmar K. Yeh T. and Darrell T. “IDeixis - Image-Based Deixis for Finding Location-Based Information”, In Proceedings of Conference on Human Factors in ComputingSystems (CHI’04), pp.781-782, 2004.

[7] D. Chen, H. Bourlard, J.-P. Thiran, Text identi3cation incomplex background using SVM, in: International Conferenceon Computer Vision and Pattern Recognition, 2001,pp. 621–626.

[8] R.K. Srihari, Z. Zhang, A. Rao, Intelligent indexing andsemantic retrieval of multimodal documents, Inf. Retri. 2 (2/3)(2000) 245–275

[9] http:// www.mathworks.com/products/matlab/

[10] http://www.mathworks.in/

[11] http://www.intelligent-systems.info/classes/ee509/gui.htm

[12] http://www.ele.uri.edu/~hansenj/projects/ele585/OCR.

VI.APPENDIX

Pseudo Code-OTE Based Text Verification

Begin Function TextOrNot
Input: Four coordinates of the dilated regions (X,X1,Y,Y1) and the edge map image (H)
Output: Text Or Not
for i= X:X1 for j=Y:Y1
hcount=hcount+1
end
end
*tcount=(X1-X) * (Y1-Y)*
Ratio=hcount/tcount
If (Ratio >=0.065)
Result= TRUE (Its Text)
end
end

Basics of OCR

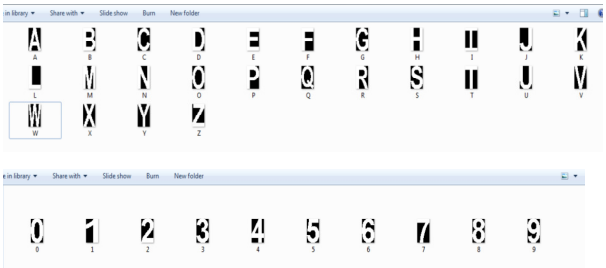

Fig.5. Template Files of Alphabets and numbers