



Character Recognition in Natural Scene Images

Nikita Tiwari, Vasundhara Rodge, Komal Zagade

Department of E&TC, BVCOEW,
Pune, Maharashtra,
India

Abstract- *The content based image analysis is done by the text detection and localization of natural scene images. This approach is for resilient text detection which localizes texts in natural scene images. This system is essentially divided into three region 1) Pre-processing 2) Text Localization 3) Optical Character Recognition. From preprocessing step, get grayscale image and resize them. After that recognized text is localised to find confident text region. Optical character recognition technique is used for character recognition. Then compare the resulting data with the records on database to get the information about whatever the data is written. Matlab is used for implementing and simulating the system. The developed system successfully detects and recognise the data on natural scene images is observed from the experiment.*

Keywords: DCT, OCR,

I. INTRODUCTION

Character recognition in natural scene images is the challenging task because of the need to find out where exactly the text is located and how much of the scene is not relevant to character recognition in addition to that there are different font styles such as thick-thin ,outlines-solid lines, colours and textures etc,furthermore variations in lightning conditions. Though its challenging but widely used task. In today's digitized world we are using the digital cameras, mobile phones for capturing the images therefore in coming years content based image analysis techniques are receiving more attention. It finds wide applications in translation, content based web image search and mobile text recognition.

In this system text recognition and detection is done by using three steps pre-processing, text localization and opticalcharacter recognition. In 1997 V.Wu,R.Manmath and E.cRiseman used the technique 'Finding text in images', in which text pixels are grouped by morphological operators into text regions however they failed to remove non text region. In 'Automatic caption localization in compressed video' technique given by researchers Y. Zhong H.J.Zhang and A.K.Jain,described the wavelet decomposition and Discrete cosine transforms (DCT) for feature extraction but this approach will not work efficiently ,there is no utilisation of color information. H. P. Li, D.Doermann, and O. Kia projected an algorithm for detecting texts in video by using first and second order moments of wavelet decomposition responses as local region features classified by a neural+++l network Classifier in their technique 'Automatic text detection and tracking in digital video', but Poor OCR performance semantic in dexing need extensive training. 'Comprehensive method for multilingual video text detection, localization, and extraction'techniqueproposed by M. Lyu, J.Q. Song, and M. Cai, described that the candidate text edges of various scales with a Sobel operator can be detected. Non text edges are filtered using local thresholding method , and the text regions are then grouped into text lines by recursive profile projection testing only focuses on English and Chinese language, but it cannot detect motion texts due to the assumption of stationary text and no horizontally aligned texts cannot be localized. Researchers R. Lienhart and A.Wernicke in 2002 calculates the gradient map with color derivative operators by the procedure 'Localizing and segmenting text in images and videos', but confined to a small area text lines are then scaled to a fixed height of 100 pixel.

Our approach is to make the system more resilient by finding confident region in localization step. This paper is organized in the following steps 1) Pre-processing 2) Text Localization 3) Optical Character Recognition.

A. Related work

Below flowchart shows the steps of system. Mainly system is divided into three parts as preprocessing, text localization and optical character recognition. In pre-processing step the natural scene image which is content based will get converted into the grayscale image as result. This image will get ahead to the text localization ,here we will get the confident region from the image consisting of text. Line cropping takes place in text localization step,in the line cropping we are separating the lines and then the output of this step is pass to the OCR , after that letter cropping occurs and the characters will get recognized by the OCR

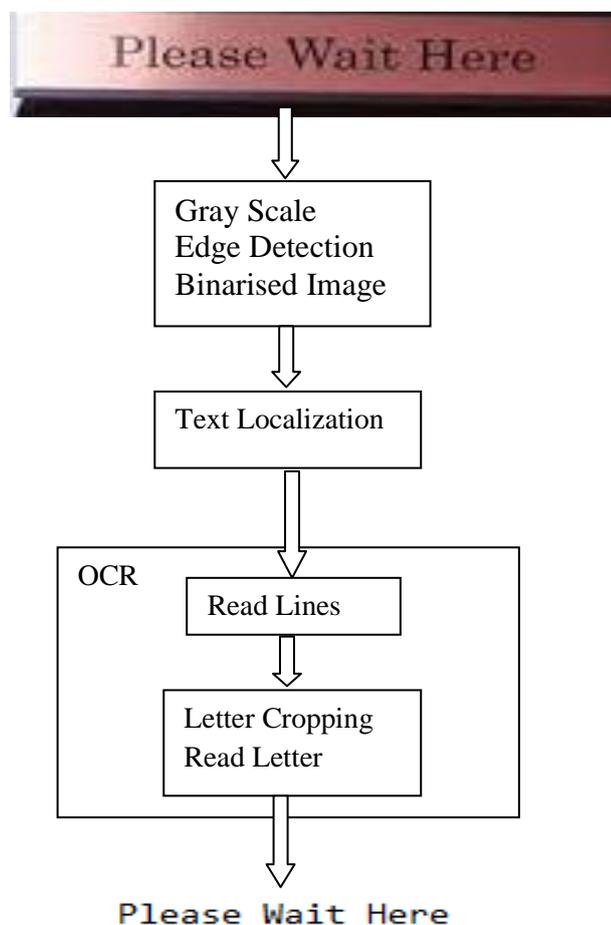


Fig: Flow chart of system

B. Algorithm used for implementation

- Input image:-take natural scene content based image as input image as stored in database or capture image from the camera provided.
- Conversion:-Convert the given rgb image into gray scale image. We have to take the rgb values for each pixels and make it output a single value which reflects the brightness of that pixel for converting rgb image into grayscale. If the input image is already in grayscale leave it as such.
- Finding confident text region :- The image will pass to the text localization ,here we will get the confident region from the image consisting of text. Line cropping takes place in text localization step. Text region can be find by using canny edge detection and dilation.Canny edge detection algorithm is the optimal edge detector. We are using canny edge detector because of its low error rate, the edge points be well localized and to have only one response to a single edge.

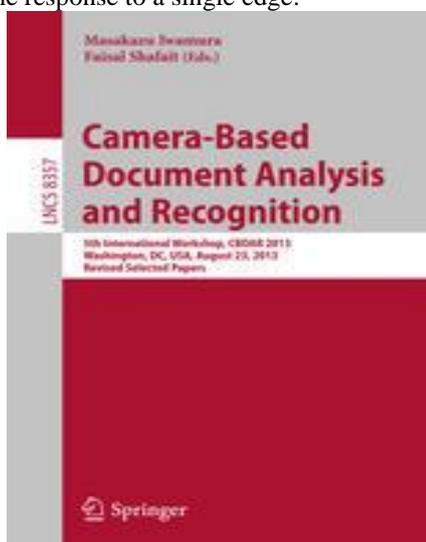


Fig:-Original image

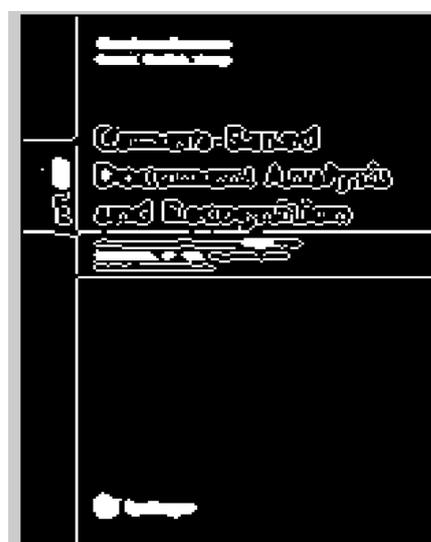


Fig:-Image After localisation
first smoothes the image to eliminate noise.

Based on these criteria, the canny edge detection

- Finding the edge direction is trivial once the gradient in the x and y directions are known. However, you will generate an error whenever sumX is equal to zero. The formula for finding edge direction is just :

$$\theta = \tan^{-1}(Gy/Gx)$$

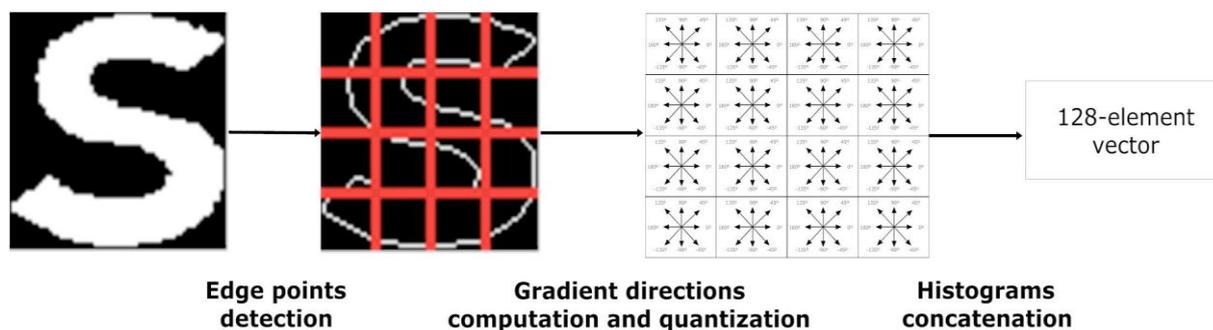


Fig:-Feature Detection

- Dilation is one of the two basic operators in the area of mathematical morphology. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of a foreground pixel (i.e. white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller.
- The next step is cropping text region.
- After that line cropping is done by vertical projections of multiple lines there in given image.
- Followed by line cropping word is cropped. Words are separated by space vector.
- The next step is character recognition by template matching and correlation.

II. CONCLUSION

In resilient approach major focus is on the localization and text detection part. In this paper we have produced a text detection and recognition system applied it to images of text in natural scenes. We demonstrated that with larger banks of features we are able to achieve increasing accuracy with top performance comparable to other systems. Our results point out that it may be possible to achieve high performance using sophisticated feature learning algorithms currently being developed by machine learning researchers.

REFERENCES

- [1] Muhammad Tahir Qadri, Muhammad Asif, "AUTOMATIC NUMBER PLATE RECOGNITION SYSTEM FOR VEHICLE IDENTIFICATION USING OPTICAL CHARACTER RECOGNITION," 2009 *International Conference on Education Technology and Computer*.
- [2] Alvaro Gonz'alez, Luis M. Bergasa, J. Javier Yebes, Sebasti'an Bronte Dept. of Electronics, University of Alcal'a, Spain" A CHARACTER RECOGNITION METHOD IN NATURAL SCENE IMAGES," 21st *International Conference on Pattern Recognition (ICPR 2012) November 11-15, 2012*.
- [3] Chirag Patel, Atul Patel, Dharmendra Patel," OPTICAL CHARACTER RECOGNITION BY OPEN SOURCE OCR TOOL TESSERACT," *International Journal of Computer Applications (0975 – 8887) Volume 55– No.10, October 2012*.
- [4] Y. Zhong, H.J.Zhang, A.K.Jain," AUTOMATIC CAPTION LOCALIZATION IN COMPRESSED VIDEO" *IEEE Transactions on Pattern Analysis and Machine Intelligence* 4 April 2000.
- [5] Kavita Ahuja, Preeti Tuli," OBJECT RECOGNITION BY TEMPLATE MATCHING USING CORRELATIONS AND PHASE ANGLE METHOD" *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 2, Issue 3, March 2013.
- [6] Victor Wu, R.Manmatha, Edward M.Riseman," FINDING TEXT IN IMAGES" *Indian Journal of Applied Research*, volume 3, Dec 13.
- [7] R. Lienhart and A.Wernicke," LOCALIZING AND SEGMENTING TEXT IN IMAGES AND VIDEOS" *Circuits and Systems for Video Technology, IEEE Transactions* 2002.