

TEXT DETECTION AND RECOGNITION FROM IMAGES

NAGASUDHAD¹, DR.MADHAVEELATHA Y²

1Research Scholar ,Dept of ECE JNTUCEH

2Dept of ECE MRECW Hyderabad

Email-Id:nayaniharshi@gmail.com

Abstract: *Text detection from images is important for the content based image analysis. To understand the contents of an image or valuable information, it is necessary to analyze the text appearing on it. Various methods have been proposed over past years for text detection and extraction from different types of images like born digital image, scene image and document image. Text detection and extraction is hard due to the intricate background, different alignment of text size, orientation in images. In this paper we propose Wavelet with Canny edge detection, and connect component hieristic filter are used for the detection of text and character recognition form natural scene images and document images.*

Key words: Text detection, character recognition, text extraction, scene image, document image.

I. Introduction

With the advances in multimedia technology, content based indexing and text extraction in images plays a major role. It contains different contents in it such as caption, text, scene etc. Among all the contents of the image text is found to be one of the most important features to understand the image contents. Text in images can be

used as indexing purpose, document processing [1,2], video content summary [3-5], video retrieval[6]and video understanding[7].

The text information can be extracted in two stages: Text detection and text recognition. Text detection detects the text regions as external regions of an image and text recognition retrieves the text information from these external regions[8]. Text extraction from images have many useful applications in document analysis, detection of vehicle license plate , diagrams, maps, charts etc. Key word based image searching, content based retrieval , name plates street signs text based video indexing and document retrieving etc.

Images can be broadly classified into document images, caption text images and Scene text images.

A document image usually contain text and few graphical components. These are acquired by scanning journal, printed document, degraded document images, hand writted document and book cover etc.

Caption text is also known as overlay text which is artificially superimposed on the video/image at the time of editing and it describes the content of the image/video .

Scene text appears within the scene which is then captured by the recording device. i.e. text is present in the scene when the image or video is shot.

Retrieving the contents from images is very challenging because when extracting text with variation in fonts, size, color, alignment, orientation, illumination and back ground. Problem of text extraction is very difficult because of these deviations.

Some of the problems of text extraction process are listed below.

Font, color and size: The size, font and color of the text can vary in different images. This makes the application of many text extraction methods. But the existing methods can deal with limited range of font styles and sizes.

Complex backgrounds: The complexity of background may vary from simple to much complex ones. The back ground consists of varying colors and textures.

Noise, blur and compression: Low resolution images mostly suffer from blur and loose sharp transitions at the text boundaries. GIF that is limited to 8 –color palette makes considerable quantization artifacts and dithered color. Compressed artifacts also degrade the quality of edges. Compression of images makes significance distortion to characters and their boundaries.

II Architecture of text extraction and recognition

Text detection and recognition process comprises of five steps namely text detection, text localization, text tracking, segmentation or binarization and character recognition.

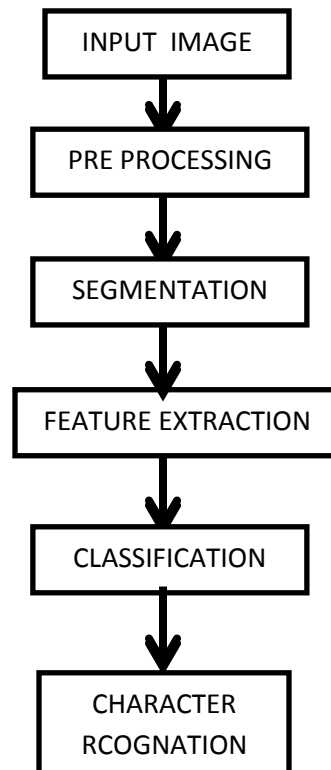
Text detection: This phase takes image or video as input and decides whether it contains text or not. It also identifies the text regions in images.

Text localization: It merges the text regions to formulate the text objects. It also generates the tight bounding boxes around the text objects.

Text tracking : Text tracking is performed to reduce the processing time for text localization and maintain the integrity of position across adjacent video frames. Although the location of text in image is identified by bounding boxes, the text still needs to be segmented from background to facilitate its recognition.

Text segmentation: In which the text components are segmented from the background. The output is text binarization is the binary image, where text pixels and background pixels appear in two different binary levels.

Character recognition: The text extracted can be transformed into plain text using OCR technology.



Architecture of text extraction and recognition

III Related Work:

For text detection in images a large number of approaches have been developed, which are classified into two categories: Region based approaches and texture based approaches.

Region based approaches: Region based methods use the properties of the color or gray scale in a text region or their differences with the corresponding properties of the background. This method uses a bottom up approach by grouping small components into larger components

until all regions are identified in the image. A geometric analysis is needed to merge text components using spatial arrangement of the components so as to filter out non-text components and mark the boundaries of the text regions.

Shivakumara et al[9] proposed an edge based technique for text detection in images with text present in the horizontal direction. The frame was segmented into 16 non-overlapping blocks. Mean and median filter and edge analysis was used to identify the candidate text blocks. Using block growing method, the complete text block

was obtained. Finally based on the vertical and horizontal bar feature, the true text regions are detected. Inshivakumara et al [10], filters and edge analysis were used for initial text detection. The straightness and curviness edge features were used for false positive elimination.

Text detection using a cascade Adaboost classifier with HOG and multi scale local binary pattern feature was proposed by Pan et al [10]. Text localization was done using window grouping technique. Within each located text line, local binarization is done to extract candidate CCs and non text CCs are filtered using Markov Random field model and MLP in order to get the final text line.

Texture based approaches: Texture based approaches use distinctive properties of the text that separate them from the background. The techniques based on Gabor filters, Wavelet, FFT, spatial variance etc. can be used to detect the texture properties of a text region in an image. These approaches are used in complex background and expensive.

Connected component based approaches are fast and good for images which have high contrast texts and plain background just like methods in the document image analysis. To improve the performance of text detection mixture of texture based and region based approaches are used.

Zhaon et al. [11] used wavelet transform and sparse representation with discriminative dictionaries for text detection. Shivakumara et al [12,13] also used Haar wavelet in both

the works. In [14] they also used color features along with Wavelet-Laplacian method to detect text. Shivakumara et al [15] used wavelet median moment feature with k means clustering to obtain text pixels.

The wavelet transform is used in different fields such as image processing, signal processing, video compression and so on. In the frequency domain an image is decomposed into different components and computed with low pass and high pass filters and by one dimensional discrete waveform transform (DWT) an image is divided into two parts that is coarse and detailed information. The texture features are extracted from discrete wavelet transform.

Proposed methodology: In this method we propose a method that makes use of Wavelet with Canny edge detection and connected component heuristic filtering method are used to extract the information from images.

The first step is to take an input image and convert into gray scale image. Then filter is used to eliminate the noise in the original image before the detection of edges. Wiener filter is used for smoothing process. After the pre-processing the next step is to detect the edges in the image. Canny edge detector is applied for edge detection.

Once the edge points are detected, links between these edges are identified. This process is called edge linking. Connected components of the edges detected and edge linking and merging over then Discrete wavelet transform and connected component

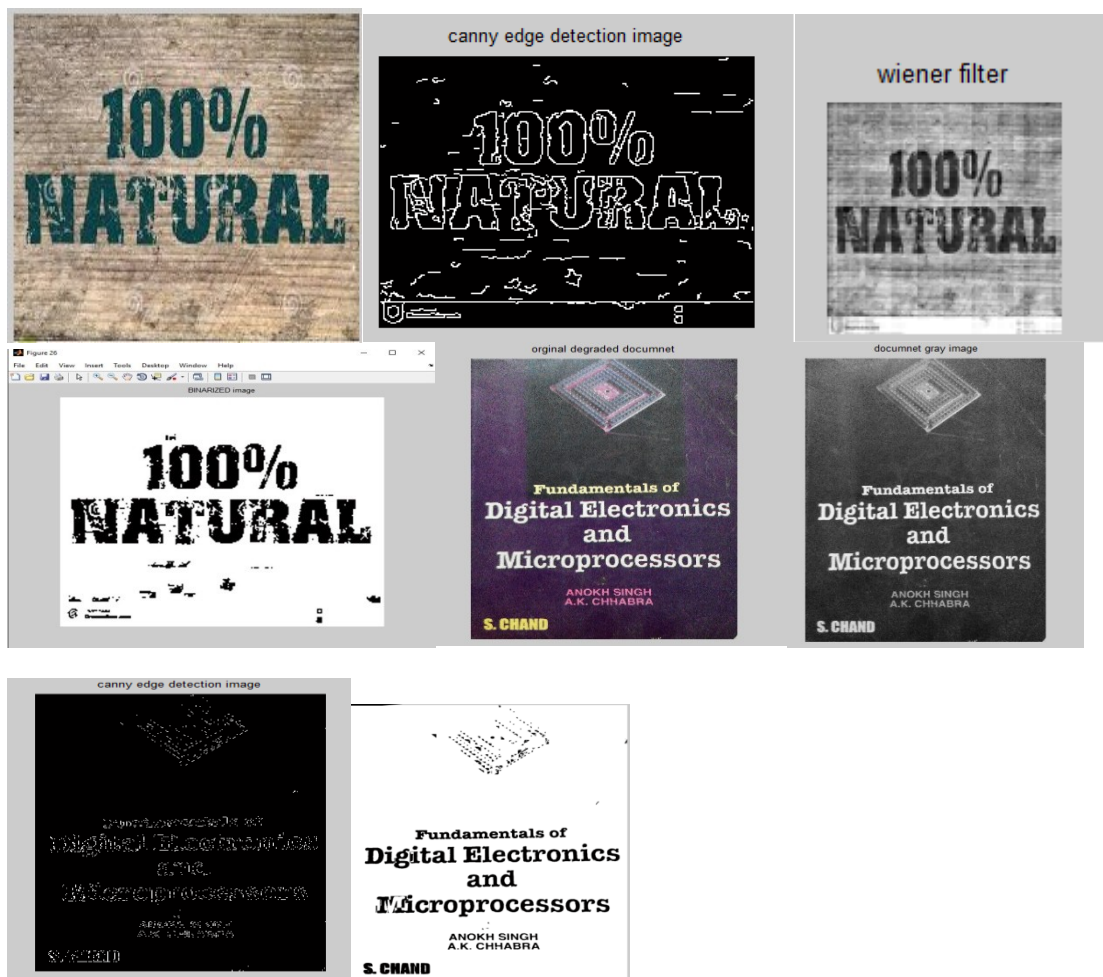
heuristic filtering is used to remove non text regions from the image.

IV. RESULT AND DISCUSSION

We evaluated this proposed approach on the images for near about 20 scene images and

about 20 images for mixed document images. All the images of the natural scene and document images were captured using a digital camera under different luminosity conditions. It is evident from the results that the performance of the proposed algorithm is satisfactory on both types of images .

Input image



Output image

V Conclusion: A robust and effective algorithm for separating text from mixed documents image and scene image has been

presented. The algorithm accepts the text in various font sizes. Even low quality printings where characters are split into

several connected components can be analyzed. The main component of the proposed system is to detect text by using edge information and connected components with heuristic filtering to extract text data from images. The algorithm has been validated on scene images and mixed document images.

References:[1] Jain,K., and Yu, B., “document representation and its application to page decomposition,”IEEE Trans Pattern And Machine Intell.,vol.20 pp. 294-308,Mar. 1998.

[2] Jain, K., and Yu, B., “Automatic Text Location in Images and Video Frames”,Pattern recognition,31(12) 2055-2076,1998.

[3] Jain, K., and Zhong, Y., “Page Segmentation using Texture Analysis”,Pattern Recognition , 29 (5) 743-770, 1996.

[4]ChitrakalaGopalan and D.Manjula,”Contourlet Based Approach for Text Identification and Extraction from Heterogeneous Textural images”, International Journal of Computer Science and Engineering 2(4) pp.202-211,2008

[5]Crandali, D., Antani, S., and Kasturi, R. “Robust Detection of Stylized Text Events in Digital Video”, Proceedings of International Conference on Document Analysis Recognition,pp. 865-869,2001.

[6]Serra,J.,”Image Analysis and Mathematical Morphology”New York,Academic,1982

[7] Fa,K.C., Wang,L.S and Wang, Y.K.,”Page segmentation and identification for intelligent signal Processing “ Signal Process.,vol,45,pp.329-346,1995.

[8]Kim,H.K.,”Efficient Automatic Text Location Method and content Based Indexing and structing of video Database”, Journal of Video database”, Journal of visual Communication and image Representation 7(4) 336-344,1996.

[9]pShivakumara, W Huang, C.L Tan, “Video text detection based on filters and edge features”ICME,2009,pp 514-517

[10]Yi-Feng Pan,XHou, and C.L Liu, “A robust system to detect and localize texts in Natural Scene images,”DAS,2008 pp35-42

[11]M.Zhao,S.Li and J Kwok,”Text detection in images using sparse representation with discriminative dictionaries”, Image and Vision Computing vol28,2010,pp 1590-1599.

[12]P.shivakumara, T.Q Phan, and C.L Tan ,”A Robust Wavelet transform based technique for video text detection”ICDAR,2009,pp1285-1289.

[13]P.Shivakumara, T.Q Phan,C.L Tan “New Wavelet and color features for text detection in Video”,ICPR,2010,pp 3996-3999.

[14]PShivakumara,ADutta,C.LTan,U.Pal,”A new Wavelet Median –Moment based method for Multi-oriented Video Text Detection,”DAS,2010,pp 279-286.