# Elimination of non-Text region by Text Localization in an Image Using Wavelet Transform

Shireen K. Hussain[1], Dr. Sachin R. Gengaje[2]
[1]*Master of Engineering student, Walchand Institute of Technology, Solapur*
[2]*Ph.D. Professor, Walchand Institute of Technology, Solapur*

***Abstract-*** **Natural scenes images contain different objects. These images can be photos taken by people which are various and used in different situations. For example, scenery photos and portrait photos are commonly seen when we are in journey and daily life. However, to many people, text is the most intuitive way to get the around information. These images can also be disclaimers on the roads, construction sites or notices of hazards which are the text regions. These images might have text as well as other objects which are non text region. The aim of this paper is to focus only on text region and eliminate non text regions. To achieve this aim Text detection and localization is implemented along with eliminating non text region. In this paper, an efficient algorithm which can detect, localize and extract text in images with complex backgrounds is presented. The proposed approach is based on a method for edge detection, and the localization of text regions using projection profile analysis and geometrical properties. The present research aims at developing a text detection system from an image using edge based or texture based approach. The approach is targeted towards being robust with respect to different kinds of text appearances, including font size, color and language. The proposed method will be applied on the various images and detect regions of the text perfectly and removes the non text regions. The standard images available on the internet and used by the researchers will be used for validation of the proposed system. The entire framework is developed on MATLAB platform.**

***Index Terms-*** **Discrete Wavelet Transform, Discrete Wavelet Transform (DWT), Discrete Cosine Transform(DCT), Connected Component (CC).**

## 1. INTRODUCTION

Text detection and localization is the process of determining text locations in the image. Cameras mounted on various hand-held devices have become very popular. And natural scene images usually taken by digital cameras are focused on more and more in computer vision. These photos taken by people are various and used in different situations.

For example, scenery photos and portrait photos are commonly seen when we are in journey and daily life. However, to many people, text is the most intuitive way to get the around information. In consequence, the extraction of text in natural scenes is one of these important technologies in computer vision. Natural scenes images contain different objects. Extracting text information from natural scene images has many challenging issues. The challenges lie within various factors, such as the variation of the light intensity, alignment of text, color, font size, and camera angles. With more and more digital devices, images have now become the most popular media type in our daily life. Since text embedded in video contains much semantic information related the image content, it plays an important role in content-based multimedia indexing and retrieval systems. Many text detection approaches have been proposed in the past several years; however, due to low resolution and complex backgrounds of images and various sizes, colors, styles and alignments of text, text detection and extraction is still challenging. Text localization and recognition in real-world (scene) images is an open problem which has been receiving significant attention since it is a critical component in a number of computer vision applications like searching images by their textual content, reading labels on businesses in map applications (e.g. Google Street View) or assisting visually impaired [2]. Several approaches for text detection in images have been proposed in the past. Based on the methods being used to localize text regions, these approaches can be categorized into two main classes: connected component based methods and texture based methods [12]. A text images can be classified into three types namely document image, scene text image, and caption image. Document image is acquired by scanning book covers, printed document etc. Scene text image sometime is referred to graphics text and it finds in natural images that contain advertisements such as street name, road signs etc. Caption image contains text which is inserted in this image. Caption text is always referred as artificial text. Many applications need text localization and segmentation from natural scene images, ranging from automatic detection of traffic signs that help in transportation system, and helping visually impaired people, to multimedia indexing and retrieval. Text localization methods have been classified as Region based methods and Texture based methods. These methods work based on color differences between text regions and their background. Region based methods are divided in two sub methods: connected component based method (CC), and edge detection based method. Region based methods work based on color differences between text regions and their background. Region based methods are divided in two sub methods: connected component based method (CC), and edge detection based method. In first method, a text is considered as a set of distinct connected components which have their specific intensity and color distributions. The second method depends on finding maximum intensity changes between text and background. The main differences between two methods are edge based method is useful to process low contrast text image and with different text size, while connected components methods are simpler to implement. The main drawback of CC method has failed in locating text regions in images which have complex background. While texture based methods work by extracting texture features of image firstly, and then a classification process is applied in the second stage to detecting text regions. Discrete transformations such as discrete wavelet transform (DWT) and discrete cosine transform (DCT) are used. These methods suffer from high complexity in nature, but it is robust in processing complex background.

This paper gives a brief review of the recent related works. It presents the proposed method. Also illustrates the experimental results.

## II. PROCEDURE FOR PROPOSED SYSTEM

In the proposed project, there are two processes used, edge based algorithm and connected components algorithm which are combined to get a robust technique.

In this section, a solution for detecting text in an image is provided. As presented in Fig.1 the proposed text detection algorithm passes through the following steps.

1) Preprocessing.
2) Text Localization.
3) Text detection.

After performing the last step of text detection, the non text image will be eliminated using thresholding techniques and morphological operation.

1) *Preprocessing*

The input images used are having text of different font, color and language. Even the size of input image may vary. So all the input image is resized to 256x256 as shown in Fig.1 which is the input image.

Further, if the input image is RGB, it must be converted into YUV color space by forming weighted sum of the R, G, and B components as in Equation 1.

$$Y = 0{:}299*R+0{:}587*G+0{:}114*B$$
$$U = 0{:}14713*R-0{:}28886*G+0{:}436*B \qquad (1)$$
$$V = 0{:}615*R-0{:}51499*G-0{:}10001*B$$

Our method uses only luminance components (Y) for next processing steps.


Fig.1 Preprocessing (a).Input image (b).Y image (c).U image (d).V image

The Fig.3 shows the YUV image which is the result of Equation 1. Further only (a).Y image will be used in the next step.

2)Text Localization

Detection of text is done using Discrete Wavelet transform. One-dimensional DWT is applied on an image, an image is decomposed into two parts which is coarse and detail elements by using low pass (L) and high-pass (H) filters. Image signal is decomposed into four sub bands when two dimensional DWT is applied on rows at first and then on columns of an image. These sub bands are approximation sub-band (LL), and details sub bands (LH, HL, and HH).

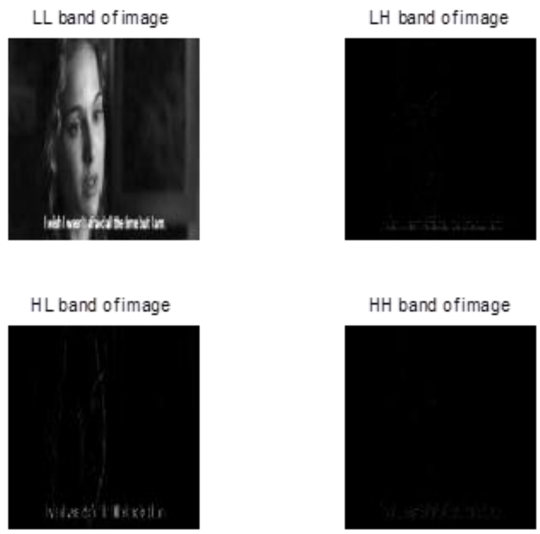


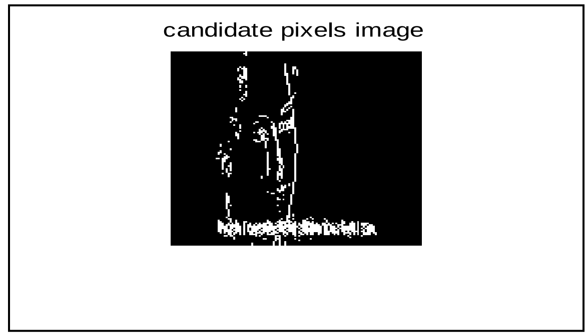

Fig.2. Single level decomposition of DWT.

The above Fig.2. shows the Single level decomposition of DWT. The text pixels have high variation around its neighbor pixels, therefore a technique based on edge detection will be applied. The edges found in high frequencies therefore is used in the next step for detection of text region. A wavelet edge pixel E(i,j) at pixel in location i and j can be defined by taking the average of corresponding pixels in three high frequency sub bands (LH, HL and HH) which is shown in Fig.3.(a)

$$\sum_{k=1}^{3}((Dk(i,j)^2)\ \ /3$$

Then each pixel will be a candidate text pixel if its value is larger than threshold α, which is a parameter and its value is selected by local statistics such as mean and standard deviation of an image. The value of α=43.89 for image in Fig.3.(b) shows candidate text pixel detection.


(a)




(b)

Fig.3. (a) Edge pixels detection of image. (b) Candidate pixels of image.

3) *Text Localization:*

Vertical and horizontal projections are one of the amplitude segmentation methods. They convert image contents into one-dimensional representations. They computed parallel to the coordinate axis, so, they are useful methods. Vertical and horizontal projections of a binary candidate text pixels image Cp(i,j), are performed for horizontal projection Phor of row as the sum of pixel values in that row and all the columns in the image. While vertical projection Pver of column is computed as the sum of pixel values in that column and all the rows in the image.

To segment text regions, two thresholds are used horizontal and vertical thresholds, which denoted as Th and Tv respectively. They are defined as:

Th=(mean(Phor)+min(Phor))/2

Tv=(mean(Pver)+min(Pver))/2

If Phor is greater than horizontal threshold, then Phor of that row can be considered as a part of a candidate text region; otherwise this row is suppressed. Similarly, if Pver is greater than vertical threshold, then Pver of that column can be considered as a part of a candidate text region. Fig.7. shows projection of candidate text image.

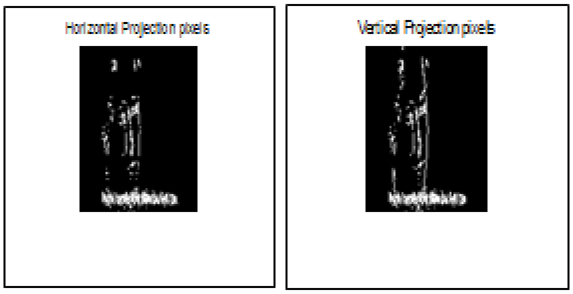
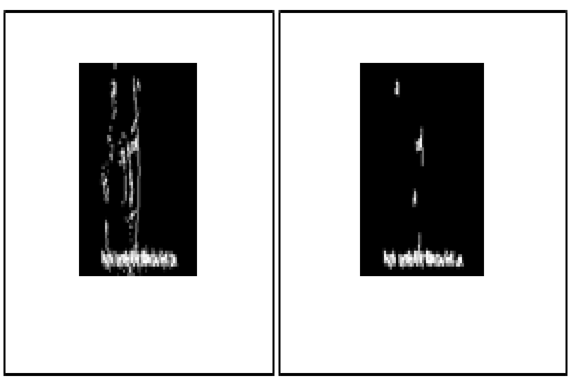



Fig.4. Projection of candidate text image (a) horizontal projection (b) horizontal projection

The next step perfomed is connected components. In this step, horizontal and vertical images are combined to be one image, which is shown in Fig.8-(a). And then non-text regions which are false detected are eliminated using geometrical features of regions. To get image regions a 4-connected component is applied on the combined projection image. Area geometric feature of each region is computed and compared with threshold (thr1) which computed according to Equation (2).

thr1=MaxArea/10 (2)

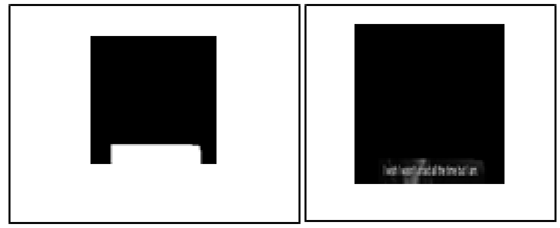
where Maxarea represents a largest region area.If the area of each region is less than threshold, then it is considered as non-text region and its pixels should be discarded. Fig.5 shows the result for this connected components step.



(a) (b)

Fig.5. Connected components step (a) combined projection image (b) result image after discard non-text regions

As shown in Fig.5-(b), the resultant image may contain non-text pixels. So a morphology operation are performed where dilation followed by opening operation are applied to further remove nontext pixels. Fig. 6 shows the detection of text regions and eliminated non text region after applying morphology operations.



(a) (b)

Fig.6.(A)Eliminated Nontext Region.(B)Text Region Extracted

## III. RESULT

Several experiment tests were carried out on text images. Every color image must be converted into gray scale image before applying text localization method. Figures shown below are several performed experiments. The result shown are in gray scale.


## IV.OBJECTIVE

The main objective of this project is to achieve this suitable system which can detect, localize and extract text in images. The main objective can be achieved by Studying and analyzing various methods for the steps of pre-processing of an image, analyzing complex backgrounds, text detection and non text detection pixels, and removing the non text regions.

## V.CONCLUSION

The proposed system is validated using set of images available using suitable performance metrics. The proposed method applied on the various images and detect regions of the text perfectly removes the non text regions.. The proposed method is more robust in dealing with the contrast, complex background images and images of different fonts and size of text. Using wavelet transform gives good tools for detecting candidate text regions.

## REFERENCES


[1] Shivananda V. Seeri, J.D. Pujari, P.S. Hiremath, Multilingual text localization in natural scene images using wavelet based edge features and fuzzy classification, Int. J. Emerg. Trends Technol. Comput. Sci. 4 (1) (January-February 2015) 210e218.

[2] T. Kumuda, L. Basavaraj, Novel technique for text detection and localization in natural scene images, Int. J. Eng. Res. Technol. (IJERT) 3 (4) (April e 2014) 2675e2680.

[3] Xiaoqian Liu, Ke lu, Weiqiang Wang, Effectively localize text in natural scene images, in: 21st International Conference on Pattern Recognition, November 11e15, 2012, pp. 1197e1200. Tsukuba, Japan.

[4] L. Neumann, J. Matas, Scene text localization and recognition with oriented stroke detection, in: IEEE International Conference on Computer Vision (ICCV), 2013, pp. 97e104.

[5] Keechul Jung, Kwang In Kim, Anil K. Jain, Text information extraction in images and video: a survey, Pattern Recognit. 37 (5) (2004) 977e997.

[6] Wen-Nung Lie, Guo-Shiang Lin, Sheng-Lung Cheng, "Dual Protection Of JPEG Image Based On Informed Embedding And Two Stage Watermark Extraction Techniques", IEEE Transactions on information forensics and security, Vol. I, 2006, pp. 330-341.

[7] Biao-bing-huang and Shao-ziantang,"A Contrast Sensitive Visible Watermarking Scheme", IEEE feature article, 2006, pp. 60-66.

[8] Enpingli, Huaqing Liang, Xinxin Niu, "Image Watermarking Scheme Based On Wavelet Tree Quantization Robust To Geometric Attacks", IEEE Proceedings of Sixth World Congress on Intelligent Control and Automation, pp. 10256-10260.

[9] Der-Chyuanlou, Hao-kuan Tso, Jiang-lungliu, "A Copyright Protection Scheme for Digital Images Using Visual Cryptography Technique", Computer Standards and Interfaces, 2006, pp. 125-131.

[10] Zhiquan Lu and Xiao-Ping Zhang,"Robust Image Watermarking Based on the Wavelet Contour Detection", IEEE Transactions, Vol.2, 2005, pp.1165-1168.

[11] Chip Hong Chang and Mingyan Zhang, "Fuzzy-Art based adaptive digital watermarking scheme", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 15 no. 1, 2005, pp. 65-81.