Available online at www.sciencedirect.com

ScienceDirect

Karbala International Journal of Modern Science 2 (2016) 138–144
http://www.journals.elsevier.com/karbala-international-journal-of-modern-science/

Wavelet transform based technique for text image localization

Suhad A. Ali [a,*], Ashwaq T. Hashim [b]

[a] *Babylon University, Science College for Women, Computer Science Dep., Babylon, Iraq*
[b] *Technology University, Control and Systems Engineering Dep., Baghdad, Iraq*

Received 10 January 2016; revised 8 March 2016; accepted 21 March 2016
Available online 3 May 2016

Abstract

In this paper, a robust technique based on discrete wavelet transform, edge detection, and morphology operation for scene text detection is proposed. There are several stages in the proposed method. In the first stage, a single wavelet decomposition LH, HL and HH subbands are applied for detecting edges in original scene text image. The projection technique is applied in the second stage to preliminary detect text and non-text pixels. In third stage, 4-connected components are applied, and then area geometric feature is used as threshold to remove non-text region. At last stage, morphological operations are applied to connect isolated text components and to remove non-text regions. The proposed method is applied on a various images such as images of low contrast, complex background images and images of different fonts and size of text. The experimental results show that the proposed method can detect regions of the text perfectly.

Keywords: DWT; Image projection; Morphological operator; Connected components

1. Introduction

Natural scenes images contain different objects. A text object is important one among these objects, because it shows the important meanings (semantic) of image. 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 inserted in this image. Caption text is always referred as artificial text.

Many applications need text localization and segmentation from natural scene images [1]; ranging from automatic detection of traffic signs that help in transportation system [2], and helping visually impaired people [3], to multimedia indexing and retrieval [4]. Text localization and recognition problem has been recently receiving significant attention because text localization method achieved a localization recall of (62%) [4]. Text localization methods have been classified into two classes [5]:

1) **Region based methods**: These methods work based on color differences between text regions and their background. Region based methods are divided in

* Corresponding author.
E-mail addresses: suhad_ali2003@yahoo.com (S.A. Ali), ashwaqtalib@yahoo.com (A.T. Hashim).
Peer review under responsibility of University of Kerbala.

http://dx.doi.org/10.1016/j.kijoms.2016.03.004

two sub methods: connected component based method (CC), and edge detection based method [6]. 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 [7]. 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 is failed in locating text regions in images which have complex background.

2) **Texture based methods**: These methods work by extracting texture features of image firstly, and then a classification process is applied in the second stage to detecting text regions [8]. 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 [9].

The rest of the paper is organized as follows; Section 2 gives briefly a review of the recent related works. Section 3 presents the proposed method. Section 4 illustrates the experimental results and discussions and Section 5 gives the conclusions and suggestion for future works.

2. Related works

This section surveys many methods which are related to our proposed method:

- In 2010, A. Angadi at al. proposed an algorithm for texture features based on discrete cosine transform (DCT). The method is applied on 100 natural scene images, it is inefficient when image background is more complex like trees, vehicles [10].
- In 2010 Epshtein et al. proposed a method to detect texts in many languages with different fonts based on stroke width transform (SWT) [11].
- In 2011 C. Yi and Y. Tian proposed a method for extracting text strings with arbitrary orientations. It is based on text image partitioning and connected components [12].
- In 2012, Seeri et al. proposed an algorithm for Kannada text images using combination of techniques such as median filter, sobel edge detector, connected component labeling, and order static filter. It fails to extract very small characters [13].
- In 2013, H. Koo and D. H, Kim proposed an algorithm for text region detection based on two classifiers, the first one is used for generating candidate word regions and the second for filtering out non-text regions [14].
- In 2014, Raj et al. proposed an algorithm for natural scene text image detection using connected components (CC). It fails for small slanted/curved text [15].

3. The proposed method

As presented in Fig. 1 the proposed text detection algorithm passes through many steps.

3.1. Preprocessing

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

$$\begin{aligned} Y &= 0.299*R + 0.587*G + 0.114*B \\ U &= -0.14713*R - 0.28886*G + 0.436*B \\ V &= 0.615*R - 0.51499*G - 0.10001*B \end{aligned} \tag{1}$$

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

3.2. Discrete wavelet transform

Discrete Wavelet transform is referred as a multi-resolution decomposition approach. In the frequency domain; it decomposes a signal depending on a family of

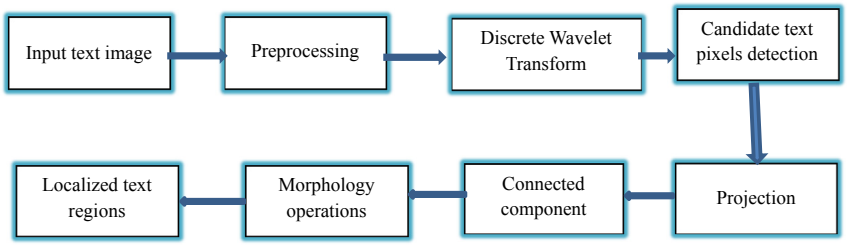

Fig. 1. The proposed method steps.