International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-1 Issue-5, October 2012

Localization of Text in Complex Images Using Haar Wavelet Transform

Neha Gupta, V.K Banga

Abstract*: In this paper, a new hybrid approach is developed which locate text in different backgrounds. However, variation of text due to differences in size, style, orientation and alignment, as well as low image contrast and complex background make the problem of automatic text localization extremely challenging. The text localization algorithm system is designed to locate text in different kinds of images and eliminates the need to devise separate method for various kinds of images. Firstly, the color image is converted into grayscale image. After that, Haar Discrete Wavelet Transform (DWT) is employed. Haar DWT decompose image into four sub image coefficients, one is average and other three are detail. Now, Sobel edge detector is applied on three detail components, the resultant edges so obtained are combined to form edge map. The morphological dilation is performed on binary edge map and further label the connected components. Finally, using some specific condition, the text is obtained in bounding box.

Index Terms*: Bounding box, Discrete Wavelet Transform, Haar wavelet, Sobel edge detector.

I. INTRODUCTION

Now days, with the advancement of the digital technology, the more and more databases are multimedia in nature. The databases usually contain images and videos in addition to the textual information. The textual information is very useful semantic information because it describes the image or video and can be used to fully understand images and videos. Text localization can be done in three kinds of images namely:

1. Document image
2. Scene text image
3. Caption text image

Document images may be in the form of scanned book covers, CD covers or video images. Text in images or videos is classified as scene text and caption text. Scene text is also called as graphics text. Natural images that contain text are called scene text. The name of caption text is artificial text and it is one in which text is inserted or superimposed in the image [7]. The diagram showing three types of images as in Fig. 1. There are two different approaches have been used for text localization from complex images namely region based approach and texture based approach

Fig. 1(a) : Document image

Fig. 1(b) : Scene text image

Fig. 1(c) : Caption text image

A. Region based approach

This approach uses the properties of the color or gray scale in a text region or their differences regarding the background. This method is basically divided in two sub categories: edge based and connected component (CC) based methods. The edge based method is mainly focus on the high contrast between text and background. In this method, firstly text edges are identified in an image and are merged. Finally, some heuristic rules are applied to discard non-text regions. Connected component based method considers text as a set of separate connected components, each having distinct intensity and color distributions. The edge based methods are robust to low contrast and different text size where as CC based methods are somewhat simpler to implement, but they fail to localize text in images with complex backgrounds [1].

B. Texture based methods

As we know that, text in images has distinct textural properties which can be used to differentiate them from the background or other non text regions [3]. This method is based on the concept of textural properties. In this method, Fourier transforms. Discrete cosine transform and wavelet decomposition are generally used. The main drawback of this method is that it is highly complex in nature but, in other hand, it is more robust than the CC based methods in dealing with complex background.

Revised Manuscript Received on 30 October 2012
*Correspondence Author(s)
Neha Gupta*, Department of Electronics and Communication, PTU University/ Amritsar College of Engineering and Technology, Amritsar, India.
Dr. V.K.Banga, Head of Department of Electronics and Communication, PTU University, Amritsar College of Engineering and Technology, Amritsar, India.

II. PROPOSED ALGORITHM

The block diagram of the proposed text localization algorithm is shown in Fig.2. The input image may be a color or gray scale image. If the image is color image then, preprocessing operation is applied on the image as shown in the flowchart. In our algorithm, input data is a color image which is entered to the system and the segmented text on a clear background is the output [9].

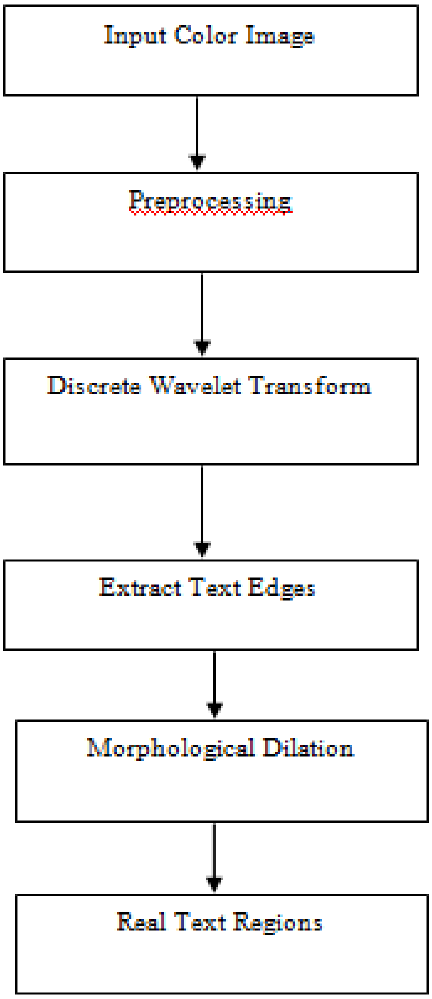

Fig. 2 : Text Localization Algorithm

A. *Preprocessing*

If the input image is a gray-level image, the image is processed directly starting at discrete wavelet transform. If the input image is colored , then its RGB components are combined to give an intensity image. Usually, color images are normally captured by the digital cameras. The pictures are often in the Red-Green-Blue color space. Intensity image Y is given by:

$Y = 0.299R + 0.587G + 0.114B$

Image Y is then processed with 2-d discrete wavelet transform. The Y is actually Value component of the Hue-Saturation-Value (HSV) color space. The RGB color image and its gray scale image is shown in Fig. 3. In this step, there is conversion from RGB color space into HSV color space, after that Value component is extracted from HSV color space using above expression.

The noise of the image is reduced by using a median filtering that is applied on the above grayscale image. After this filtering step, a great part of noise will be removed while the edges in the image are still preserved.

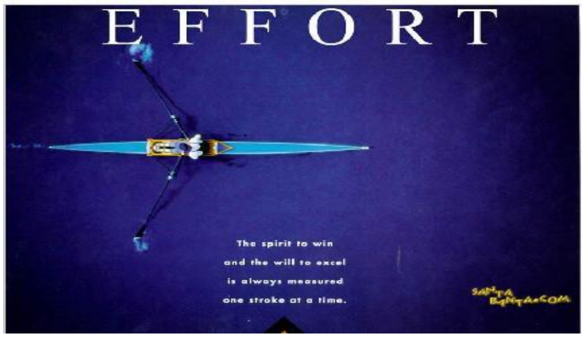

Fig. 3(a): Color image

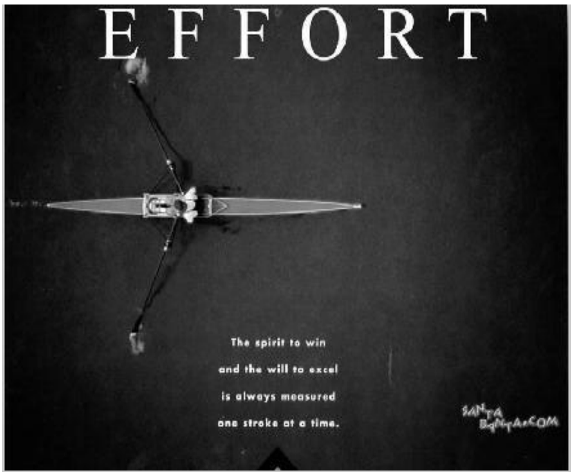

Fig. 3(b) : Gray scale image

B. *Discrete Wavelet Transform(DWT)*

In this algorithm, we are using Haar discrete wavelet transform which provides a powerful tool for modeling the characteristics of textured images. Most textured images are well characterized by their contained edges. It can decompose signal into different components in the frequency domain [12]. We are using 2-d DWT in which it decomposes input image into four components or sub-bands, one average component(LL) and three detail components(LL, HL, HH) as shown in Fig. 4 and 5[9] .The detail component sub-bands are used to detect candidate text edges in the original image. Using Haar wavelet, the illumination components are transformed to the wavelet domain. This stage results in the four LL, HL,LH and HH sub image coefficients. The traditional edge detection filters can provide the similar result as well but it cannot detect three kinds of edges at a time. Therefore, processing time of the traditional edge detection filters is slower than 2-d DWT. The reason we choose Haar DWT because it is simpler than that of any other wavelets.

Fig. 4 : Results of 2-d DWT decomposition

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Some of the following advantages are as follows:

1. Haar wavelets are real, orthogonal and symmetric.
2. Its coefficients are either 1 or -1.
3. Haar wavelets are real, orthogonal and symmetric.
4. Its coefficients are either 1 or -1.
5. It is the only wavelet that allows perfect localization in the transform domain.

Fig. 5 : DWT coefficients

C. *Localization of text edges*

In this localization process, three detail sub-components are used to detect edges of the text blocks. By finding the edges in the three sub images namely horizontal sub image, vertical sub image and diagonal sub image, fusing the edges contained in each sub image. In this way, candidate text regions can be found. In this algorithm, we use Sobel edge detector because it is efficient to locate strong edges that are needed in this application. The next step is to form an edge map using weighted 'OR' operator. To get binary edge map, thresholding is to be applied as shown in Fig. 6. The thresholding operation is performed to remove some non-text regions.

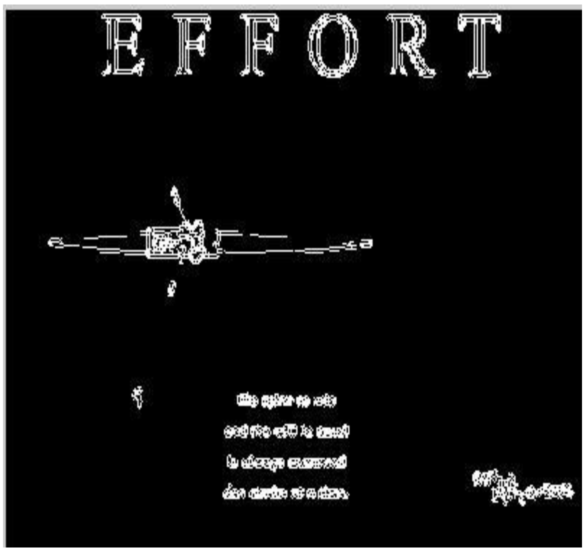

Fig. 6 : Binary edge map

D. *Morphological dilation*

First of all, the image is resized by using command

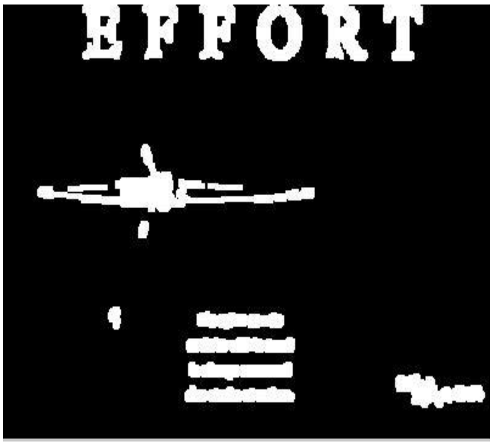

Fig. 7 : Dilated image

imresize in matlab. Now, Morphological dilation operation is performed on binary edge map obtained after thresholding. Basically, the function of dilation is to fill the gaps inside the obtained text regions as hown in Fig. 7. The structuring element we use in dilation is square of size 4x4.

E. *Real text regions*

The thresholding value we use here is 300 which has been experimentally set up.The first step is to label the connected components by using bwlabel function in matlab. The connectivity we use is 8. Now, various shape measurements are computed by using regionprops command . Actually, this command measures a set of properties for each connected component (object) in the binary image. But, we only consider area and bounding box shape measurements as per our requirement. Area is a scalar, it is actually number of pixels in the region. This value might differ slightly from the value returned by bwarea, which weights different patterns of pixels differently. Bounding box is the smallest rectangle containing the region. It include width and height of the rectangle and it also specifies the upper left corner. A new value is computed by multiplying height and width of bounding box. The ration of this new value and area is taken. If the ratio is less than 2, then the regions so obtained are text regions. Finally, copy RGB values in that position so that real text regions are obtained with background as shown in Fig. 8.

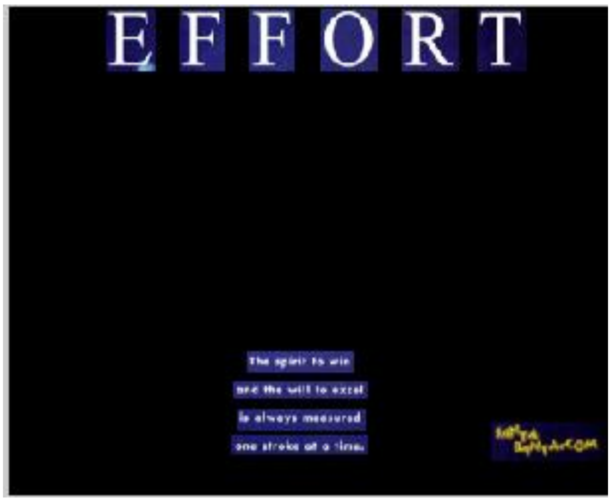

Fig.8 : Real text regions

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

III. EXPERIMENTAL RESULTS

The algorithm is implemented using MATLAB. Experiments were carried out on various images. Each image is in BMP or JPEG format. Since the images are colored, they are converted into gray-level before applying the discrete wavelet transform. The proposed method can correctly locate text regions in complex images. Basically, this method successfully detect 100% of the text regions. The results of the proposed algorithm when run on different kinds of images such as scene text and caption text , wallpapers are shown below in Fig. 9, 10, and 11 respectively.

Fig. 9 : Scene text image

Fig. 10(a) : Caption text image

Fig. 10(b) : Caption text image

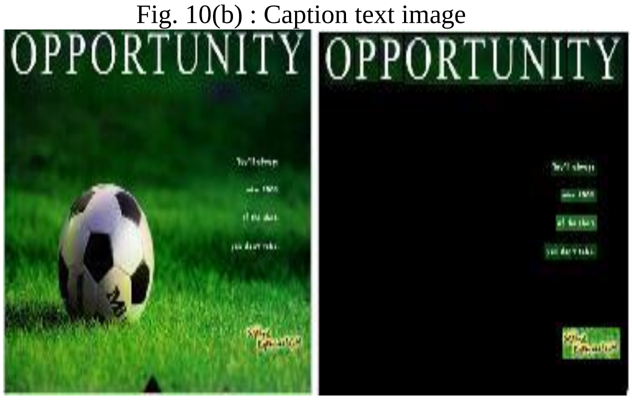

Fig. 11: Wallpaper

IV. CONCLUSION

In this paper, we present a relatively simple and effective algorithm for text detection and extraction by applying DWT to the images [13]. As it requires less processing time which is essential for real time applications [11]. Mostly all the previous methods fail when the characters are not aligned well or when the characters are too small. They also result in some missing characters when the characters have very poor contrast with respect to the background [12]. But this algorithm is not sensitive to image color or intensity, uneven illumination and reflection effects. This algorithm can be used in large variety of application fields such as vehicle license plate detection to detect number plate of vehicle, mobile robot navigation to detect text based land marks, object identification, identification of various parts in industrial automation, analysis of technical papers with the help of charts, maps, and electric circuits etc.. This algorithm can handle both scene text images and printed documents. Our main future work involves a suitable existing OCR techniqueto recognize text. The binary output can be directly used as an input to an existing OCR system for character recognition. Also, the algorithm only analysis text box not a single character . Therefore, it requires less processing time which is essential for real time applications.

REFERENCES

1. Keechul Jung, Kwang In Kim and Anil K. Jain, "Text information localization in images and video: A Survey", Elsevier, Pattern Recognition, vol.37 (5), pp 977–997, 2004.
2. H.K.kim, "Efficient Automatic Text Location Method and Content Based Indexing and Structuring Of Video Database", Journal of Visual Communication Image Representation,vol.7(4),pp336-344,1996.
3. Mohieddin Moradi, Saeed Mozaffari, and Ali Asghar Orouji, "Farsi/Arabic Text Localization from Video Images by Corner Detection", 6th, IEEE, Iranian conference on Machine Vision and image processing, Isfahan, Iran, 2010.
4. Chung-Wei Liang and Po-Yueh Chen, "DWT Based Text Localization", International Journal of Applied Science and Engineering, pp.105-116, 2004.
5. Nikolaos G. Bourbakis, "A methodology for document processing: separating text from images", Engineering Applications of Artifcial Intelligence 14, pp. 35-41, 2001.
6. C. Strouthopoulos, N. Papamarkos, A. E. Atsalakis, "Text localization in complex color documents", The Journal of Pattern Recognition 35, pp.1743–1758, 2002.
7. Chitrakala Gopalan and D. Manjula, "Contourlet Based Approach for Text Identification and Localization from Heterogeneous Textual Images", International Journal of Electrical and Electronics Engineering 2(8), pp. 491-500, 2008.
8. Dr.N.Krishnan, C. Nelson Kennedy Babu 2, S.Ravi 3 and Josphine Thavamani, "Segmentation Of Text From Compound Images", International Conference on Computational Intelligence and Multimedia Applications", Tamil nadu , India, vol. 3, pp.526-528, 2007.
9. S.Audithan, RM. Chandrasekaran, "Document Text Localization from Document Images Using Haar Discrete Wavelet Transform", European Journal of Scientific Research ISSN 1450-216X, Vol.36, 2009.
10. Xiaoqing Liu and Jagath Samarabandu, "Multiscale edgebase Text localization from Complex images", IEEE Multimedia and Expo 2006, International Conference, Tronto, Canada, pp. 1721-1724, 2006.
11. Roshanak Farhoodi and Shohreh kasaei, "Text Segmentation From Images With Textured and colored Background", 13th Iranian conference on Electrical Engineering, 2005 .

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

12. Xiao-Wei Zhang, Xiong-Bo Zheng, Zhi-Juan Weng, "Text Localization Algorithm Under Background Image Using Wavelet Transforms", Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, Hong Kong, pp.30-31, Aug. 2008.
13. R. lienhart, F. Stuber, "Automatic Text Segmentation and Text Recognition for Video Indexing", The Journal Multimedia Systems , vol(8), pp 69-81,Jan 2000.
14. M. Padmaja, J. Sushma, "Text Detection in Color Images", International Conference on Intelligent Agent & Multi-Agent Systems, Chennai, 22-24 July 2009, pp. 1-6, 2009.

AUTHOR PROFILE

Neha Gupta is scholar in Department of Electronics & Communication Engineering in Amritsar College of Engineering and Technology, Amritsar, Punjab, India. She did his B.E (Electronics and Communication Engineering) from Jammu University, J&K, India. and pursuing M.Tech (Electronics and Communication Engineering) from Amritsar College of Engineering and Technology, Amritsar, Punjab, India. She has 4 national and 1 international conferences to her credit

Vijay Kumar Banga is working as Professor and Head of the Department of Electronics & Communication Engineering, Amritsar College of Engineering and Technology, Amritsar, Punjab, India. He obtained his B. E (Electronics and Instrumentation Engineering) from Punjabi University, Patiala, Punjab, India, M.Tech (Electronics and Instrumentation) from Panjab University, Chandigarh, India and Ph.D. in Electronics (Artificial Intelligence) form Thapar University, Patiala., India. Presently, he has 13 years of research and UG & PG teaching experience. He has 35 research papers to his credit in various international journals and conferences. He is member of Board of Studies in Electronics and Communication Engineering of Punjab Technical University, Jalandhar, Punjab, India. His areas of research and interest include Robotics, Artificial Intelligence, automated control systems and Genetic algorithms.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)