World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:1, 2011

Journals Subheadlines Text Extraction Using Wavelet Thresholding and New Projection Profile

Davod Zaravi, Habib Rostami, Alireza Malahzaheh, S.S Mortazavi

***Abstract*—**In this paper a new robust and efficient algorithm to automatic text extraction from colored book and journal cover sheets is proposed. First, we perform wavelet transform. Next for edge detecting from detail wavelet coefficient, we use dynamic threshold. By blurring approximate coefficients with alternative heuristic thresholding, achieve effective edge,. Afterward, with ROI technique get binary image. Finally text boxes would be extracted with new projection profile.

***Keywords*—** Text extraction; colored cover sheet; wavelet threshold, region of interest (ROI).

I. Introduction

ALTHOUGH text extraction from color images and color documents is one of interests for researchers, but investigate around cover sheet of journals and books seems to be mint domain. Searches where deals with it object are so few. In this paper we try to propose a new robust and effective approach which could detect and extract sub titles on cover sheet of book and journals.

Growth in the volume of literature makes the search in printed journals and books covers more difficult. They may have various non text and the text ones which can be printed in various font, size, color or spacing and style**.**

Mean of mid, sub and general headlines and titles are whose not so large font that categorized in out of size and so small font that categorized in describer context for subtitles.

The paper is organized as follows: next section, discussed around thresholding and its kinds. Wavelet thresholding is mentioned too. Section 3 contains related works. So many papers might be found on color images text detection. But so few are deals with colored books, journals or magazines, unfortunately.

S.S Mortazavi is with Elect. Engg Dept. Islamic Azad University, Bushehr branch, Iran. (e-mail:mortazavi_s@scu.ac.ir)

Section 4 involves our algorithm. After preprocessing it has three main steps:

1- Wavelet decomposition: finding sub bands, performing dynamic thresholding and blurring contains it three sub steps.
2- Region of interest (ROI) binarization: in this step we propose a new selecting yardstick to how this region must be selected.
3- Getting bounding box with implementing new projection profile and other heuristic features.

These procedures illustrated in figure 1as flowchart. It demonstrates 5 basic steps:

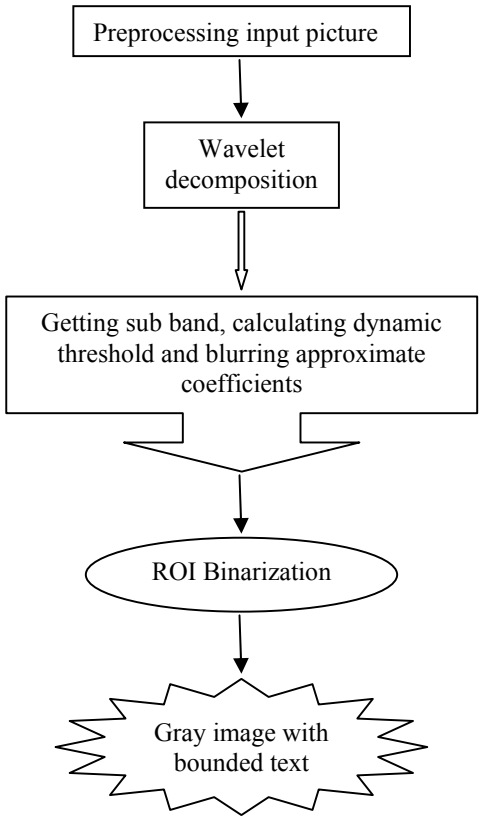

Fig. 1 Flowchart of proposed algorithm

Finally in section 5 conclusion ends paper.

II. Thresholding

A. Thresholding

Thresholding can be used in many widespread applications, such as: extracting logous, musical scores, endoscopic images, graphical context, cell images or knowledge representation [1], and in this paper dealing with text extraction.

Binarization in many image processing procedures is the first step. Local binarization methods can be improved by calculating local thresholds within separate windows or areas. In different image processing applications object pixels are substantially different from background [1]. Binarization goal is converting given color or gray scale input into a bi level representation [2].

In Ref. [1] Mehmet Sezgin and Bulent Sankur present a survey over image thresholding methods. They categories the thresholding methods in six groups, according to the information:

- Histogram shape-based methods, where in this method, for example, the peaks, curvatures and valleys of the smoothed histogram were be analyzed.
- Clustering based methods, where in it the gray-level samples are clustered in two parts: background and foreground (object), or are modeled as a mixture of two Gaussians, alternately.
- Methods based on entropy: result in algorithms that use the entropy of the foreground and background regions, the cross-entropy between the original and binarized image, etc.
- Object attribute-based methods whose search a measure of similarity between the binarized and the gray-level images, fuzzy shape similarity, other artificial methods and edge coincidence categorized in this category.
- Spatial methods that use higher-order probability distribution and/or correlation between pixels.
- Local methods adapt the threshold value on each pixel to the local image characteristics.

Our proposed approach could be categorized in both several method clusters, Because of from one hand we cluster input to text and non text pixels and other hand calculate soft threshold. We use histogram for that too. Performing dynamic threshold on each (sub band coefficient) pixel can put proposed algorithm in other method cluster.

B. Wavelet thresholding

Recently there has been significant investigations in using the wavelet transform as a tool for improving thresholding [1, 3, 4, 5].

The discrete wavelet transform (DWT) corresponds to basis decomposition, then it could obtain a specific and unique representation of the signal. In many discrete wavelet transforms (DWT), because of coefficients that are either 1 or -1, Haar operates the fastest among all wavelets [5].

In all wavelet thresholding methods basic steps are:

- Employing the DWT on the image, and calculating coefficient
- Thresholding the wavelet coefficients. (Threshold may be universal or sub band adaptive).
- Computing the IDWT to get the thresholded image.
- There are two thresholding functions frequently used, i.e. a hard threshold, a soft threshold [3].

III. Related works

As explained in previous section, in ref. [1] Mehmet Sezgin and Bulent Sankur described 40 thresholding methods.

Faisal Shafait et al. [2] present a fast approach to compute local threshold. Their approach uses integral images to compute mean and variance in local windows. They compaires their approach with Sauvola and otsu method, and achieve the same results as Sauvola, but in a time, reached close to global binarization schemes like otsu.

S.Sudha et al. [3] present a wavelet thresholding algorithm for adaptive threshold selection for denoising ultrasonic images. Wavelet thresholding results efficient thresholds. This kind of thresholding is mentioned in previous section.

Lakhwinder Kaur et al. [4] have proposed an adaptive threshold estimating method for image denoising purpose in the wavelet domain, named normalshrink. In it parameters required for estimating the threshold depend on subband data. Their algorithm based on the generalized Guassian distribution (GGD) modeling of subband coefficients.

S. Audithan and RM. Chandrasekaran [5], first detect edges by wavelet transfom then generating line feature vector graph that was based on the edge map. After extracting stroke information, generate text region and filter them according line features finally. Authors in that paper proposed a mean based thresholing. We use it in our algorithm as a dynamic thresholding.

K. Sobotka et al. [6] directly studying colored book and journal covers. In this research, they develop two methods for extracting text hypotheses, and then the results of both methods are combined to robustly distinguish between text and non text elements.

Gholamreza Aghajari and Jamshid Shanbehzadeh in [7] introduce new projection profile. That technique with counting alternately changes in binary image provide proper feature of text regions. We use new projection profile as improving feature. Our experimental result illustrate this feature could get better output.

IV. Methodology

For preprocessing, RGB component of received colored image would be combined to intensity image as follow:

$$Y= [0.299\ 0.587\ 0.114]*[RGB]^{T}$$

A. Wavelet decomposition

This step includes:

1- Obtaining sub bands.
2- Calculating dynamic threshold for each pixel in details sub bands; and,
3- Blurring approximate sub band (to use details only as edges).

B. Calculating dynamic threshold

Let us describe it further. We select an appropriate dynamic threshold value and preliminarily remove the non text edges in the each detail component sub band. We obtain the target threshold value by performing an equation on each pixel with its neighboring pixels:

$$g1 = [-1\ 0\ 1]$$
$$g2 = [-1\ 0\ 1]^T \quad (1)$$

Let us consider a mask for illustrating effect of mask such as follow:

	1	
1	0	-1
	-1	

However, this mask combined from two ones that perform horizontal and diagonal on each pixel and two its neighbor, then select maximum one of this two. If "es(i,j)" denotes "each sub band pixels in i'th row and j'th column" Thus for each pixel would have an equation such as:

$$s(i, j) = Max(abs[g1 * *es(i, j), abs[g2 * *es(i, j)]) \quad (2)$$

An example for equation (2) could be understand from figure 2 considering equation (3).

$$s(p8) = \max(|p9 - p7|, |p13 - p3|) \quad (3)$$

$$\begin{bmatrix} P1 & P2 & P3 & P4 & P5 \\ P6 & P7 & P8 & P9 & P10 \\ P11 & P12 & P13 & P14 & P15 \\ P16 & P17 & P18 & P19 & P20 \\ P21 & P22 & P23 & P24 & P25 \end{bmatrix}$$

Fig. 2 Example for calculating "s" matrix for calculating dynamic threshold

"s" is a new matrix must be used in next equation finding threshold:

$$T = \frac{\sum (es(i, j) \times s(i, j))}{\sum s(i, j)} \quad (4)$$

This target obtains for ever pixels individually and aims to different value for each pixel. Finally, algorithm performs each threshold on its each detail component pixel [5]. Any pixel is lower than itself threshold has be eliminate and be zero. However this is gray scale yet, and only in rang pixels remained.

C. Blurring

Approximate in wavelet transform aimed from approximating image, and details from it changing. Then details can get edges, and approximate can get a new image closed same to original one. Thus we blur approximate sub image to get edges. But point is what kind blurring is suitable for this purpose? This sub stage needs a thresholding value, too. Our experimental aimed us this value affect text extraction, sizably. In other hand we lend any image get different value as proper constant blurring threshold. Then it is private for each image. Studying many cases points this parameter has significant relation between horizontal coefficients threshold (aimed from previous equations) and mean of intensity of whole intensity image. According it, we scale blurring threshold from 2 to almost 3 times rather than horizontal coefficient threshold. Then, recombining thresholded detail coefficients and blurred approximated one, a gray scale image obtain. This method bright recombined image according original image brightness.

D. Binarization

This step contains ROI (region of interest) technique. In it, a certain region has been selected and get 1s logical value and unfiltered values for pixels where the binary mask contains 0s [8]. This region could be selected arbitrary at certain color, intensity, sector or etc. Thus, a binary image returns. Now, in our set of magazine and journals coversheets where is interested region?! Let us, with focusing on another image feature answer it: that is histogram. An attractive point in these images is their histogram, integrated on limited region such seems like narrow band in histogram plot. It is interesting region in our approach. Real binarization performs by this technique.

New projection profile used after some heuristic modifying like removing long straight lines and count all binary level changes [5]. Old projection profile counts pixels have 1 value but this new one counts changes. Thus could be robust against some effects aimed of noise. So we can estimate the possibility of the text presence in every line especially for texts with normal font size via thresholding. Finally, from processed image, such text candidates were be selected when have following conditions:

Width > 10 pixels & Length > 2*width

V. CONCLUSION

In this work we have introduced a relatively simple and effective algorithm. We perform our algorithm on 80 pictures would been collected from internet. They contain various quality ones and might be scanned or be a digital improved image and might have few sub headlines or messy with so lot ones. Their backgrounds could be simple or complexity complicated. However, result was satisfying: 91.2 percent hit rate.

Fig 3. Illustrate an example image. In this image we can see some features of each image: preprocessed gray scale, each of threefold wavelet coefficient sub bands and image histogram.

Fig 4. Outlines result of Fig. 4. and some other examples.

REFERENCES

[1] Mehmet Sezgin , Bulent Sankur "Survey over image thresholding techniques and quantitative performance evaluation" Journal of Electronic Imaging 13(1), 146–165 (January 2004).

[2] Faisal Shafait, Daniel Keysers, Thomas M. Breuel, "Efficient Implementation of Local Adaptive Thresholding Techniques Using Integral Images", Proc. SPIE, Vol. 6815, 681510 (2008).

[3] S. Sudha, G. R.Suresh and R. Sukanesh. "Speckle Noise Reduction in Ultrasound Images by Wavelet Thresholding based on Weighted Variance", International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009 ,1793-8201.

[4] S. Sudha, G.R. Suresh, R. Sukanesh, "Wavelet Based Image Denoising Using Adaptive Thresholding," iccima, vol. 3, pp.296-300, International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), 2007.S. Audithan, RM. Chandrasekaran "Document Text Extraction from Document Images Using Haar Discrete Wavelet Transform" European Journal of Scientific Research;SSN 1450-216X Vol.36 No.4 (2009), pp.502-512.

[5] K. Sobotka, H. Kronenberg, T. Perround,H. Bunke, "Text extraction from colored book and journal covers" International Journal on Document Analysis and Recognition, Volume 2, Number 4 / June, 2000.

[6] Gholamreza Aghajari, Jamshid Shanbehzadeh, "A Text Localization Algorithm in Color Image via New Projection Profile" International MultiConference of Engineers and Computer Scientists 2010.

[7] "Image Processing Toolbox for Use with MATLAB®" User's Guide Verson3,©COPYRIGHT 1993 - 2002 by The MathWorks, Inc.

[8] Chitrakala Gopalan, D.Manjula, "Sliding window approach based Text Binarisation from Complex Textual images" International Journal on Computer Science and Engineering Vol. 02, No. 02, 2010.

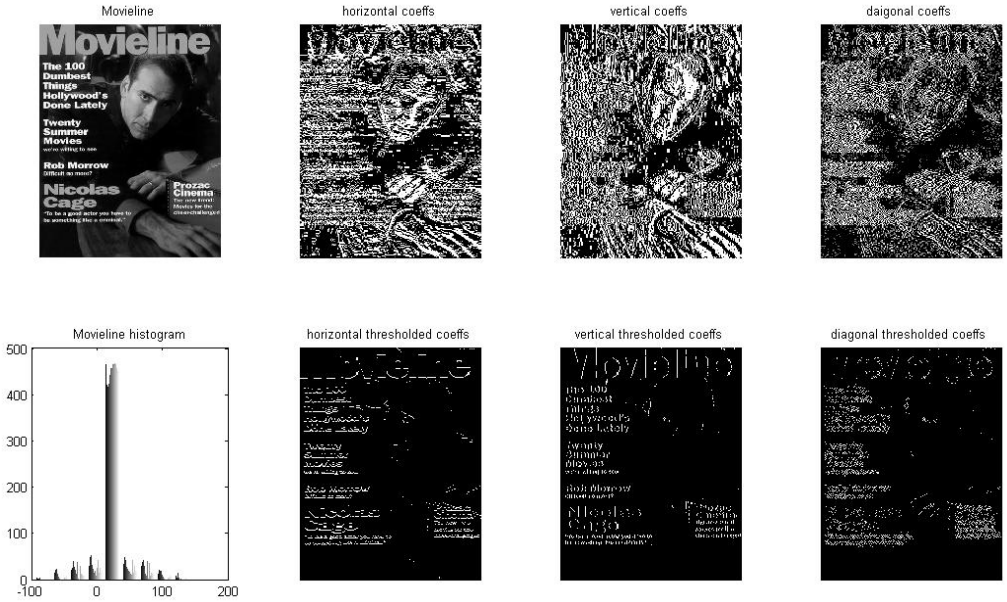

Fig. 3 image and its features: preprocessed gray scale, histogram and threefold wavelet coefficients before and after thresholding

Fig. 4 Some other example of text extraction from proposed algorithm results and their histograms