*International Journal of Engineering Research and Development*
***e**-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com*
*Volume 3, Issue 2 (August 2012), PP. 47-53*

# Document Image Segmentation for Analyzing of Data in Raster Image

Dr .P.Sengottuvelan[1], Mr.R.Arulmurugan[2], Mr.R.Lokeshkumar[3]
*[1,2,3]Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, India*

***Abstract*—This paper focuses on the needs of developing an automated Digital Library Management system .The purpose is to automate the task of analyzing data containing in raster image documents for the purpose of intelligent information retrieval in digital library. An efficient and computationally fast method for segmenting text and graphics part of document images based on multi-scale wavelet analysis and statistical pattern recognition is presented. The extracted text is further classified into Title, Author name, name of the publication etc and being stored in the database for further Library related operations. We do not assume any a priori information regarding the font size, scanning resolution, type of layout, etc. of the document in our segmentation scheme.***

***Keywords*——Document segmentation, daubechies wavelet, Multiscale wavelet analysis, priori Information, Fourier transform.***

## I. INTRODUCTION

In Today's world, automated processing and reading of documents has become an imperative need with efforts have been made to store the documents in digitized form, but that requires an enormous storing space, even after compression using modern techniques. Documents can be more effectively represented by separating the text and the graphics/image part and storing the text as an ASCII (character) set and the graphics/image part as bit-maps. Document image segmentation plays an important role because this facilitates efficient searching and storage of the text part in documents, required in large databases. Consequently, several researchers have attempted different techniques to segment the text and graphics part in document images [1]. Several useful techniques for text–graphics segmentation are given in, the most popular amongst these being the top-down and bottom-up approaches[2]-[4].

The most common top-down techniques are run-length smoothing and projection profiles. Top-down approaches first split the document into blocks, which are then identified and subdivided appropriately in terms of columns first and then into paragraphs, text lines, and maybe also words[5]-[8]. Some assume these blocks to be only rectangular. The top-down methods are not suitable for skewed texts, as these methods are restricted to rectangular blocks, whereas the bottom-up methods are typically variants of the connected components which iteratively group together components of the same type starting from the pixel level and form higher level descriptions of the printed regions of the document (words, text lines, paragraphs etc.). The drawbacks with the connected components method is that it is sensitive to character size, scanning resolution, inter-line, and inter-character spacing. A wavelet-based tool has been designed by them for distinguishing text from non text regions and characterization of font sizes [8]. Some of the common difficulties that occur in documents are given below:

- Differences in font size, column layout, orientation, and other textual attributes.
- Skewed documents and text regions with different orientations.
- Degraded documents due to improper scanning.
- Combinations of varying text and background gray levels.
- Text regions touching or overlapping with non-text regions.
- Irregular layout structures with non-convex or overlapping object boundaries.
- Multicolumn document with misaligned text lines and different languages.

Thus to develop a full fledged system all the above said difficulties should be overcome.

## II. PROPOSED SYSTEM

In the system thus proposed a technique called document image segmentation is used where text such as Title, Author name, name of the publication etc.. is being extracted from the image being scanned (front cover of the book) classified accordingly and is being stored in the database for further Library related operations. Thus the proposed system avoids the need for manual entry of the information in the database.


***Fig.1***.*Raster Image*

A working principle of the proposed system

- The image is being scanned from book using a good quality scanner.
- The textual regions are being separated and classified accordingly.
- Then the segmented text is being stored in the database for further library related operations.

## III. WAVELET THEORY

The wavelet transforms has many unique features that had made it a popular method for the purpose of image processing. The wavelet transform performs a high degree of decor relation between neighboring pixels and it also provides a distinct localization of the image in the spatial as well as frequency domain. The transform also provides an elegant sub-band framework in which both high and low frequency component of the image can be analyzed separately.

### A. Wavelet transform vs. Fourier transform

***1. Similarities***

The fast Fourier transform(FFT) and the discrete wavelet transform(DWT) are both linear operations that generate a data structure that contains $\log_2 n$ segments of various lengths, usually filling and transforming it into a different a data vector of length 2n.The mathematical properties of the matrices involved in the transforms are similar as well. The inverse transform matrix for both FFT and DWT is the transpose of the original. As a result, both transforms can be viewed as a rotation in function space to a different domain. For the FFT, this new domain contains basis functions that are sines and cosines. For the wavelet transform, this new domain contains more complicated basis functions called wavelets, mother wavelets, or analyzing wavelets.

Both transforms have another similarity, the basis functions are localized in frequency, making mathematical tools such as power spectra useful at picking out frequencies and calculating power distributions.

***2. Dissimilarities***

The most interesting dissimilarity between these two kinds of transforms is that individual wavelet functions are localized in space. Fourier sine and cosine functions are not. This localization feature, along with wavelet localization of frequency, makes many functions and operators using wavelets sparse when transformed into the wavelet domain. This sparseness, in turn, results in a number of useful applications such as data compression, detecting features in images, and removing noise from time series.

An advantage of wavelet transforms is that the width of the windows varies. In order to isolate signal discontinuities, one would like to have some very short basis functions. At the same time in order to obtain detail frequency analysis, one would like to have some very long basis functions. A way to achieve this is to have short high-frequency basis functions and long low-frequency ones. Thus wavelet analysis provides immediate access to information that can be obscured by other time frequency methods such as Fourier analysis.

### B. Wavelet Transform Applied on Images

To use the wavelet transform for image processing we must implement a 2D version of the analysis and synthesis filter banks. In the 2D case, the 1D analysis filter bank is first applied to the columns of the image and then applied to rows. If the image has N1 rows and N2 columns, then after applying the 1D analysis filter bank to each column we have two sub band images, each having N1/2 rows and N2 columns; after applying the 1D analysis filter bank to each row of both of the two sub band images, we have four sub band images, each having N1/2 rows and N2/2 columns. This is illustrated in the diagram below. The 2D synthesis filter bank combines the four sub band images to obtain the original image of size N1 by N2. This is shown in figure.2.

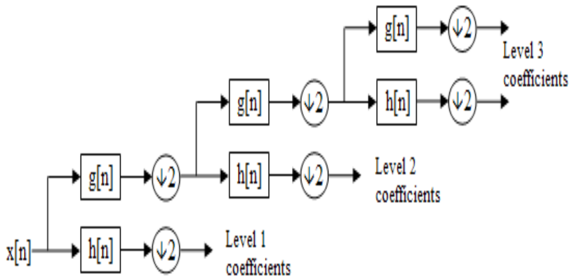



***Fig .1:*** *Multi resolution wavelet decomposition of an image.*

The following figure.3 is the decomposition of an image into four frequency bands by using wavelet transform as explained above.

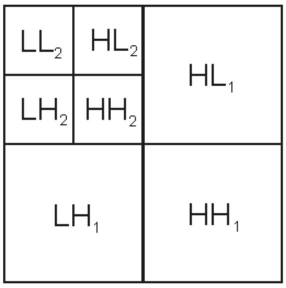



***Fig.3: Decomposition of an image into four frequency bands by wavelet transforms***

The continuous wavelet transform of a function is given as

$$Wf_a(b) = \int f(x)\Psi^*{}_{a,b}(x)dx.$$

Applying wavelet transform for our sample image,

In this paper we have taken a sample image (front cover of the book).we must use a good quality scanner such that the scanned image is free from any noise or distortion. If so the noise is being eliminated by using suitable filtering techniques. The scanned image can be in anyone of the formats like JPEG, BMP and Tiff. Here we have used JPEG format. Image decomposition is achieved at many levels and many type of wavelets like harr, daubechies, coiflets, symlets etc can be used. Here we have used daubechies wavelet for decomposing an image into various levels. The following figure illustrates the wavelet transform applied on a sample image.

**Level 1 Filter type: db1**


**Level** 2 **Filter type**: db1

**C. Wavelet Scale–Space Features**

The feature-extraction scheme that we have used has a multi-channel filtering and a subsequent nonlinear stage followed by a smoothing filter (both these constitute the local energy estimator) [8] [9].The objectives of the filtering and that of the local energy estimator are to transform the edges between textures into detectable discontinuities.

The filter bank, in essence, is a set of bandpass filters with frequency- and orientation-selective properties. In the filtering stage, we make use of an eight-tap, four-band, orthogonal– and linear phase wavelet transform following to decompose the textured images into M×M corresponding to different direction and scales [14]. The *one-dimensional* (1-D) four-band wavelet filter impulse responses denoted by $\psi_r$ are given in Table I, and their corresponding transfer functions are represented by $H_r$ for r =1,2..4. In this paper, we extend the decomposition to the 2-D case by successively applying the M -band transform separably in the horizontal and vertical directions without downsampling (i.e., an overcomplete representation). The size of the filter is an important factor. The filter length is increased with increased level of decomposition. The sequences of low-pass and bandpass filters of increasing width corresponding to an increased level of decomposition are expanded by inserting an appropriate number of zeros between taps of filters. So, if the filter length becomes large, it is possible that it may bias the decomposition of the image. We have chosen an eight-tap filter for suitability of the size of the image that we have considered in this study (i.e., 512×512)

***Table I - Coefficients for Eight-tap Four-band wavelet***.

| **No. of Taps(n)** | $\Psi_1(n)$ | $\Psi_2(n)$ | $\Psi_3(n)$ | $\Psi_4(n)$ |
|---|---|---|---|---|
| 0 | -0.067371764 | -0.09419511 | -0.09419511 | -0.067371764 |
| 1 | 0.09419511 | 0.067371764 | -0.067371764 | -0.09419511 |
| 2 | 0.40580489 | 0.56737176 | 0.56737176 | 0.40580489 |
| 3 | 0.56737176 | 0.40580489 | -0.40580489 | -0.56737176 |
| 4 | 0.56737176 | -0.40580489 | -0.40580489 | 0.56737176 |
| 5 | 0.40580489 | -0.56737176 | 0.56737176 | -0.40580489 |
| 6 | 0.09419511 | -0.067371764 | 0.067371764 | 0.09419511 |
| 7 | -0.067371764 | 0.09419511 | -0.09419511 | 0.067371764 |

The objective of the filtering is to find out about the discontinuities that exist within the image. The spectral response is strongest along the direction perpendicular to the edge of an image, while it decreases as the direction of the filter approaches that of the edge [16]. Therefore we can perform edge detection by using 2-D filtering of the image as follows:

- Horizontal edges are detected by high-pass filtering on columns and low-pass filtering on rows.
- Vertical edges are detected by low-pass filtering on columns and high-pass filtering on rows.
- Diagonal edges are detected by high-pass filtering on columns and high-pass filtering on rows.
- Horizontal–diagonal edges are detected by high-pass filtering on columns and low-pass filtering on rows.
- Vertical–diagonal edges are detected by low-pass filtering on columns and high-pass filtering on rows.

A typical edge-detection filter corresponding to a particular direction covers a certain region in the 2-D spatial-frequency domain. Based on this concept, several wavelet-decomposition filters are designed which are given by the summations $\Sigma_{Reg} H_{r,c}$, where Reg denotes the frequency sector of a certain direction and scale.

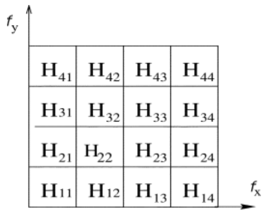



***Fig.4:*** *Frequency bands corresponding to decomposition filters.*

The basic decomposition scheme followed in this paper is given in the following figure.5.

**D. Local Energy Estimation**

The next step is to estimate the energy of the filter responses in a local region around each pixel. The local energy estimate is utilized for the purpose of identifying areas in each channel where the band pass frequency components are strong resulting in a high energy value and the areas where it is weak into a low energy value. Although energy is usually defined in terms of a squaring nonlinearity, in a generalized energy function, however, other alternatives are also used.

We have studied several nonlinear operators. These include the *magnitude operation, average absolute deviation and standard deviation* calculated over small overlapping windows around each pixel. This nonlinear operator is independent of any parameter, i.e., independent of the dynamic range of the input image and also of the filter amplification.

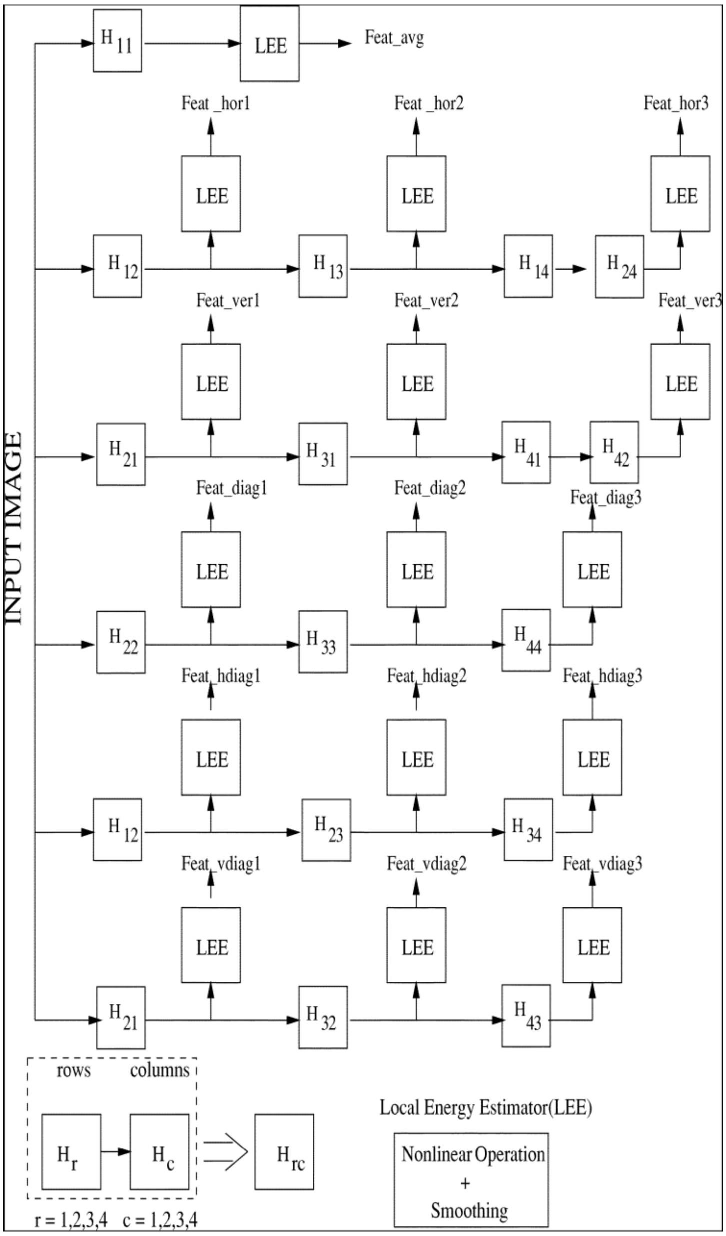



***Fig.5:*** *Basic decomposition scheme*

The images resulting from these operations are the features denoted by $Feat_{hor_i}, Feat_{ver_i}$ etc. for *i=1, 2, 3*.as shown in Fig.5

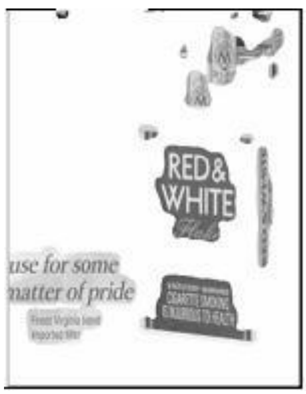
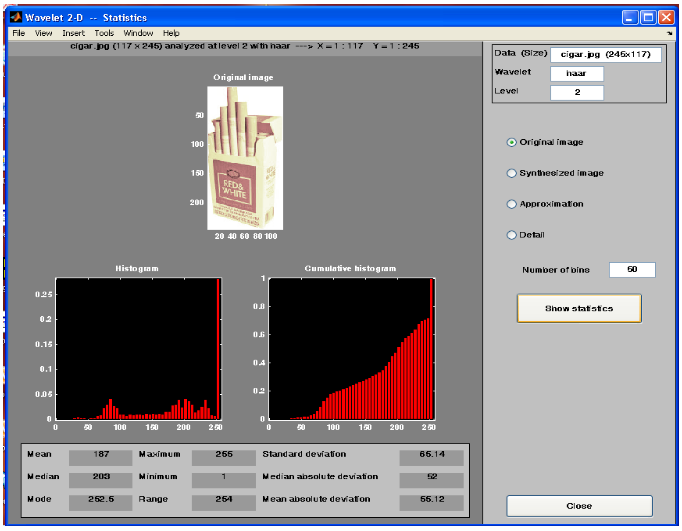




***Fig.6:*** *(a) Test image with document skewed and text regions with different orientations.* *(b) Segmented result.*




It is to be noted that all of our experiments were performed with no *a priori* knowledge about the input image. We did not have any information about the font size or format of the text. While knowledge about these can definitely improve the segmentation results, for this, we can make use of supervised segmentation.

## IV. CONCLUSION

We have segmented the text regions from the scanned image (Front cover of any image).The segmented text must further be subjected to further analysis for identifying the characters that facilitates intelligent information retrieval that could be used efficiently in our system. It is quite apparent that there is a need for digitization of documents for making it easily accessible via computers and networks, but it is not absolutely necessary to align the document in raster direction.

Thus the system overcomes the difficulties of the conventional bar code scanning wherein the user must manually enter the details of every image that comes to the Cover design. This method of automation is more efficient and requires less human labour. Thus an automated image information system is developed which minimizes manual interference considerably saving time and money yet the system has its own difficulties like recognizing text with different font sizes and styles may be quite difficult.

## REFERENCES

**[1].** S. N. Srihari, "Document image understanding," Proc. IEEE Computer Society Fall Joint Computer Conf., pp. 87–96, Nov. 1986.

**[2].** P. Chauvet, J. Lopez-Krahe, E. Taflin, and H. Maitre, "System for an intelligent office document analysis, recognition and description," Signal Processing, vol. 32, no. 1–2, pp. 161–190, 1993.

**[3].** F. M. wahl, K. Y. Wong, and R. G. Kasey, "Block segmentation and text extraction in mixed text/image documents," Comput. Graph. Image Process, vol. 20, pp. 375–390, 1982.

**[4].** F. Shih, S.-S. Chen, D. Hung, and P. Ng, "A document image segmentation, classification and recognition system," in Proc. Int. Conf. Systems Integration, 1992, pp. 258–267.

**[5].** O. Iwaki, H. Kida, and H. Arakawa, "A segmentation method based on office document hierarchical structure," in Proc. Intl. Conf. Systems, Man and Cybernetics, 1987, pp. 375–390.

**[6].** D. Wang and S. N. Srihari, "Classification of newspaper image blocks using texture analysis," Compu. Vis., Graph, Image Process, vol. 47, pp. 327–352, 1989.

**[7].** G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis for technical journals," Computer, vol. 25, no. 7, pp. 10–22, 1992.

**[8].** Mausumi Acharyya and Malay K. Kundu, " Document Image Segmentation Using Wavelet Scale–Space Features" IEEE transactions on circuits and systems for video technology, vol. 12, no. 12,December 2002.

**[9].** M. Krishnamoorthy, G. Nagy, S. Seth, and M. Viswanathan, "Syntactic segmentation and labeling of digitized pages from technical journals," IEEE Trans. Pattern Anal. Machine Intell., vol. 15, no. 7, pp. 737–747,1993.

**[10].** G. Loum, P. Provent, J. Lemoine, and E. Petit, "A New method for texture classification based on wavelet transform," in Proc. 3rd Int. Symp. Time-Frequency Time-Scale Analysis, June 1996, pp. 29–32.

**[11].** S. G. Mallet, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Machine Intel, vol. 11, July 1989.

**[12].** O. Rioul and M. Vetterli, "Wavelets and signal processing," IEEE Signal Processing Mag., vol. 8, pp. 14–38, Oct. 1991.

**[13].** O. Alkin and H. Caglar, "Design of efficient M-band coders with linear phase and perfect reconstruction properties," IEEE Trans. Signal Processing, vol. 43, pp. 1579–1590, July 1995.

**[14].** F. Y. Shih and S. Chen, "Adaptive document block segmentation and classification," IEEE Trans. Syst., Man, Cybern. B, vol. 26, pp. 797–802, Oct. 1996.

**[15].** M. Vetterli and J. Kovacevic, Wavelets and Subband Coding. Englewood Cliffs, NJ: Prentice-Hall, 1995, ch. 7.

**[16].** D. Wang and S. N. Srihari, "Classification of newspaper image blocks using texture analysis," Comput. Vis., Graph, Image Process, vol. 47, pp. 327–352, 1989.

**[17].** Y. Meyer, *Wavelets and Operators*. Cambridge, U.K.: CambridgeUniv. Press, 1992.
**[18].** O. J. Kwon and R. Chellappa, "Segmentation based image compression,"*Opt. Eng.*, vol. 32, pp. 1581–1587, 1993.
**[19].** M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Processing*, vol. 4, pp. 1549–1560, Nov.1995.
**[20].** A. Laine and J. Fan, "Frame representation for texture segmentation,"*IEEE Trans. Image Processing*, vol. 5, pp. 771–779, May 1996.
**[21].** T. Chang and C. C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Processing*, vol.2, pp. 42–44, Apr. 1993.