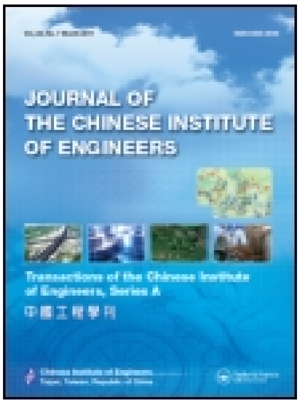


Journal of the Chinese Institute of Engineers
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcie20

Wavelet-based images compression of color document by fuzzy picture-text segmentation

Bing-Fei Wu [a], Chung-Cheng Chiu [b] & Wen-Long Lin [b]

[a] Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C. Phone: 886-3-5712122 ext. 54313 Fax: 886-3-5712122 ext. 54313 E-mail:

[b] Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C.
Published online: 03 Mar 2011.

Short Paper

WAVELET-BASED IMAGES COMPRESSION OF COLOR DOCUMENT BY FUZZY PICTURE-TEXT SEGMENTATION

Bing-Fei Wu*, Chung-Cheng Chiu, and Wen-Long Lin

ABSTRACT

In this paper, we present a new compression method for color document images based on the wavelet transform and fuzzy picture-text segmentation. This approach addresses a fuzzy picture-text segmentation method, which separates pictures and texts by using wavelet coefficients from color document images. The picture components and the text components are encoded by zerotree wavelet coding and by the modified run-length Huffman coding, respectively.

Key Words: wavelet transform; zerotree coding; fuzzy picture-text segmentation; color document image compression.

I. INTRODUCTION

A typical digital image encoder initially converts the input image data into coefficients by means of one of the transform procedures, such as DCT or FFT. The combination of discrete wavelet transform (Antonini *et al*., 1992; Mallet, 1989) and zerotree coding (Shapiro, 1993; Said and Pearlman, 1996) was proposed to compress a pure image. However, using those methods on text has produced poor results. Since the characteristics of text and pictures are different, it is not suitable to compress them by the same method like JPEG (Leger *et al*., 1991) or discrete wavelet transform. Many approaches devoted to processing monochrome documents have been proposed in the past. Wahl *et al*. (1982) designed a system for document analysis and a constrained run length algorithm (CRLA) for block segmentation. Nagy *et al*. (1986) presented a system with the *X*-*Y* tree and formal block-labeling schema to accomplish document analysis. Fletcher *et al*. (1988) proposed a robust algorithm to separate text from mixed text/graphics document images. Kamel and Zhao (1993) presented two new extraction techniques. Tsai (1985) proposed an approach to automatic threshold selection. Some other systems based on the prior knowledge of some statistical properties of various blocks (Fisher *et al*., 1990; Akiyama and Hagita, 1990; Shih *et al.*, 1992; Pavlidis and Zhou, 1992; Zlatopolsky, 1994), or texture analyses (Wang and Srihari, 1989; Jain and Bhattacharjee, 1992) have also been developed. Suen and Wang (1996) presented a text string extraction algorithm, which uses the edge-detection technique and text block identification to extract text strings. Haffner *et al*. (1998) proposed an image compression technique called "DjVu" that is specially geared toward the compression of document images in color.

In this paper, we present a compression method for color document images by using a new fuzzy picture-text segmentation algorithm.

*Corresponding author. (Tel: 886-3-5712122 ext. 54313; Fax: 886-3-5712385; Email: bwu@cc.nctu.edu.tw)

The authors are with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C.

II. THE CHARACTERISTICS OF COEFFICIENTS IN WAVELET TRANSFORM

After the second level of wavelet transform, seven frequency bands (LL2, LH2, HL2, HH2, LH1, HL1 HH1) are obtained. After the wavelet transform, the coefficients extracted from text components hold more edge information than the coefficients from picture components. Therefore, the edge feature is a good parameter for segmenting picture-text components.

The number of colors, also a useful feature, can be used for color-document image segmentation. Because the coefficients of the LL2 band are very similar to the original image after wavelet transform, the color number can be obtained by counting the color number of *UV*-planes from the coefficients of LL2 band.

The fractal dimension indicates the complexity of images. Because the picture components are more complicated than the text components, the picture components have higher fractal dimensions than the text components. The coefficients of the LL2 band are applied only to obtain the fractal dimensions from text components and picture components. In this way, the processing time to compute fractal dimensions can be reduced.

III. FUZZY PICTURE-TEXT SEGMENETATION ALGORITHM

We propose a fuzzy picture-text segmentation algorithm to separate text components and picture components from color document images. The flowchart of the algorithm is shown in Fig.1. Details of the algorithm are explained in the following subsections.

1. Spreading and Region Growing for Block Extraction

Before the process, we need to convert the coefficients of the LL2 band into bi-level data, and use the thresholding method to decide the location of foreground and background. We use the Constrained Run Length Algorithm (CRLA) to remove noise pixels. The algorithm was proposed by Wahl *et al.* (1982) to preserve the pixel when it comes from valid continuous pixels. The CRLA is performed in horizontal and vertical directions, and the bi-level images, “Mv” and “Mh”, are obtained, respectively. Then, we apply the “*OR*” operator on Mv and Mh pixel by pixel, and get a bi-level spreading image, Mhv. Therefore, the methods of thresholding, CRLA and logic operation are called the spreading process.

The steps of region growing are described below:

Step 1. Merge the pixels of image Mhv row by row.

Step 2. Compare the pixels merged from Step 1 with the current blocks. If there exists any overlap between the pixels and blocks, the pixels and blocks are merged into the same block. If there is no overlap, make a new block for the foreground pixels.

Step 3. After region growing, every block will be checked. If the block is neither growing bigger nor coming a new one, stop the block’s growth and regard it as an isolated block.

Step 4. Check if there is any overlap between blocks or not. Merge overlapping blocks into the same block.

Step 5. If Mhv comes to the last row, then go to Step 6; if not, return to Step 1.

Step 6. Change all existing blocks into isolated blocks. If there is any overlap between blocks, merge the overlapping blocks into the same block.

Step 7. Delete those smaller noise blocks.

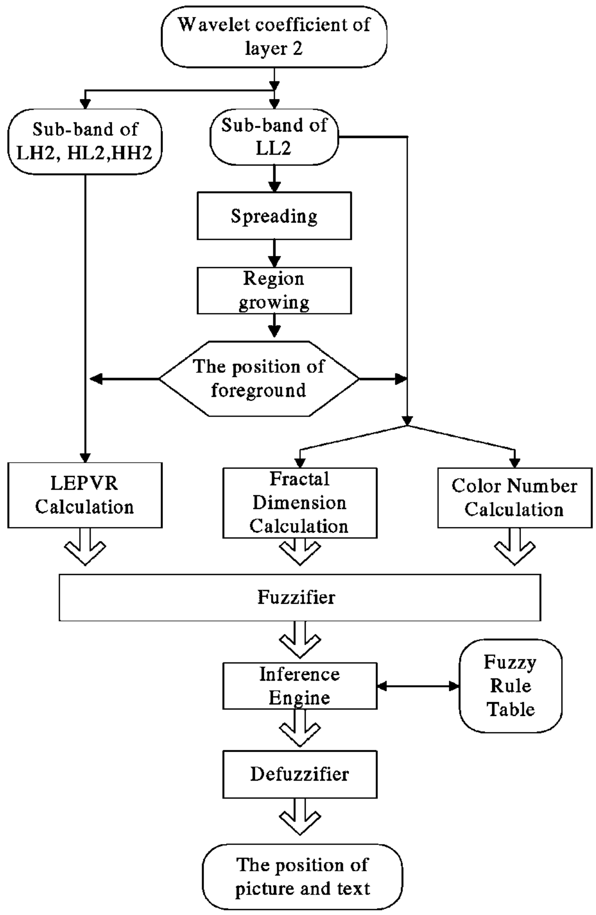


Fig. 1 The flowchart of fuzzy picture-text segmentation algorithm

2. The Calculation of Local Edge Projection Variance Ratio

The edge projection is calculated from the binary image which combines the binary images of high frequency bands (LH, HL, and HH) using the logical *OR* operator. The local edge projection variance ratio is defined by

Local edge projection variance ratio (LEPVR)

$$= \frac{1}{Mean} \times \sqrt{\sum_i (P(i) - LMean(i))^2}$$

, where $P(i)$ is the magnitude of the ith projection, *Mean* is the projection average, and

$$LMean(i) = \frac{1}{N} \times \sum_{i-N/2}^{i+N/2-1} P(k) \,,$$

N is the local average width.

3. The Calculation of Color Number

The color number is calculated in the *UV*-plane of the LL2 band as follows:

Step 1. Calculate the histogram of the *UV*-plane.

Step 2. Set LB as the threshold value, calculate effective color candidates (*ColorCandidate*) whose pixel numbers are larger than LB.

$$LB = \begin{cases} 10 & , \text{ if } 1000 < TotalPixel \\ \dfrac{TotalPixel}{100} & , \text{ if } 400 \le TotalPixel \le 1000 \\ 3 & , \; Otherwise \end{cases}$$

Step 3. Define the effective mean value (*MeanPixel*) of each color candidate as:

$$MeanPixel = \frac{TotalPixel}{ColorCandidate}$$

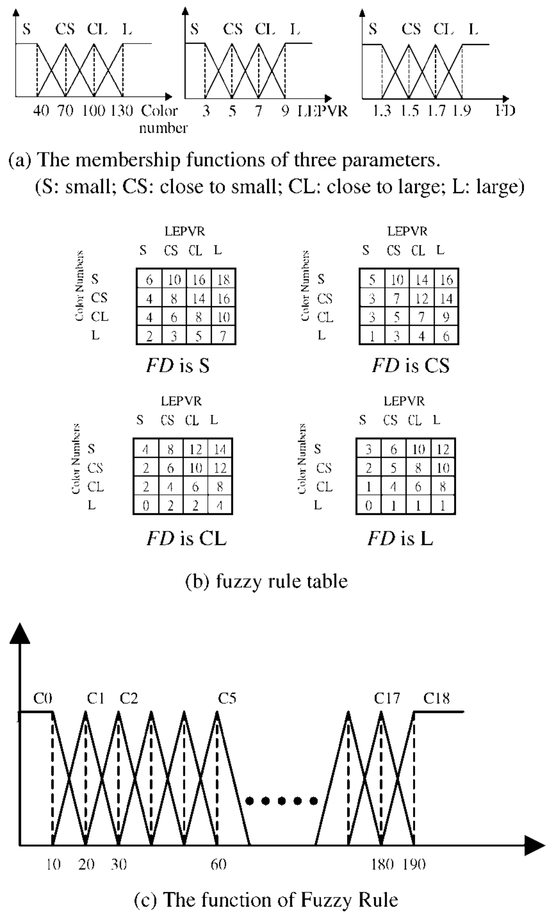
Step 4. Set *MeanPixel*/2 as the threshold value. Calculate the number of colors whose peak histogram values are larger than the threshold value. This number of colors is called *color number* in the *UV*-plane. We can calculate the color numbers of *A* and *B* in the *UV*-plane. We define the value of $A \times B$ to be the color number of the foreground block.

4. The Calculation of Fractal Dimension

We use the Two Dimension Box Counting method (Buczkowski *et al.*, 1998) which is easier and faster to calculate than others.

5. Fuzzy Logic Decision System

The fuzzy membership functions are shown in Fig. 2(a). By using Hamming Distance, and the reliability of three parameters, we set a Fuzzy Rule Table (4×4×4 cube) as shown in Fig. 2(b). The fuzzy



FD is S

Color Numbers \ LEPVR	S	CS	CL	L
S	6	10	16	18
CS	4	8	14	16
CL	4	6	8	10
L	2	3	5	7

FD is CS

Color Numbers \ LEPVR	S	CS	CL	L
S	5	10	14	16
CS	3	7	12	14
CL	3	5	7	9
L	1	3	4	6

FD is CL

Color Numbers \ LEPVR	S	CS	CL	L
S	4	8	12	14
CS	2	6	10	12
CL	2	4	6	8
L	0	2	2	4

FD is L

Color Numbers \ LEPVR	S	CS	CL	L
S	3	6	10	12
CS	2	5	8	10
CL	1	4	6	8
L	0	1	1	1

Fig. 2 Functions and rule table of fuzzy logic decision system

number of the fuzzy rule table is relative to the rule function in Fig. 2(c).

Let us now discuss the basic steps involved in the design of the fuzzy logic decision system.

Step 1. Calculate the FD, color number, and LEPVR.

Step 2. Find the probability P_S, P_{CS}, P_{CL}, and P_L for *FD*, color number, and LEPVR from Fig. 2(a). Therefore, the $P_x(FD)$, Py(color number), and P_z(LEPVR) are obtained, where x, y, and $z \in \{S, CS, CL, L\}$.

Step 3. Use the combination of $P_x(FD)$, P_y(color number), and P_z(LEPVR) to find fuzzy rules from the fuzzy rule table.

Step 4. Calculate the defuzzified value by the center of area method. The center of area method (Klir *et al.*, 1995) is selected to define the defuzzified value.

The threshold value is set at 100 : if the defuzzified value of foreground blocks is smaller than 100, it belongs to picture components; and if the defuzzified value of foreground blocks is larger than or equal to 100, then it belongs to text components.

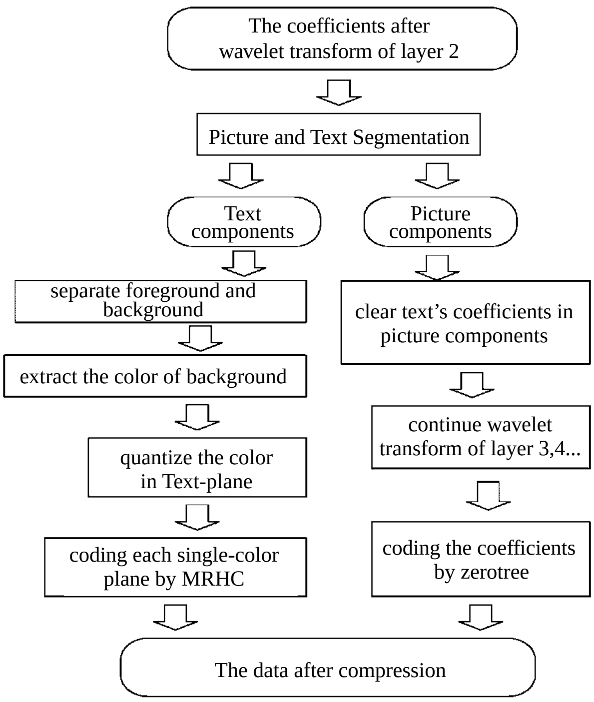


Fig. 3 The flowchart of proposed compression algorithm

IV. COLOR DOCUMENT IMAGES COMPRESSION METHOD

After the fuzzy picture-text segmentation algorithm is applied, the document images are classified into text components and picture components. In this paper, zerotree coding is used to compress the coefficients of wavelet transform for picture components. For text-components, we have to extract the colors of text from the original document image, making each color of text component a single color plane and using Modified Huffman Code to compress each single color plane. However, the run-length of blank pixels between text-lines can be very long, and may reduce the coding efficiency of Huffman Code. To solve this problem, we use Modified Run-length Code to deal with the long run-length, and Modified Huffman Code to code the short run-length. The flowchart of the proposed compression algorithm is shown in Fig. 3.

1. Color Quantization in Text Components

Text components are separated into several single-color planes. Therefore, we have to decide the number of colors in the text components by calculating the foreground histogram ($His[i]$) and then deciding the locations where the pixels are gathered. The steps are listed as follow:

Step 1. Let $His[i]$ pass the low-pass filter, and set the result as $Fhis[i]$.

Step 2. Find the maximum values of $Fhis[i]$.

Step 3. Get rid of the maximum value which is smaller than *TH_LowBound*.

Step 4. Set a threshold of duration (*TH_UV*), and calculate the local-maximum between ±*TH_UV*. Here the value of *TH_UV* is set to 15 and *TH_LowBound* is set to 10.

2. Using Modified Run-Length Huffman Code (MRLHC) to Encode Text

The text components are segmented into several single-color planes. Those planes are compressed by Modified Run-Length Huffman Code. The algorithm is designed as follows:

$$\text{Run-Length} = 1728\times \textit{Multi-code} + 64\times \textit{Makeup code} + \textit{Terminate code}$$

Run-length

$$= \begin{cases} 0\sim63, & \text{only } \textit{Terminate code} \\ 63\sim1727, & \textit{Makeup code}\text{'1}\sim\text{26'} + \textit{Terminate code} \\ 1728\sim(2^{16}\times1728), & \textit{Makeup code}\text{'27'} + \textit{Modified Run Length code} \end{cases}$$

Modified Run Length code

$$= \textit{Continue-code} + \textit{Multi-code} + \textit{Makeup code} + \textit{Terminate code}.$$

The first four bits are *Continue-code* which indicate the number of bits which can be read in *Multi-code*. By using this approach, the maximum Run-Length is represented being as long as $1728\times2^{16}=27\times2^{22}$.

3. Compression Method of Picture Components

When the blocks of text are extracted from the color document image, many gaps are located in the color document image. The boundaries of those gaps produce many high-frequency coefficients after wavelet transform. The zerotree coding algorithm will waste bits to encode those high-frequency coefficients. In order to improve the efficiency of compression, those gaps must be compensated for with appropriate coefficients. Our method is to directly compensate the text components by the average of neighboring data in the LL2 band, and set the

(a) (b) (c)

Fig. 4 (a) The original image(512×1024), (b) JPEG image(CR=104.1), (c) Proposed method (CR=112.3)

coefficients of text components in HLi, LHi, and HHi (i=1, 2) bands as zeroes. After compensating the text components in the sub-band coefficients, the wavelet transform is adopted for the coefficients of the LL2 band continuously. Then, the zerotree coding algorithm is used to encode the coefficients of wavelet transform.

V. EXPERIMENT RESULTS AND CONCLUSION

The proposed coding algorithm was simulated on Window 2000 (Pentium III 700, 128 MB RAM) with programs written in C++ language. In our study, we used the 24-bit true color image format and 200 dpi in processing. Each pixel in a 24-bit true color image is characterized by its R, G, and B color values, and 8 bits represent every value. Fig. 4 shows the comparison of the JPEG and the new compression algorithm. The time needed for performing fuzzy picture-text segmentation in Fig. 4(c) is 0.02 sec. We can find that the picture color compressed by JPEG loses very seriously, the block effect is very obviously and the text contour is blurred. Experimental results show that the compression algorithm based on fuzzy picture-text segmentation has achieved better and clearer quality of pictures and texts than JPEG.

This paper has proposed a new compression method with promising performance on color document images. The method uses different compression algorithms based on fuzzy picture-text segmentation for sub-images with different characteristics. The fuzzy picture-text segmentation algorithm is based on the coefficients of wavelet transform. It quickly finds the text components and picture components from the coefficients of wavelet transform. We have also compared our method with JPEG. The results show that the new compression method has achieved better and clearer quality than JPEG.

ACKNOWLEDGEMENTS

This work was supported by the National Science Council of the Republic of China under Grants "Human Technology (HT)-Intelligent Transportation Systems (ITS) : EX-91-E-FA06-4-4".

NOMENCLATURE

FD	fractal dimension
LEPVR	local edge projection variance ratio
MRLHC	modified run-length Huffman code

REFERENCES

Antonini, M., Barlaud, M., Mathieu, P., and Daubechies, I., 1992, "Image Coding Using Wavelet Transform," *IEEE Transation on Image Processing*, Vol. 1, No. 2, pp. 205-220.

Akiyama, M. T., and Hagita, M. N., 1990, "Automated Entry System for Printed Documents," *Pattern Recognition*, Vol. 23, No. 11, pp. 1141-1154.

Bottou, L., Haffner, P., Howard, P. G., Simard, P., Bengio, Y., and Cun, Y., 1998, "High Quality Document Image Compression with DjVu," *Journal of Electric Imaging*, Vol. 7, No. 3, pp. 410-425.

Buczkowski, S., Kyriacos, S., Nekka, F., and Ccartililer, L., 1998, "The Modified Box-Counting Method: Analysis of Some Characteristic Parameters," *Pattern Recognition*, Vol. 31, No. 4, pp. 441-418.

Fletcher, L. A., and Kasturi, R., 1988, "A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, pp. 910-918.

Fisher, J. L., Hinds, S. C., and D'amato, D. P., 1990, "A Rule-based System for Document Image Segmentation," *Proceedings of 10th IEEE international conference on Pattern Recognition*, Vol. 1, pp. 567-572.

Haffner, P., Bottou, L., Howard, P. G., Simard, P., Bengio, Y., and Cun, Y. Le, 1998, " Browsing through High Quality Document Images with DjVu," *Proceedings. IEEE International Forum on Research and Technology Advances in Digital Libraries*, pp. 309-318.

Jain, A. K., and Bhattacharjee, S., 1992, "Text Segmentation Using Gabor Filters for Automatic Document Processing," *Mach. Vis. Appl.*, Vol. 5, pp. 169-184.

Kamel, M., and Zhao, A., 1993, "Extraction of Binary Character/Graphics Images from Grayscale Document Images," *CVGIP:Graphical Models and Image Processing*, Vol. 55, No. 3, pp. 203-217.

Klir, G. J., and Yuan, B., 1995, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ.

Leger, A., Omachi, T., and Wallace, G. K., 1991, "JPEG Still Picture Compression Algorithm," *Optical Engineering*, Vol. 30, No. 7, pp. 947-954.

Mallet S, C., 1989, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, " *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674-693.

Nagy, G., Seth, S. C., and Stoddard, S. D., 1986, "Document Analysis with an Expert System," *Pattern Recognition Practice II*, pp. 149-159.

Pavlidis, T., and Zhou, J., 1992, "Page Segmentation and Classification," *CVGIP: Graph. Models Image Process.*, Vol. 54, No. 6, pp. 484-496.

Said, A., and Pearlman, W. A., 1996, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Transations on Circuits and Systems for Video Technology*, Vol. 6, No. 3, pp. 243-250.

Shapiro, J. M., 1993, "Embedded Image Coding Using Zerotrees of Wavelets Coefficients," *IEEE Transation on Signal Processing*, Vol. 41, No. 12, pp. 3445-3462.

Shih, F. Y., Chen, S. S., Hung, D. C. D., and Ng, P. A., 1992, "A Document Segmentation, Classification and Recognition System," *Proceedings of IEEE International Conference on System Integration*, pp. 258-267.

Suen, H. M., and Wang, J. F., 1996, "Text String Extraction from Images of Colour-printed Documents," *IEE Proc.-Vis. Image Signal Process.*, Vol. 143, No. 4, pp. 210-216.

Tsai, W. H., 1985, "Moment-Preserving Thresholding : A New Approach," *Computer Vision, Graphics, and Image Pressing*, Vol. 29, pp. 377-393.

Wahl, F. M., Wong, K. Y., and Casey, R. G., 1982, "Block Segmentation and Text Extraction in Mixed Text/Image Documents," *Computer Graphics and Image Processing*, Vol. 20, pp. 375-390.

Wang, D., and Srihari, S. N., 1989, "Classification of Newspaper Image Blocks Using Texture Analysis," *Computer Vision Graph. Image Process.*, Vol. 47, pp. 327-352.

Zlatopolsky, A. A., 1994, "Automated Document Segmentation," *Pattern Recognition Lett.*, Vol. 15, No. 7, pp. 699-704.

Manuscript Received: Sep. 13, 2001
Revision Received: May 20, 2002
and Accepted: Jun. 22, 2002