SUBBAND CODING OF BINARY TEXTUAL IMAGES FOR DOCUMENT RETRIEVAL[1]

Ömer N. Gerek, A. Enis Çetin

Bilkent University,
Dept. of Electrical Engineering,
Bilkent, Ankara TR-06533, Turkey
E-mail: gerek@ee.bilkent.edu.tr

Ahmed H. Tewfik

Dept. of Electrical Engineering,
University of Minnesota,
Minneapolis, MN 55455, USA

ABSTRACT

Efficient compression of binary textual images is very important for applications such as document archiving and retrieval, digital libraries and facsimile. The basic property of a textual image is the repetitions of small character images and curves inside the document. Exploiting the redundancy of these repetitions is the key step in most of the coding algorithms. In this paper, we use a similar compression method in subband domain. Four different subband decomposition schemes are described and their performances on textual image compression algorithm is examined. Experimentally, it is found that the described methods accomplish high compression ratios and they are suitable for fast database access and keyword search.

1. INTRODUCTION

Fast database search and retrieval is an essential requirement for digital document libraries. Widely used transform domain coding and adaptive predictive coding methods for image compression neither enable direct pattern matching or keyword search in the coded bit-stream nor provide high compression ratios. Compression ratios for document images can be improved by taking into account both the image characteristics and the application domain. There are a number of meth-ods proposed for document image compression and archiving [1] - [6]. Highest compression methods can be obtained using optical character recognition (OCR) methods [1]. Unfortunately they are usually not reliable and some of the document analysis applications require faithful reproductions of the original documents.

The textual image compression methods described in [2] - [6] are appropriate for fast keyword search in image databases and they can achieve compression ratios of 60:1 to 100:1.

The basic procedure for textual image compression can be described in a sequence as follows:
1) Find and extract all the characters in the image,
2) add it to the library consisting of the separate character images,
3) find the locations of the characters and remove them from the image,
4) compress (i) the constructed library and (ii) the symbol locations.
A further step is proposed in [6] to encode the residue image and in this way, lossless compression can be achieved.

In this work, the problem of textual image compression is considered in the subband domain. The subband domain characteristics of the binary textual images is suitable both for obtaining higher compression levels and for fast keyword search. Our approach is based on finding the repetitions of small character images in the subband images. The final compression ratio is higher than the method described in [6], and the time required for encod-

---

Figure 1: Repetition places of the letter "a"

ing and keyword search decreases approximately by a factor of 2^{2M}, where M is the level of subband decomposition, compared to the direct use of the textual image compression method described in [6].

2. SUBBAND TECHNIQUES FOR IMAGE COMPRESSION

In this section, the subband domain techniques used in textual image compression method is described.

A two-dimensional image is decomposed into four subband images ll, lh, hl, and hh with sizes one fourth of the original image after one level of subband decomposition [7] - [9]. Different characteristics of the subband images enable us to treat each subband image separately. In this way, one can utilize the spatial correlation and the quantization in individual subband images.

Four subband decomposition methods are used for the document image compression in this work. The first scheme is based on the Haar Wavelet Transform [9]. The good time localization property of this filter bank is suitable for the analysis of textual images which consist of sharp edges. Furthermore, the number of gray levels in subband images is not too high as compared to other linear subband decomposition filter bank structures. The Haar subband images of a binary image have 5 gray levels after one stage decomposition.

The next decomposition scheme is based on the work by Swanson and Tewfik [10]. In this work, the binary images are decomposed into bi-

nary transform images by using modulo-2 operations. This scheme shares many of the important characteristics of the real wavelet transforms. Typically, the binary wavelet transform (BWT) yields an output similar to the thresholded output of a real wavelet transform operating on the image.

The third scheme is based on the non-linear subband decomposition method of Egger et al. [11] (Fig. 2). In order not to increase the number of levels in subimages, the Galois Field - 2 (GF-2) arithmetic is used in our method. It is also shown that the GF-2 arithmetic based structure achieves perfect reconstruction (PR) [12]. This filter bank structure uses of non-linear filters instead of the standard linear filters as shown in Fig. 2. Order statistics filters (M) with appropriate regions of support and modulo-2 operations are used in this structure. This method is suitable in document analysis because of the edge preserving property.
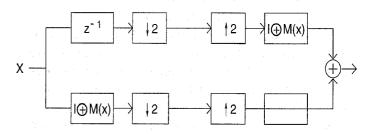

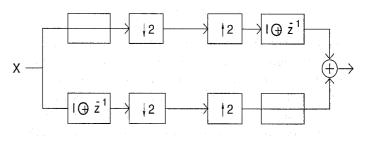Figure 2: One stage nonlinear subband decomposition with order statistics filter

Figure 3: One stage nonlinear subband decomposition with xor

The final decomposition method which is introduced in this paper also uses GF(2) arithmetic. The filtering operations perform the simple logical operation "xor" between two consecutive elements of the image data. In this scheme, the non-linear function (M) of the third scheme is replaced by a

simple z^{-1} (Fig. 3).

The lowpass synthesis filters in the filter banks in Fig 2 and 3 have the half-band property. Let $G_0(k)$ be the lowpass synthesis filter ($[I+M(x)](k)$ for the 3rd, $[I + D_{-1}](k)$ for the fourth scheme). Assume that $H(k)$ is the signal produced by filtering the down- and up-sampled signal $x(k)$ with $G_0(k)$, then the output of the filter has the following property :

$$H(2k) = c \times x(2k) \qquad (1)$$

where c is an arbitrary constant, specifically 1 for these cases.

In Fig. 4, a letter image is decomposed into binary subband images. The upper left image is shows the BWT results original image of the letter "a", the upper right shows the non-linear median subband decomposition, the lower left shows the Haar decomposition and the lower right shows the XOR filter.

3. TEXTUAL IMAGE COMPRESSION IN THE SUBBAND DOMAIN

Subband filtering is performed to obtain better compression ratios for textual images. When the subband images of the document are compressed according to the textual image compression method, four character libraries corresponding to four subband images are generated (Fig. 5). In the generation of the library, only the 'll subband image is used, and the boundary coordinates for the extracted characters are used for all four subband images. In this way, the compression time is reduced because the subband images have smaller sizes.

The total number of bits to represent the four library of symbols is smaller than the original symbol library (OSL) generated by using the method in [6] and these subband library (SL) images can be compressed more efficiently.

The efficiency of the subband domain textual compression is accomplished by making use of the high correlation between subband images. The textual images mainly consist of large regions of white and black pixels. As a result, the edges of the character images are at the same locations for

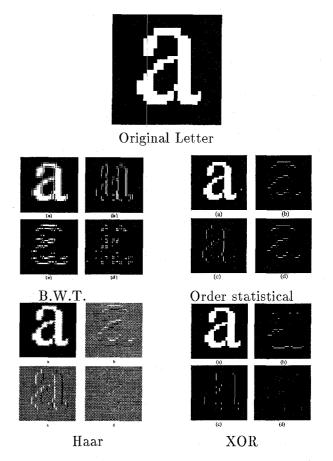all subbands. In this way, a cross-band scheme is used to achieve high compression for the OSL.

Original Letter

B.W.T.   Order statistical

Haar   XOR

Figure 4: Four different decompositions of the letter "a"

Figure 5: Symbol libraries of subband images, ll, lh, hl, and hh

A query in a digital library corresponds to a pattern search, and the pattern search can be carried out over the character library of the ll subband image. In this way, the keyword search time is reduced by a factor of 2^2 except for the Haar decomposition (because the subimages are not binary in this case). The ll image can also be used for fast preview purposes to decrease the band-

width usage. Furthermore, the time required for encoding is reduced if the textual image compression is performed in smaller size images.

4. SIMULATIONS AND CONCLUSIONS

The test document is a Times Roman font printed text with 11 points font size. This document is scanned at 300 dpi and it has a size of 2500×720 pixels. The direct use of the textual image compression procedure [6] yields a compression ratio (CR) of 63.47:1. In Haar decomposition case, the four library of symbols could be coded with a CR of 102.60:1 which shows a significant improvement over the previous methods. In the binary subband decompositions, the CR is 108.53:1 for the median filter and 105.05:1 for the xor filter. This result is even better than the result obtained by the Haar wavelet. Since the nonlinear subband decomposition yields binary images, the keyword search for the subband images is faster in the nonlinear decomposition cases. Keeping the binary property of the image is also suitable for any kind of analysis on the subband images. Furthermore, the encoding and decoding times for these operations are very small because only the logical operations are needed for the analysis, synthesis and textual coding parts.

5. REFERENCES

[1] V. K. Govindan and A. P. Shivaprasad, "Character Recognition - A Review," *Pattern Recognition*, vol. 23, no. 7, pp. 671-683, 1990.

[2] R. N. Ascher and G. Nagy, "A means for achieving a high degree of compaction on scan-digitized printed text," *IEEE Trans. Comput.*, Vol. C-23, No. 11, pp. 1174-1179, Nov. 1974.

[3] W. K. Pratt, P. J. Capitant, W. H. Chen, E. R. Hamilton, and R. H. Wallis, "Combined symbol matching facsimile data compression system," *Proc. IEEE* vol. 68, no. 7, pp. 786-796, July 1980.

[4] M. J. J. Holt, " A fast binary template matching algorithm for document image data compression," in *Pattern Recognition*, J. Kittler Ed., Berlin, Germany, Springer Verlag, 1988.

[5] O. Johnsen, J. Segen, and G. L. Cash, "Coding of two-level pictures by pattern matching and substitution," *Bell Syst. Tech. J.*, vol. 62, no. 8, pp. 2513-2545, May 1983.

[6] Ian H. Witten, Timothy C. Bell, Hugh Emberson, Stuart Inglis, and Alistair Moffat, "Textual Image Compression: Two-Stage Lossy/Lossless Encoding of Textual Images," *Proceedings of the IEEE*, Vol. 82, No.6, June 1994.

[7] E. H. Adelson, E. Simoncelli, and Hingorani, "Orthogonal pyramid transforms for image coding," *Proc. SPIE Conf. VCIP*, pp. 50-58, Cambridge, MA, 1987.

[8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transforms," *IEEE Trans. ASSP*, 1991.

[9] J. W. Woods, Ed., *Subband Image Coding*, Kluwer, 1991.

[10] M. D. Swanson and A. H. Tewfik, "A Binary Wavelet Decomposition of Binary Images," Submitted to *IEEE Trans. Image Processing* (IP-941).

[11] O. Egger, W. Li, and M. Kunt, "High Compression Image Coding Using an Adaptive Morphological Subband Decomposition," *Proc IEEE*, vol. 83, no. 2, pp.272-287, February 1995.

[12] Ömer N. Gerek, Metin Nafi Gürcan, A. Enis Çetin, "Binary Morphological Subband Decomposition For Image Coding," *IEEE Int. Symp. on Time-Frequency and Time Scale Analyis*, 1996.