University of Wisconsin – Madison

Electrical Computer Engineering

ECE533 Digital Image Processing

Embedded Zerotree Wavelet Image Codec

Team members

Hongyu Sun

Yi Zhang

December 12, 2003

Table of Contents

TABLE OF CONTENTS .... **II**

ABSTRACT **1**

I. INTRODUCTION **2**

II. TRANSFORM CODER **3**

2.1 Discrete Wavelet Transform *3*

2.2 Transform Coder *5*

III. ZEROTREE OF WAVELET COEFFICIENTS **7**

3.1 Zerotree structure *7*

3.2 Scanning of the wavelet coefficients *9*

IV. THE EZW ALGORITHM **10**

4.1 EZW algorithm features *10*

4.2 A simple example *11*

V. EXPERIMENTAL RESULTS **14**

VI. CONCLUSION **15**

VII. INDIVIDUAL CONTRIBUTIONS **16**

REFERENCE **17**

ABSTRACT

This project focuses on implementing image compression by the embedded zerotree wavelet algorithm (EZW). The EZW algorithm has been successfully applied in image compression, by which a high compression ratio can be reached while still keeping the quality of the image. The wavelet transformation first decorrelates the image samples and get transform coefficients. Then the significance map of the wavelet coefficients are quantized and encoded to symbol stream by the successive-approximation entropy-coded quantization (SAQ). The symbol stream is then represented as bit stream for transmission losslessly and efficiently.

Implementation of the EZW algorithm in this project can be considered as success. The compressed image has a remarkable smaller size compared with the original image. However, due to the limit of our algorithm, the wavelet coefficients were rounded to integers to be processed. This quantization leads to the information loss and hence a decreasing in image quality. So future work can be done to enhance the data process capacity and furthermore improve the image quality and compression ratio.

I. Introduction

Image compression methods employing wavelet transforms has been successfully implemented to provide high compression rates while maintaining good image quality. The main contribution of wavelet theory and multiresolution analysis is that it provides an elegant framework in which both anomalies (such as edges and object boundaries) and trends (areas of high statistical spatial correlation) can be analyzed on an equal footing [1]. Wavelet is a signal representation in which some of the coefficients represent long data lags corresponding to a narrow band, low frequency range, and some of the coefficients represent short data lags corresponding to a wide band, high frequency range. Using the concept of scale, data representing a continuous tradeoff between space and frequency is available.

The significance map is a binary decision indicates if a coefficient of a 2-D discrete wavelet transform has a zero or nonzero quantized value [1]. The coding of the significance map, or in other words, the positions of those coefficients that will be transmitted as nonzero values is one of the important aspects of low bit rate image coding. After quantization followed by entropy coding, the zero symbol, which occur with the most likely probability, must be extremely high in order to achieve very low bit rates. Therefore, a large fraction of the bit budget is spent on encoding the significance map. This results in a significant improvement in encoding the significance map, and hence, a higher efficiency in compression.

The successive approximation is a compact coarse-to-fine logarithmic multiprecision representation of the significant coefficients. By employing the SAQ, the EZW coder generates a representation of the image that is coarser-to-finer in both the spatial domain and frequency domain simultaneously.

II. Transform Coder

2.1 Discrete Wavelet Transform

Wavelets are generated from on single function f by dilations and translations

$$j_{s,u}(x) = 2^{s/2} j(2^s x - u)$$

Using the wavelet transform, any signal f can be represented as a superposition of wavelets.

$$f(x) = \sum_{s,u \in Z} W_{s,u} j_{s,u}(x)$$

where $W_{s,u} = \int j_{s,u}(x) f(x) dx$ is the wavelet representation of the signal.

Wavelet transform provides a hierarchical signal representation, each coefficient corresponds to a spatial area and a frequency range. It is identical to a hierarchical subband system, where the subbands are logarithmically spaced in frequency and represent octave-band decomposition. At each level, the signal can be further decomposed into a coarser approximation and a corresponding added detail.

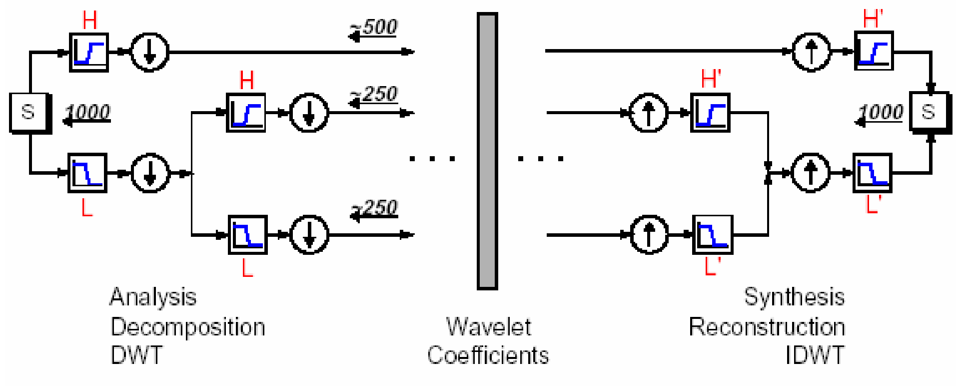

Figure 1 Multistep Analysis and Synthesis of Wavelet Coefficients Process

The coefficients in the same orientation and corresponding to the same spatial region in the image can be organized as a tree, where each parent node has four children nodes, which are in the higher frequency band corresponding to the same spatial region. Such tree-structured representation provides and efficient means for exploitation of wavelet coefficients clustered both in spatial and frequency domain.

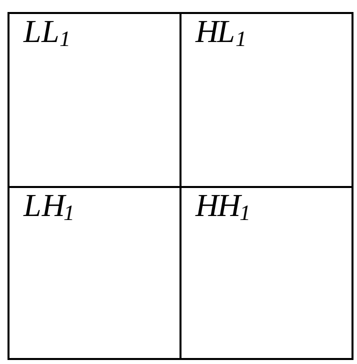

Figure 2. A One-Level Wavelet Decomposition

At the first stage of wavelet transform, the original image is divided into four subbands by separable application of vertical and horizontal filters. Each coefficient represents a spatial area corresponding to approximately a 2×2 area of the image. The low frequencies (L) represent *0<|? |<p/2*, and the high frequencies (H) represent

$p/2<|?|<p$, as shown in Figure 2. The HL_1, LH_1 and HH_1 represent the finest scale wavelet coefficients.

Then, the subband LL_1 is further decomposed by separable vertical and horizontal filters, as shown in
Figure **3**. This process is continued till the final level is reached. For example, a 256×256 image will have 8 levels of subbands. At each level, there are 3 subbands, and a remaining low frequency subband representing all coarser scales. The coefficients in higher level, or say coarser scale, represent larger spatial area but a narrower frequency band. In subband coding systems [6], the coefficients from a given subband are usually grouped together for the purposes of designing quantizers and coders.

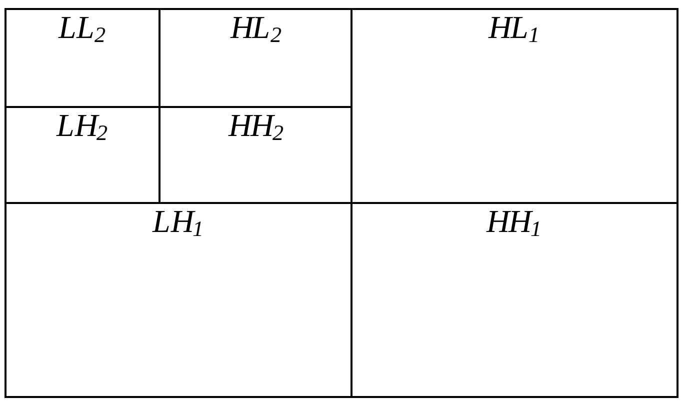

Figure 3. A Two-level Wavelet Decomposition

2.2 Transform Coder

Basically, a low-bit rate image encoder consists of three components: a transformation, a quantizer and data compression, as shown in Figure 4.

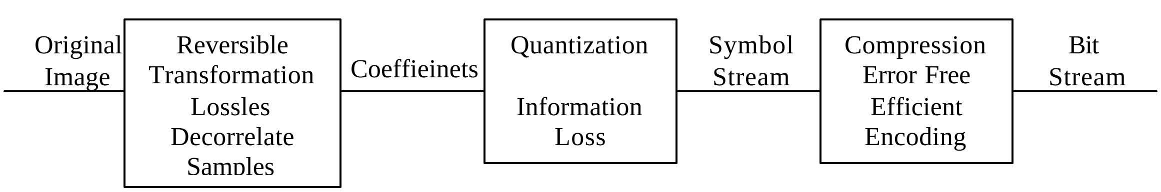

Figure 4. Transform coder

The original image is operated by some lossless transformation, generating transform coefficients. And the transformation is supposed to decorrelate the samples, i.e. the output transform coefficients are expected to be uncorrelated, or even better to be independence, though in practical this may not be exactly. The ideal situation is that after the transformation all resulting coefficients are statistically independent. If the mean is extracted and coded separately, then zero-mean, independent distributed random variables can be used to approximately model the coefficients.

The less correlated coefficients are quantized to some particular quantization bins, forming the symbol stream. The quantization step is expected to achieve small entropy of the resulting distribution of bin indexes such that the symbols can be entropy-coded at a low bit rate. This is the stage where information loss takes place, and directly affects the quality of the compressed image. An entropy code is designed based on modeling probabilities of bin indices as the fraction of coefficients in which the absolute value of a particular bin index occurs [1]. Assuming the symbols are independent with each other, the entropy of the symbols, *H*, can be expressed as

$$H = -p\log_2 p - (1-p)\log_2(1-p) + (1-p)[1 + H_{NZ}]$$

where p is the probability of a zero quantized transform coefficient, and H_{NZ} is that given nonzero values, the conditional entropy of the absolute values of the quantized coefficients. The first two terms on the right hand side of the equation are the 1st-order binary entropy of the significance map, and hence represent the fraction of encoding the significance map in the whole entropy. The last term is the conditional entropy of the distribution of nonzero values multiplied by the probability of a coefficient being nonzero, and hence represents the fraction of encoding those nonzero values of the entropy. In any case, the part of encoding the significance map always has a higher priority in the total cost of encoding, i.e. we scarify the cost of encoding nonzero values to guarantee the encoding of significance map. When the target rate decreases, the probability of zero in creases leading to the encoding cost of significance map increases. In other words, the cost of determining the positions of the few significant coefficients is the most important part of the bit budget at low rates, and its fraction in total cost will increase as the target rate decreases.

The last block of a transform encoder is data compression. In this step, the symbol stream is losslessly represented by bit stream with high efficiency. To improve the compression of significance maps of wavelet coefficients, the data structure zerotree is used in this project.

III. Zerotree of Wavelet Coefficients

3.1 Zerotree structure

The zerotree is defined as in the following way. First, a threshold T is defined and used to compare with all coefficients. A wavelet coefficient x is insignificant if $|x|<T$, and is significant if $|x|>T$. A hypothesis is that if a wavelet coefficient in a coarser scale is insignificant with respect to the given threshold value T, then all the coefficients of the same orientation in the same spatial location at finer scales are likely to be insignificant with respect to T as well [1]. Zerotree is based on this hypothesis, which is often true by empirical evidence. In a hierarchical subband system, except the finest scaled subband, all coefficients at any scale are related to a set of coefficients at the next finer scale of similar orientation. The coefficient at the coarser scale is called the parent, and all coefficients representing the same spatial area at the next finer scale of similar orientation are called children [1]. All the coefficients for a given parent at all finer scale representing the same spatial area of similar orientation form the set of descendants. Similarly, all the coefficients for a given child at all coarser scale representing the same spatial area of similar orientation form the set of ancestors. The parent-child dependencies of subbands are shown in the Figure 5 below. The arrows point from the subband of the parents to the subband of the children. The top left subband represents the lowest frequency, the coarsest scale. The bottom right subband represents the highest frequency, the finest scale. Any coefficient is an element of a zerotree if all its descendants are insignificant with respect to the given threshold value. If an element of a zerotree is not predicatably insignificant from the discovery of a zerotree root at a coarser scale at the same threshold, it is called the zerotree root.

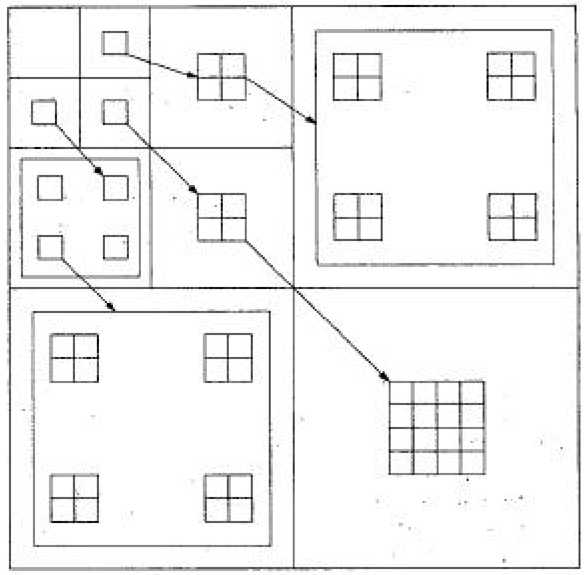

Figure 5. Parent-Child Dependencies of Subbands

3.2 Scanning of the wavelet coefficients

When scanning the coefficients, there are two principles. First, parent node is always scanned before all its children nodes. Second, all positions in the same subband re scanned before the scan goes to the next finer subband. The process starts from the lowest frequency subband LL_N, where N is the coarsest scale. Then subband HL_N LH_N and HH_N are scanned. After that, the scan goes to the next finer scale N-1, and HL_{N-1} LH_{N-1} and HH_{N-1} are scanned. The procedure is shown in Figure 6.

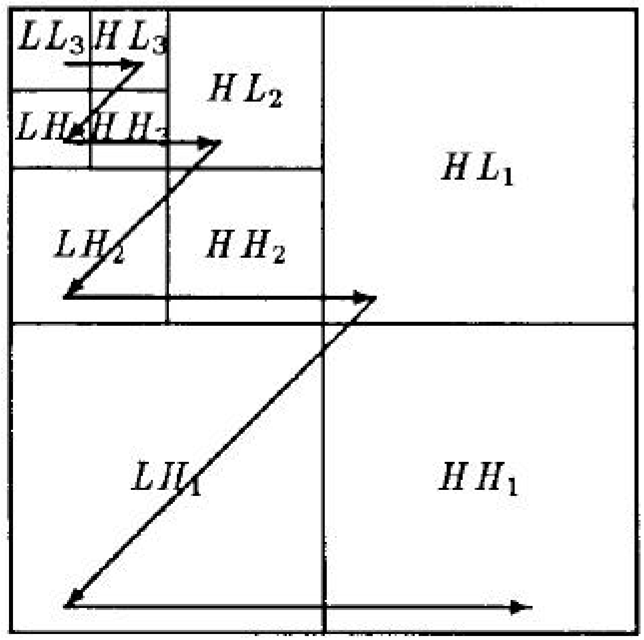

Figure 6. Scanning Order of the Subbands for Encoding a Significance Map

The significance map then can be encoded with four symbols: 1) zerotree root, 2) isolated zero, i.e. the coefficient itself is insignificant but has some significant descendants, 3) positive significant and 4) negative significant.

IV. The EZW Algorithm

4.1 EZW algorithm features

The embedded zerotree wavelet algorithm (EZW) is a simple and effective image compression algorithm. It has several notable advantages:

1) wavelet transforms,

2) zerotree coding,

3) importance defined by a decreasing sequence of uncertainty intervals,

4) adaptive arithmetic coding using small alphabets.

Its most remarkable feature is that the bits in the bit stream are generated in order of importance, yielding a fully embedded code. This embedded code provides information about the nonzero wavelet coefficients, i.e. the significance map, to distinguish the image from a null image. The encoder can terminate the encoding at any point to meet a target rate. And the decoder can stop decoding at any point in the given bit stream as well, but still recover the same image encoded at the bit rate. A simple example is explained in detail to clarify the operation in EZW algorithm.

4.2 A simple example

Suppose we've done the wavelet transform of an 8×8 image and the wavelet transform coefficients are in

Figure 7. The successive-approximation entropy-coded quantization (SAQ) method is applied to encode the coefficients. The procedure is described as following.

63	-34	49	10	7	13	-12	7
-31	23	14	-13	3	4	6	-1
15	14	3	-12	5	-7	3	9
-9	-7	-14	8	4	-2	3	2
-5	9	-1	47	4	6	-2	2
3	0	-3	2	3	-2	0	4
2	-3	6	-4	3	6	3	6
5	11	5	6	0	3	-4	4

Figure 7. Wavelet Coefficients of an 8×8 Image

The top left corner coefficient 63 is the biggest value among all coefficients and represents the DC coefficient. According to the largest coefficient magnitude, the first dominant pass threshold *T* can be any number between 31.5 and 63. Choose $T_0 = 32$. The lowest frequency coefficient with magnitude 63 is greater than the threshold and hence is a positive symbol. At the decoder end, this coefficient will be decoded to the center value of the uncertainty interval [32, 64), which is 48. Next, according to the scan order, the coefficient -34 is to be processed. Its magnitude is also greater than 32 and therefore is a negative symbol. It will be decoded to -48 in the reconstruction of the image. The coefficient of *LH*3 is -31, which is insignificant with respect to the threshold 32. But it has a descendant of magnitude 47 in *LH*1. Therefore, -31 is an isolated zero. So far, the 1st dominant pass has completed the process on the finest scale and is going to move to the 2nd scale. At *HL*2 subband, except 49 is a positive significance, all other three are zerotree root. At *LH*2 subband, all coefficients are insignificant. However, 14 has a significant child. Therefore, 14 is an isolated zero, like -31, and all other three are zerotree roots. Coefficients in *HH*2 subband are all zerotree roots since themselves and their descendants are all insignificant. Then move to the finest scale. Both *HL*1 and H*H*1 have all zeros. At L*H*1 subband, except 47 is a positive significant, all others are zeros. This completes the 1st dominant pass of the sample 8×8 image.

After the 1st dominant pass, the 1st subordinate pass is applied to the significance map. The subordinate pass only looks at the nonzero values and refine them. The uncertainty intervals are refined as [32, 48) and [48, 64). Any significant coefficients greater than 48 will be encoded as symbol "1", and as symbol "0" if lying between 32 and 48. Each symbol "1" will later be decoded as 56, which is the center of the interval [48, 64), while symbol "0" will later be decoded as 40, the center of the interval [32, 48).

During the 2nd dominant pass, the threshold value is set to be 16 since all the coefficients to be processed are those insignificant ones after the 1st dominant pass, which are between [0, 32). The procedure is the same as the 1st dominant pass and results are shown in Table 1. The 2nd dominant pass processes all the significant coefficients after the 2nd dominant pass, including those from the 1st dominant pass. The uncertainty intervals are refined as [16, 24), [24, 32), [32, 40), [40, 48), [48, 56) and [56, 64). In other words, uncertainty intervals starts from the threshold value, and has interval length 16/2 = 8. Results are shown in Table 1.

The SAQ processing continues and can cease at any time.

Table 1. Process of the 1st and 2nd Dominant and Subordinate Passes

Subband	Coefficient Value	Symbol and Reconstruction Value: 1st Dominant Pass (Threshold T=32)		1st Subordinate Pass interval center (56 40)		2nd Dominate Pass (Threshold T=16)		2nd Subordinate Pass (60 52 44 36 28 20)
LL3	63	POS	48	1	56	0		60
HL3	-34	NEG	-48	0	40	0		36
LH3	-31	IZ	0	0		NEG	-24	28
HH3	23	ZTR	0	0		POS	24	20
HL2	49	POS	48	1	56	0		52
HL2	10	ZTR	0	0		ZTR	0	0
HL2	14	ZTR	0	0		ZTR	0	0
HL2	-13	ZTR	0	0		ZTR	0	0
LH2	15	ZTR	0	0		ZTR	0	0
LH2	14	IZ	0	0		ZTR	0	0
LH2	-9	ZTR	0	0		ZTR	0	0
LH2	-7	ZTR	0	0		ZTR	0	0
HL1	7	Z	0	0		ZTR	0	0
HL1	13	Z	0	0		ZTR	0	0
HL1	3	Z	0	0		ZTR	0	0
HL1	4	Z	0	0		ZTR	0	0
HL1	-1	Z	0	0		ZTR	0	0
HL1	47	POS	48	0	40	0		44
HL1	-3	Z	0	0		ZTR	0	0
HL1	-2	Z	0	0		ZTR	0	0

V. Experimental Results

For any given wavelets coefficients as input with all elements integers, the EZW algorithm encodes the significance map and then the decoder reconstructs the coefficient matrix. The decoder output is proved to be exactly the same as the input coefficients.

The EZW coder was applied to the 256×256 grey level (8bit/pixel) test image "cameraman.tif", which is shown in Figure 8(a). The coding result has 8030 zeros and 1675 ones, totally 9705 bits. The compressed file uses a 0.138 bit/pixel. Hence the compression ratio reaches 54.0225, which is a remarkable efficiency. The peak-signal-to-noise-ratio (PSNR) is 29.5474 dB. The PSNR is calculated with the following formula:

$$PSNR = 10\log_{10}\left[\frac{MaxGreyLevel \times MN}{\sum_{xy} | g(x,y) - f(x,y) |}\right]$$

where g(x,y) is the compressed image, f(x,y) is the raw image, M is the image width, N is the image height and MaxGreyLevel is the maximum value of f(x,y).

The reconstructed image has a worse quality compared with the original one. Edges are blurred and some details are smeared. But we believe this is a reasonable quality for a high compression ratio of 54:1. Note that the codec itself did not introduce any error, but the quantization makes the compression lossy.

(a) Original image at 8 bpp **(b) Reconstructed image at 0.138bpp, 54:1 compression, PSNR=29dB**

Figure 8. Performance of EZW Coder Operating on "Cameraman.tif"

VI. Conclusion

1) The implementation of EZW algorithm is successful. Given any wavelet coefficients with all elements integer, the algorithm encodes the coefficients to bit stream with high efficiency and the decoder reconstruct the coefficients exactly the same as the original ones.
2) According to the implement of the EZW algorithm on the test image, great compression ratio is achieved with an acceptable image quality. The very high compression ratio and small PSNR value prove the success of the implementation.
3) One of the reasons that the reconstructed image quality is not so good is that due to the limit of our code, the wavelet coefficients must be integer. Therefore all the coefficients are rounded to the nearest integer and hence information is lost. This is where information loss occurs in our algorithm. Future work can be done to remove this deficiency.

VII. Individual Contributions

Tasks	Contribution in %	
	Yi Zhang	Hongyu Sun
Report	60	40
Coding	40	60
Slides	50	50
Total	50	50

Signature: __________________________ (Yi Zhang)

__________________________ (Hongyu sun)

Reference

[1] J. Shapiro, “Embedded Image Coding Using Zerotrees of Wavelet Coefficients”, *IEEE Trans. on Signal Processing*, Vol. 41, pp. 3445-3462, 1993

[2] T. C. Bell, J. G. Cleary, and I. H. Witten, *Text Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1990.

[3] J. M. Shapiro, “An Embedded Wavelet Hierarchical Image Coder,” *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, San Francisco, CA, Mar. 1992.

[4] R. A. DeVore, B. Jawerth, and B. J. Lucier, “Image Compression Through Wavelet Transform Coding,” *IEEE Trans. On Information Theory*, Vol. 38, No. 2, pp. 719-746, March 1992.

[5] A. S. Lewis and G. Knowles, “Image Compression Using the 2-D Wavelet Transform”, *IEEE Trans. on Image Processing*, Vol. 1, No. 2, pp. 244-250, April 1992

[6] J. W. Woods, Ed., *Subband Image Coding*. Boston, MA: Kluwer, 1991