Modified Embedded Zerotree Scheme for Efficient Coding of Discrete Wavelet Coded Frames

N. J. Mitra, P. K. Biswas*, T. Acharya+

Department of Electronics and Electrical Communication Engineering
IIT, Kharagpur – 721302
*pkb@ece.iitkgp.ernet.in

+Intel Corporation
5000 W. Chandler Boulevard
Chandler, Arizona 85226, USA
tinku.acharya@intel.com

Abstract

Zerotree coding is an intelligent approach for quantization of Discrete Wavelet coefficients and thus achieving image compression. Embedded Zerotree Coding by Shapiro [3] is very effective in this regard, having the property that the bits in the bit stream are generated in order of importance, yielding a fully embedded code. In this paper we have proposed a Modified Embedded Zerotree (MEZT) scheme which is computationally more efficient, making it more suitable for real time video compression applications. MEZT is a bit plane wise single pass algorithm, unlike Shapiro's two pass technique. This leads to reduction in processing time. Depending on the priority of bit budget or compression ratio, the number of bit planes to be encoded can be decided. More over, since most of the necessary tasks are bit operations, it is more suitable for VLSI implementation.

Key Words: Wavelet Coefficients, Image Compression, Zerotree Coding.

1 Introduction

In today's world of multimedia and internet, video compression poses a great challenge to both the academic community and the industry people. The huge amount of data involved in video sequences along with limited storage capacity and restricted communication bandwidth makes compression an absolute necessity. Although standard compression technique like the Discrete Cosine Transform (DCT) performs well at low compression ratios (CR), at low bit rates it has been found to introduce blocking artefacts, which tend to be visually disturbing. On the other hand, subband based coding techniques like the discrete wavelet transform (DWT) has been found to be free from this shortcoming. At low bit-rates, compression based on DWT leads to a uniform degradation of the sequences, which unlike blocking artefacts, is not visually irritating. However, like DCT, DWT images are of the same size as that of the original ones. DWT only renders the image more suited for compression by localizing the energy. So before application of entropy coding another specialized coding which, can exploit the subband hierarchy of the wavelet transform, is highly recommended. To this end, Shapiro suggested the clever approach of embedded zerotree (EZT) coding in [3] to code DWT frames. Though EZT has been found to perform well in case of still images, the involved computational complexity renders it useless for real time video applications[4]. In this paper we suggest a scheme, which has a much lower computational complexity and a comparable compression performance. Further, unlike Shapiro's EZT method, it can be implemented on a SIMD architecture thus making the scheme amenable for real time applications.

2 Brief Description of Shapiro's Algorithm

Discrete Wavelet Transform [1,2] results in decomposition of the input frame into a multi-resolution subband structure as shown in `Figure 1`. While the LL

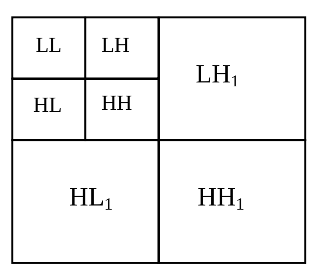

Figure 1: 2-D Wavelet Representation

US Patent Pending

subband, which retains the original image at a reduced resolution, contains most of the non-zero coefficients present in the DWT frame, all the other subbands contain only high frequency information.

Shapiro's algorithm[3] is based on the hypothesis that if at a higher subband the magnitude of a coefficient is insignificant with respect to some threshold, then there is a high probability that the magnitudes of all the coefficients at the same spatial location but at lower subbands will also be insignificant. Since the higher subbands possesses a high degree of correlation with the corresponding lower subbands, this hypothesis is largely found to hold.

The algorithm has distinct two passes: the ***dominant*** and the ***subordinate*** pass. In the first pass, the position information is coded by *significance map encoding* using a total of four symbols. In case of natural images, DWT frames contain large clusters of zeros with isolated patches of significant coefficients. Hence coding of such frames involves a significant amount of information to represent the position of the significant coefficients. In the second pass, the values of significant coefficients are coded using either 1 or 0, and refined in subsequent passes known as *successive approximation of significants*. Zerotree encoding results in an output containing some special symbols that can still be compressed by application of standard entropy coding techniques like Arithmetic Coding[5], Huffman Coding, etc.

Shapiro's EZT being embedded in nature contains all lower bit rate codes embedded in the code bit stream. This implies that the encoder can terminate the coding whenever the allocated bit budget is consumed. Similarly, the decoder can also stop decoding at any point and produce the same reconstruction possible with a lower bit rate encoding.

3 Terminology and Scanning Pattern

In the DWT generated hierarchical subband structure, a coefficient at any subband and at any level (except for the lowest level, i.e. level number 1) can be related to a set of coefficients at the lower levels at corresponding subbands. A coefficient at a higher level (coarser scale) is called the ***parent*** of all the corresponding coefficients at the same spatial orientation at lower level (finer scale), which are known as its ***children***.

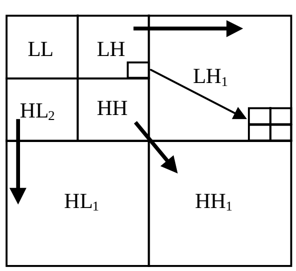

Figure 2: Parent-child Dependencies of Subbands

The parent-child relationship in the DWT hierarchy is shown in Figure 2**.** As shown in the figure, with the exception of the LL subband, any parent coefficient at any level (except the lowest level) has four children, each of which in tern also has four children and so on until the children in the lowest level are reached. The tree structure formed by a scanning order is known as ***zerotree*** with its *root* being any coefficient located at starting level (the tree shown in the figure starts from level number 2).

The scanning of the coefficients is performed in such a way so that no child is scanned before its parent. For a N-level transform, the scan proceeds in the following order: LL followed by LH_2, HL_2, HH_2 followed by subbands at the lower level i.e. LH_1, HL_1, HH_1. The scanning pattern is illustrated in Figure 3. Such a scanning pattern ensures that no child is scanned before its parent, thus ensuring maximum exploitation of the inter-subband correlation. The scanning inside each subband is done in raster scan fashion.

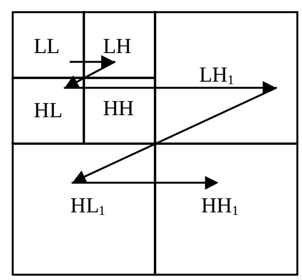

Figure 3: Zerotree Scanning Order

4 Coding Algorithm

The coding scheme begins with casting of the wavelet transformed coefficients into integers by truncation. The truncation error thus incurred can be neglected for video applications. Then the maximum-magnitude coefficient present in the DWT matrix is determined. Let it be denoted by **n**. Then the number of bits required to represent the number is given by,

$$p = \left\lceil \log_2(|n|) \right\rceil$$

The input to the next stage is the p^{th} bit plane of the wavelet transformed matrix. During the encoding process, at any bit-plane level, -1 is coded only if the coefficient being encoded is negative and its most significant bit (MSB) is being encoded. Any other bit the sign of the coefficient is not considered.

Each of the bits encountered while following the intra and inter subband scanning pattern, which has been already mentioned in the previous section, are coded using any of the four symbols as stated below. If a bit is found to be –1 (it can appear only if the above mentioned condition is satisfied), it is coded as **N** (negative). Any other significant bit, i.e. if the bit is 1, is to be coded using the symbol **P** (positive). Occurrence of 0 necessitates the checking for existence of an insignificant zerotree (IZT)* with this particular bit as its root. Such a root of IZT is to be coded using **R** (root). All the bits comprising such an insignificant IZT are marked indicating that all the necessary information has already been encoded and hence should not be scanned or coded during any of the later stages. Other occurrences of 0 are denoted by **IZ** (isolated zero). It is to be noted that for the lowest level subbands the existence of a zerotree is an impossibility and hence zeros at these levels are coded by IZ.

Once the encoding is complete for the *p*th bit level a similar scheme is followed for coding of each of (p-1)th, (p-2)th, … 1st bit-planes of the transformed matrix. Instead of going to the 1st bit plane (least significant bit plane), the coding can also be stopped at any level if the desired bit-rate is achieved. Progressive dropping of 1st, 2nd, … bit planes leads to a gradual degradation of the reconstructed image quality.

Along with these symbols, a header containing the size of the image, the number of levels of DWT along with the encoded bit plane numbers are also to be sent.

5 Decoding Algorithm

Decoding of modified EZT coded coefficients involves the performance of the same scheme but in the reverse order. The decoding is performed from the MSB to the LSB bit-planes. This results in increased resolution of the decoded data after each stage of decoding. The inter and intra subband scanning order is same as the one followed during the encoding process. P and IZ symbols mean placement of 1 or 0 at the corresponding bit-location at the correct coefficient coordinate without making any alteration to the sign of the coefficient. N not only means a similar placement of 1 but also changing the sign of the coefficient to negative. R indicates placement of zeros at corresponding bit positions of all the coefficients falling under the zerotree with that particular bit as its root. All positions in this zerotree are marked and hence ignored if encountered at a later stage while following the scanning pattern.

Both the coding and the decoding algorithm is better explained by the following example.

* A zerotree with all its elements (in this case the zerotree are composed only of bits 1, -1 or 0) as zero is termed as insignificant zerotree.

6 An Example

Let us assume that after 2 levels of wavelet transform on a 4x4 matrix the resultant matrix is of the form:

4	1	1	2
-2	0	0	1
0	3	0	0
0	1	0	0

Coding:

Step 1: The binary representation of the coefficients are as follows:

100	001	001	010
-010	000	000	001
000	011	000	000
000	001	000	000

Step 2: The maximum value is 4 and the number of bits to represent is 3. Hence, we have n=***3***.

Step 3 & 4 : The coding is done as follows:
After coding the ***3rd*** *or the* ***MSB bit-plane*** of the co-efficients generated code are

P, R, R, R.

After coding the ***2nd bit-plane*** of the coefficients the generated code are

IZ, IZ, N, R, IZ, P, IZ, IZ, IZ, P, IZ, IZ.

After coding the ***1st digits*** of the co-efficients the generated code is

IZ, P, IZ, R, P, IZ, IZ, P, IZ, P, IZ, P.

Hence the total code is obtained by the ***concatenation*** of the above sequences.

Decoding:

The decoder has been supplied with the necessary header information including size of the matrix, the number of levels of DWT and the encoded bit plane numbers.

Hence, 1st level (MSB bit plane) of decoding leads to the generation of a matrix of the form:

100	000	000	000
000	000	000	000
000	000	000	000
000	000	000	000

After ***2**nd* level of decoding the matrix is of the form:

100	000	000	010
-010	000	000	000
000	010	000	000
000	000	000	000

After ***3**rd* level (LSB bit plane) of decoding the matrix is of the form:

100	001	001	010
-010	000	000	001
000	011	000	000
000	001	000	000

7 Salient Features of The Proposed Algorithm

Unlike Shapiro's EZT algorithm, which consists of two passes for each level of encoding, the proposed scheme consists of a single pass for each bit-plane encoding. The need for maintaining a subordinate list has also been removed thus ensuring a much simpler data structure. Depending on the priority of bit budget or compression ratio, decision is taken on the number of bit-planes to be encoded. Further, the fact that most of the necessary tasks are bit-operations results in a drastic speed up in the performance. It also indicates that a VLSI implementation of this algorithm, which is more than feasible, may be very well suited for video compression applications.

8 Parallel Implementation

A simple modification of the previously described algorithm makes it suited for implementation on SIMD architecture. The sign plane of the coefficients, comprising of the signs of the coefficients, is coded as a separate bit-plane using only P and N symbols. For all the other *p* bit planes the N symbol is no longer needed as the sign information is coded separately. This implies that the bit planes require only three symbols (P, R, IZ) for their encoding and can also be done independently and hence in parallel. Codes for each bit planes also contain the corresponding bit-plane number tagged with them. The

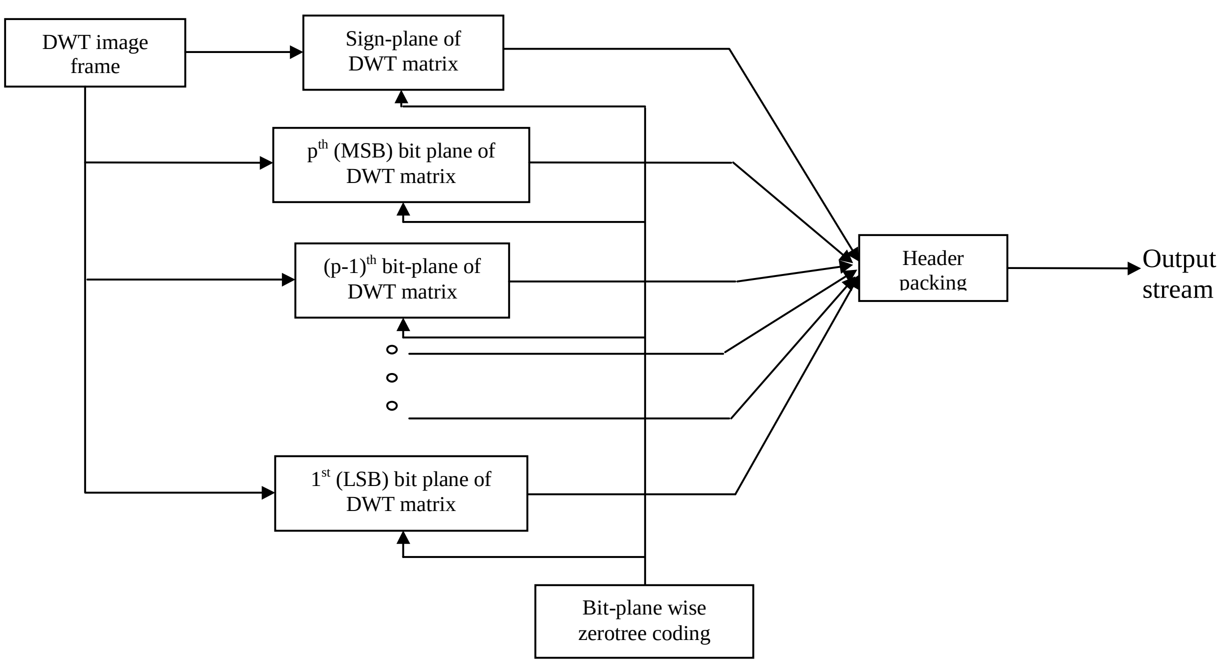

Fig. 4: Implementation of Parallel Zerotree Coding

coding scheme is illustrated in Figure 4. For the decoding part, the MSB to LSB bit-planes are decoded in parallel using the correct scaling factor depending on its bit plane number. Each of these planes are then added together and finally multiplied with the sign bit plane, which is decoded separately.

9 Results and Interpretation

The compression ratios as obtained by still grey-scale image compression are quoted in Table 1. The test images which, are the 1st frames of the named sequences, are of dimension 176*144. The results shown are after application of entropy coding viz. Arithmetic Coding[5] in this case. Some reconstructed images for the carphone sequence are shown in Figs. 5–8.

The scheme has also been tested on a zerotree based video codec. The basic design of the codec is as indicated by [4] with the EZT block being replaced by the proposed algorithm. Evaluation of the scheme on a video codec has reinforced our earlier claim of efficient encoding. Although the compression ratio has been found to be slightly less than that achieved by Shapiro's scheme, the proposed method has been found to out-perform the same when coding/decoding time is taken into consideration.

Frames from sequence	**Compression Ratio(CR)**	**Peak Signal to Noise Ratio(PSNR)**
Miss America	4.68	43.13
	7.52	39.55
	13.13	35.29
	25.60	30.90
Carphone	2.30	39.34
	3.37	36.83
	5.32	33.03
	9.50	28.54
Foreman	1.89	35.86
	2.81	34.26
	4.66	31.31
	8.80	27.48

Table 1: Performance of the Proposed Algorithm

Fig. 5 Reconstructed Carphone
CR 2.3 PSNR 39.34

Fig. 6 Reconstructed Carphone
CR 3.37 PSNR 36.83

Fig. 7 Reconstructed Carphone
CR 5.32 PSNR 36.83

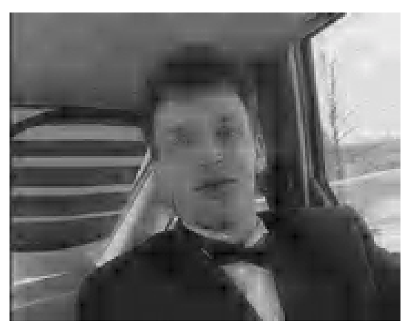

Fig. 8 Reconstructed Carphone
CR 9.50 PSNR 28.54

10 References

[1]. S.Mallat, *"A Theory for Multiresolution Signal Decomposition : The Wavelet Representation* **"** , IEEE Trans. Pattern Analysis and Machine Intell. , vol. PAMI-11, no. 7, pp. 674-693, July 1989.

[2]. Ingrid Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, 1992.

[3]. Jerome M. Shaprio**,** *"Embedded Image Coding Using Zerotrees of Wavelet Coefficients***"**, IEEE Trans. on Signal Processing, vol. 41, no. 12, pp. 3445-3459, December 1993.

[4]. Stephen A. Martucci, Iraj Sodagar, Tihao Chiang, Ya-Qin Zhang, *"A Zerotree Wavelet Video Coder"*, IEEE Trans. on Circuits and Systems for Video Technology, vol. 7, no. 1, pp. 109-118, February 1997.

[5]. Paul G. Howard, Jeffrey Scott Vitter, *"Arithmetic Coding for Data Compression"***,** Proceedings of the IEEE, vol. 82, no. 6, pp. 857-865, June 1994.