Asian Journal of Information Technology 5(6): 633-639, 2006

Image Compression Via Embedded Coder in the Transform Domain

Samai Djamel and [1]Bedda Mouldi
Department of Mechanics and Electronics, Kasdi Merbah University, Ouargla, Algeria
[1]Department of Electronics, Badji Mokhtar University, Annaba, Algeria

Abstract: Embedded Zerotree Wavelet (EZW) coder has become very popular in image compression applications, owing to its simplicity and high coding efficiency. In this paper, we first illustrate the Shapiro algorithm (EZW). In second, we combine the Discrete Cosine Transform (DCT) with an embedded zerotree quantizer in order to obtain the Xiong *et al.* algorithm (EZDCT). Our aim is to improve the bit rate gotten by EZDCT algorithm while changing the resolution level and to make a comparison with the EZW algorithm. The experiments show that the DCT-based embedded image coder gives higher Peak Signal-to-Noise (PSNR) than Joint Photographic Expert Group (JPEG) and almost similar than Shapiro's EZW coder.

Key words: Index terms-image coding, embedded coding, wavelet transform, discrete cosine transform

INTRODUCTION

Transform coding is one of the most efficient methods for image compression. In a transform based compression system two-dimensional images are transformed from the spatial domain to the frequency domain. An effective transform will concentrate useful information into a few of the low frequency transform coefficients. Human visual system is more sensitive to energy with low spatial frequency than with high spatial frequency. Therefore compression can be achieved by quantizing the coefficients so that important coefficients (low frequency coefficients) are transmitted and remaining coefficients are discarded. Very effective and popular ways to achieve compression of image data are based on Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT).

Recently the application of wavelets in image compression has received significant attention and several very efficient wavelet-based image compression algorithms have been proposed. This is due to the fact that the wavelet transform can provide a multi-resolution representation for an image or a signal with localization in both time and frequency domains. In addition, the coarse-to-fine representation of images matches the characteristics of the human visual system and makes it possible to achieve both high compression ratio and good subjective quality for the decoded image for image compression. EZW (Embedded Zerotree Wavelets)[1] uses a zerotree data structure to characterize the selfsimilarity of zeros across different scales. It can provide better image quality than DCT especially on higher compression ratio. But the implementation of the DCT is less expensive than that of the DWT. For example, the most efficient algorithm for 2-D 8x8 DCT requires only 54 multiplications[2] while the complexity of calculating DWT depends on the length of wavelet filters, which is at least one multiplication per coefficient. In this context, we give a comparative study between EZW and EZDCT (Embedded Zerotree Discrete Cosine Transform)[3] where the DCT is coupled with an embedded zerotree quantizer. In our study, we propose to vary the size of DCT blocks of 4×4 up to 64×64 instead of 8×8 only in order to have several resolution levels.

DISCRETE WAVELET TRANSFORM

The theory of continuous and discrete wavelet transforms[4,5] has inspired much basic and applied research in signal and image processing, as well as revitalizing the study of sub-band filtering[6-8]. The Discrete Wavelet Transform (DWT) is obtained by repeated filtering and sub-sampling into two bands with low- and high-pass Finite Impulse Response (FIR) filters called the analysis filters. The inverse process makes use of the synthesis FIR filters and gives perfect reconstruction if the wavelet is biorthogonal. Fig. 1 and Fig. 2 show one stage of subband decomposition/reconstruction with 2-D separable filters.

This decomposition/reconstruction derives the principle of the multiresolution introduces by[5]. A multiresolution analysis of $L^2(\Re)$ is defined as a set $M=\{V_j\}_{j\in Z}$ of embedded subspaces vector that exhibit the following properties:

1. $\{V_j\}_{j\in Z}$ is an approximation spaces sequence, i.e:
   - V_j is a closed subspace of
   - $V_j \subset V_{j-1} \quad j\in Z$

Corresponding Author: Samai Djamel, Department of Mechanics and Electronics, Kasdi Merbah University, Ouargla, Algeria

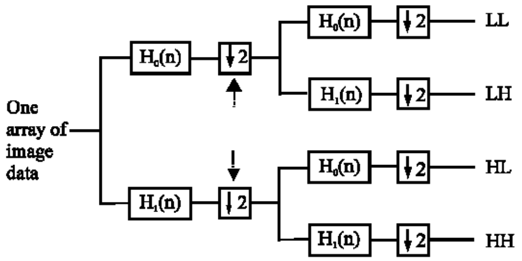

Fig. 1: One stage of subband decomposition with 2-D separable filters

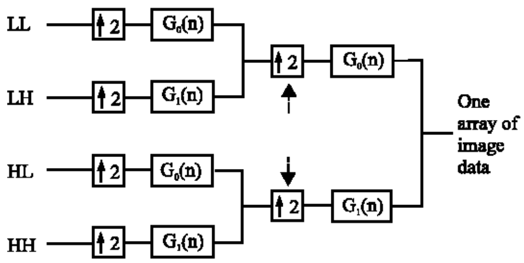

Fig. 2: One stage of subband reconstruction with 2-D separable filters

- $\bigcup_{j=-\infty}^{+\infty} V_j = L^2$ and $\bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$

2. $\forall j \in Z$, $v(t) \in V_j \Leftrightarrow v(2t) \in V_{j-1}$
3. $\exists h \in V_0$ such that $\{h(t-k)\}_{k\in Z}$ is a Riesz basis of V_0.

From the last expression, one can define a scaling function $\phi(t) \in V_0$ exists such that the set $\{\phi(t-k)\}_{k\in Z}$ is a Riesz basis of V_0.

These properties imply the existence of a sequence $\{h_k\}$ for which the scaling function satisfies the refinement equation

$$\varphi(t) = \sqrt{2}\sum_k h_k \varphi(2t-k) \qquad (1)$$

It also follows that the functions

$$\left\{\forall j \in Z, \varphi_{j,k} = 2^{-\frac{j}{2}}\varphi\left(2^{-j}t-k\right)\right\}_{k\in Z},$$

constitute a Riesz basis of V_j.

Define now W_j as complementary space of V_j in V_{j+1}, such that:

1. $V_{j-1} = V_j \oplus W_j$ with $W_j \perp V_j$.
2. $\forall j \in Z$, $w(t) \in W_j \Leftrightarrow W(2t) \in W_{j-1}$.

Consequently $\bigoplus_{j=-\infty}^{+\infty} W_j = L^2(\Re)$

A function $\psi(x)$ is a wavelet if the set of functions $\{\psi(t-k)\}_{k\in Z}$ is a Riesz basis of W_0. Since the wavelet is also an element of V_0, a sequence $\{g_k\}$ exists such that:

$$\psi(t) = \sqrt{2}\sum_k g_k \varphi(2t-k) \qquad (2)$$

The set of wavelet functions $\{\varphi_{j,k}\}_{j,k\in Z}$ is now a Riesz basis of $L^2(\Re)$. The coefficients in the expansion of a function in the wavelet basis are given by the inner product with dual wavelets $\left\{\forall j \in Z, \tilde{\psi}_{j,k} = 2^{-\frac{j}{2}}\tilde{\psi}\left(2^{-j}t-k\right)\right\}_{k\in Z}$ such that:

$$f(t) = \sum_{j,k}\left\langle f, \tilde{\psi}_{j,k}\right\rangle \psi_{j,k}(t) \qquad (3)$$

Likewise a projection on V_j is given by:

$$P_j f(t) = \sum_k \left\langle f, \tilde{\varphi}_{j,k}\right\rangle \varphi_{j,k}(t) \qquad (4)$$

Where $\left\{\forall j \in Z, \tilde{\varphi}_{j,k} = 2^{-\frac{j}{2}}\tilde{\varphi}\left(2^{-j}t-k\right)\right\}_{k\in Z}$ are the dual scaling functions. The dual functions have to satisfy the biorthogonality conditions:

$$\left\langle \varphi_{j,k}, \tilde{\varphi}_{j,k'}\right\rangle = \delta_{k-k'} \text{ and } \left\langle \psi_{j,k}, \tilde{\psi}_{j',k'}\right\rangle = \delta_{j-j'}\delta_{k-k'}$$

They satisfy refinement relations similar to (1) and (2) involving sequences $\{\hat{h}_k\}$ and $\{\breve{g}_k\}$. In case the basis functions coincide with their duals, the basis is orthogonal.

DISCRETE COSINE TRANSFORM

In JPEG-DCT method[9] 8×8 sub-block is employed because of the coding efficiency. Hence image should be divided into 8x8 sub-blocks and 8x8 two-dimensional DCT is taken for each sub-block. The Eq. (5) and (6) are the definitions of a two-dimensional discrete cosine transform and the inverse discrete cosine transforms respectively. Form these equations we can see easily that a DCT coefficient is real and the transform itself is an even function-expansion of a signal. It means that the signal is evenized implicitly. Hence from the point of view about aliasing DCT is much favorable to transform a signal.

$$F(u,v) = \frac{2C(u)C(v)}{\sqrt{MN}}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n) \cos\left\{\frac{(2m+1)u\pi}{2M}\right\}\cos\left\{\frac{(2n+1)v\pi}{2N}\right\} \qquad (5)$$

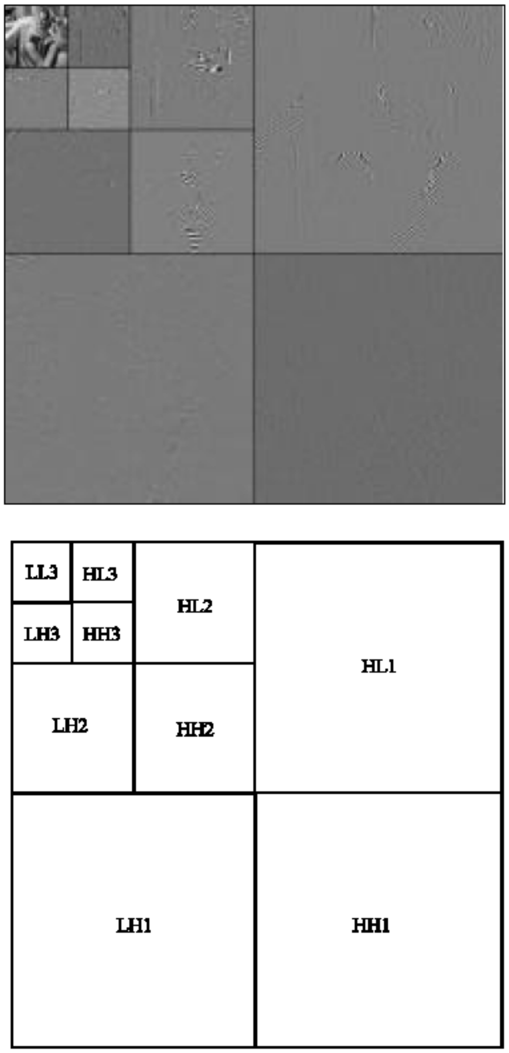

Fig. 3: Decomposition results from the third stage

$$f(m,n)=\frac{2}{\sqrt{MN}}\sum_{u=0}^{U-1}\sum_{v=0}^{V-1}C(u)C(v)F(u,v)\cos\left\{\frac{(2m+1)u\pi}{2M}\right\}\cos\left\{\frac{(2n+1)v\pi}{2N}\right\} \quad (6)$$

Where f(m,n) is the image signal and F(u,v) is the DCT coefficient.

C(k) is a normalizing coefficient as

$$C(k)=\begin{cases}1/\sqrt{2} & : \quad k=0\\ 1 & : \quad \text{otherwise}\end{cases} \quad (7)$$

The discrete cosine transform is the most widely used transform in many applications such as image transform coding, image processing and pattern recognition. This is because of its near optimum performance with respect to the redistribution of signal variance into a few low-order coefficients. The DCT is characterized by its superior energy compaction compared with other transforms.

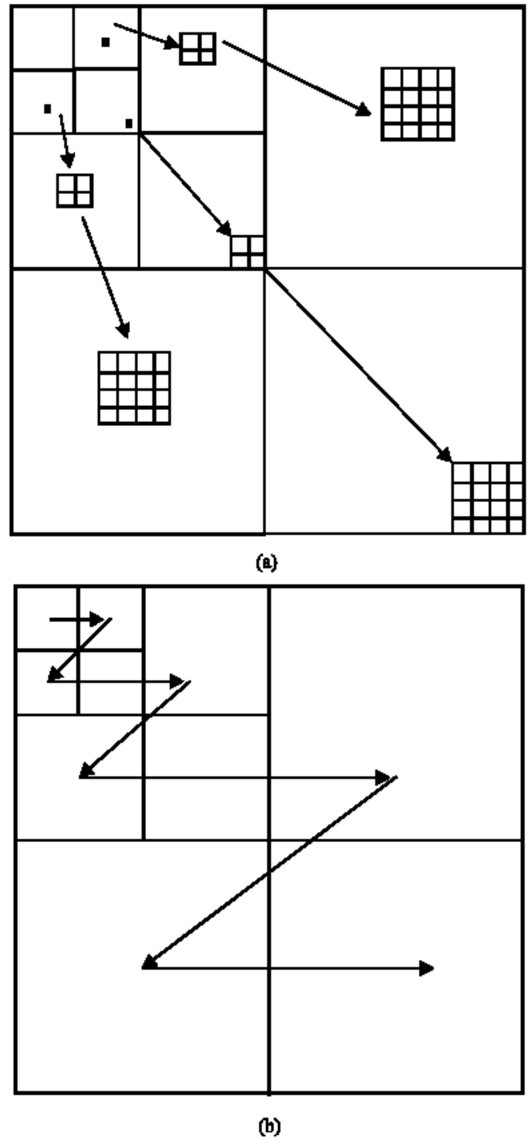

Fig. 4: (a) Structure of zerotrees and (b) Scanning order of subbands for encoding

EZW CODING SCHEME

The Embedded Zerotree Wavelet (EZW) coding scheme was proposed by Shapiro[1]. As the name suggests, this coding method has the embedding property. Essentially, there are three key elements to the EZW scheme:

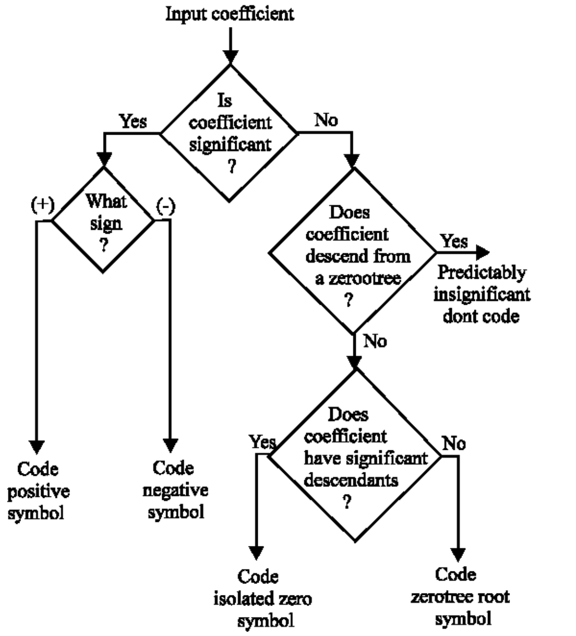

Fig. 5: Flow chart for encoding a coefficient of the significance map

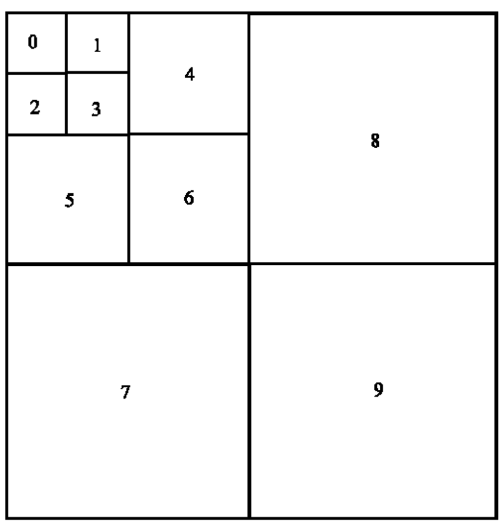

Fig. 6: 8×8 DCT block taken as three-scale tree with ten-subband decomposition

- The wavelet transform is used to form a hierarchical subband decomposition of an image.
- The absence of significant information across scales is predicted by exploiting the self–similarity inherent in images.
- Successive approximation quantization is combined with arithmetic coding[10,11] to produce the compressed bit stream.

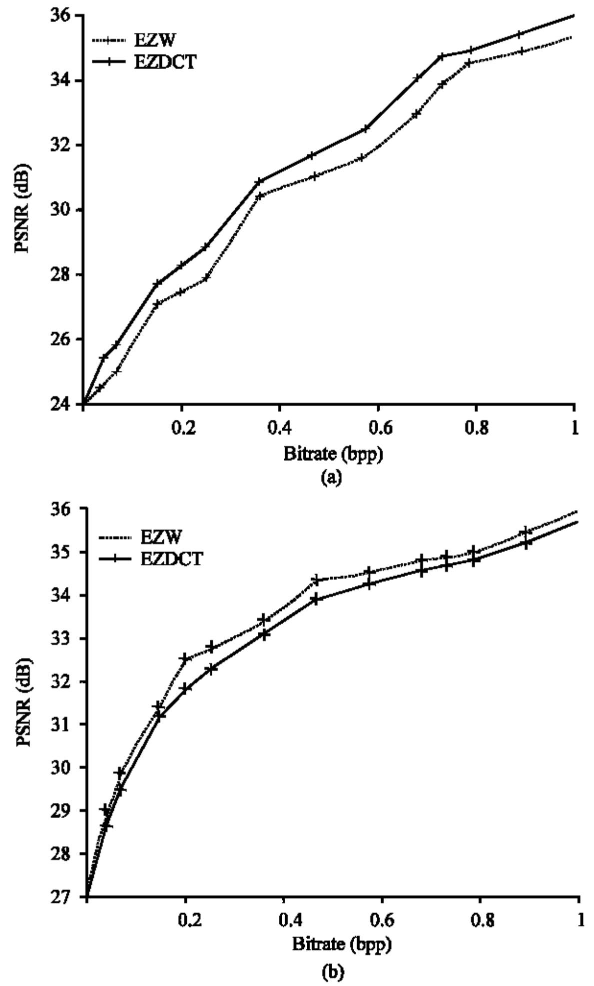

Fig. 7: Comparison results between EZW and EZDCT at scale 4 (a) Barbara (b) Peppers

The second point above is arguably the most important. The EZW scheme encodes the wavelet transform coefficients in bit significance order.

Normally, this would also require that the coefficient positions be explicitly coded. By predicting the absence of significant information across scales, however, the EZW scheme is able to avoid encoding individual coefficient positions explicitly. Not having to explicitly encode the position of each coefficient saves bits and leads to excellent compression results. The way of constructing a zerotree is as follows:

Initialization: Compute the wavelet transform of the image. As an example, Fig. 3 represents the Wavelet transform for a three stages bank filter for the image Barbara, where H represents High frequency and L represents Low frequency.

The low-frequency sub-band images are more important than the high-frequency ones. Furthermore, the coefficients of the low-pass images are large compared to the coefficients in the high-frequency images. The zerotree technique depends on this fact in designing the set of quantizers to be used at every stage in the encoding process. At the very beginning, the large coefficients are quantized by a coarse quantizer. At every next stage in the decoding process, a refinement is added to represent these coefficients more precisely in addition to represent the small coefficients that were ignored in a previous stage.

Next, we Determine the threshold 'T_0' so that it is equal to the greatest power of two lower than the maximum value of the wavelets coefficients.

$$T_0 = 2^{\left[\log_2(c_{max})\right]} \quad (7)$$

Dominant pass: We traverse the wavelets coefficients according to the way represented on Fig. 4(b), so that the zerotree are most effective possible (Fig. 4 (a)).

For each coefficient we affect one of the four following symbols by comparing them with the current threshold T_i:

Positive Significant (POS): if the absolute value of the coefficient is greater than the threshold T_i and positive sign.

Negative Significant (NEG): if the absolute value of the coefficient is greater than the threshold T_i and negative sign.

Isolated Zero (IZ): if the absolute value of the coefficient is lower than the threshold T_i having one or more descendants which are not negligible in front of T_i.

Zerotree Root (ZTR): if the value of the coefficient is lower than the threshold T_i having only negligible descendants.

The flow chart for the decisions made at each coefficient is shown in Fig. 5.

Moreover, all the descendants coefficients of a coefficient which has symbol ZTR are not coded. When the decoder receives a coefficient noted ZTR it will know that all the descendants are negligible. We transmit only for part of the coefficients the four symbols above.

Subordinate pass: The second part of the algorithm for this scan is to quantize the coefficients with symbols POS/NEG. For all the significant values given in the preceding stage, we emit the bit corresponding to 2^{i-1} to increase the precision of the transmitted significant values.

Subordinate pass is used to make the coefficients significant at the current stage negligible at the next stage to increase the chances to have zerotree.

We start again the algorithm at step B on the residue of the image by incrementing 'i' of one and by dividing the threshold 'T_i' by two. We reiterate the process until the quality standard of the image is reached or which the transferable number of bits is exceeded.

EZDCT CODING SCHEME

Typical images can be described as a set of smooth surfaces delimited by edge discontinuities. This is shown in the DCT domain by two facts.

- Signal energy due to smooth regions is compacted mostly into DC coefficients, thus resulting in negligible contributions to coefficients in the higher frequency bands.
- Due to the small compact support associated with DCT, edges can only contribute energy to a small number of AC coefficients.

The simplest form DCT-based encoder can be thought of as essentially compression of a stream of 8x8 blocks of image samples. Each 8x8 block makes its way through each processing step and yields output in compressed form into the data stream.

In[3], block-based DCT coding can treat as a depth-3 tree of coefficients. After reorganization can be further utilized to DCT-based coder in order to obtain better compression performance as EZW did in the wavelet domain. So as to improve the results obtained by the EZDCT, We opted to vary the size of the DCT blocks. So the algorithm of EZDCT is as follows:

- An input image is first partitioned into blocks, where,.
- Each block is then transformed to the DCT domain[9] and can be taken as an L-scale tree of coefficients with subbands decomposition.
- After that, the same subbands for all DCT blocks are grouped and put onto their corresponding positions. Fig. 6 demonstrates a DCT block taken as three-scale tree structure with ten-subband decomposition.
- An embedded zerotree quantizer is then applied to quantize the tree-structured DCT coefficients as was done to the wavelet coefficients in[1].

EXPERIMENTAL RESULTS AND PERFORMANCE COMPARISON

The performances of our algorithms are evaluated on several standard images. All the images are of 512×512 pixels, 8 bit/pixel. As usual, the

(a) (b) (c)

Fig. 8: Performance of the two coding algorithms at scale 4 and 0,25 bits/pixel (a) Original 512×512 8bpp Barbara image, (b) PSNR = 27,13 dB (EZW), (c) PSNR = 28,09 dB (EZDCT)

(a) (b) (c)

Fig. 9: Performance of the two coding algorithms at scale 4 and 0,125 bits/pixel (a) Original 512×512 8bpp Peppers image, (b) PSNR = 29,55 dB (EZW), (c) PSNR = 29,15 dB (EZDCT)

Table 1: PSNR results in (DB) for different scales and different rates obtained with EZW and EZDCT

		PSNR (dB)									
		EZW					EZDCT				
Scale		2	3	4	5	6	2	3	4	5	6
Barbara	0,125	22,06	24,06	24,39	24,48	24,48	22,26	24,26	25,39	25,80	25,54
Bitrate (bpp)	0,25	25,26	26,98	27,13	27,17	27,16	24,33	27,01	28,09	28,44	27,90
	0,50	29,55	30,89	31,02	31,03	31,03	27,45	30,87	31,73	31,94	31,23
	0,75	31,95	33,77	34,21	34,24	34,26	30,72	33,12	35,13	35,14	34,58
	1,00	35,25	35,67	35,80	35,81	35,82	33,22	35,69	36,51	36,63	35,83
Peppers	0,125	22,06	24,06	24,39	24,48	24,48	22,26	24,26	25,39	25,80	25,54
Bitrate (bpp)	0,25	25,26	26,98	27,13	27,17	27,16	24,33	27,01	28,09	28,44	27,90
	0,50	29,55	30,89	31,02	31,03	31,03	27,45	30,87	31,73	31,94	31,23
	0,75	31,95	33,77	34,21	34,24	34,26	30,72	33,12	35,13	35,14	34,58
	1,00	35,25	35,67	35,80	35,81	35,82	33,22	35,69	36,51	36,63	35,83
Lena	0,125	22,06	24,06	24,39	24,48	24,48	22,26	24,26	25,39	25,80	25,54
Bitrate (bpp)	0,25	25,26	26,98	27,13	27,17	27,16	24,33	27,01	28,09	28,44	27,90
	0,50	29,55	30,89	31,02	31,03	31,03	27,45	30,87	31,73	31,94	31,23
	0,75	31,95	33,77	34,21	34,24	34,26	30,72	33,12	35,13	35,14	34,58
	1,00	35,25	35,67	35,80	35,81	35,82	33,22	35,69	36,51	36,63	35,83
Goldhill	0,125	22,06	24,06	24,39	24,48	24,48	22,26	24,26	25,39	25,80	25,54
Bitrate (bpp)	0,25	25,26	26,98	27,13	27,17	27,16	24,33	27,01	28,09	28,44	27,90
	0,50	29,55	30,89	31,02	31,03	31,03	27,45	30,87	31,73	31,94	31,23
	0,75	31,95	33,77	34,21	34,24	34,26	30,72	33,12	35,13	35,14	34,58
	1,00	35,25	35,67	35,80	35,81	35,82	33,22	35,69	36,51	36,63	35,83

distortion is measured by Peak Signal-to-Noise Ratio (PSNR) defined as:

$$PSNR(dB) = 20\log_{10}\frac{255}{RMSE} \quad (8)$$

Where RMSE is the root mean-square error between the original and reconstructed images.

Table I shows EZW and EZDCT coding results with different scales and different rates. Fig. 7 compares for the Barbara and the Peppers images, the peak-SNR vs. bits per pixel results obtained by an EZW coder and an EZDCT coder when four scales dyadic are used.

Similarly, Fig. 8 and Fig. 9 show performance of the two coding algorithms for the same images and the same scale at 0,25 bpp and 0,125 bpp, respectively.

Note that the 7/9 bi-orthogonal filters in[1] are used in EZW case.

For Barbara image, EZDCT is better than EZW, while for Peppers image, EZW is slightly better than EZDCT. According the results obtained, we can say that the block-based DCT by proper reorganization and representation of its coefficients can have the similar characteristics to wavelet transform, such as energy compaction, cross-subband similarity, decay of magnitude across subband, etc. Also, if we rise the level of decomposition at four and five, the compression quality (PSNR) in will be improved by 1,18 dB and 1,4 dB, respectively for Barbara image compared with[3]. The increase of the level of decomposition to four and to five permits us to have some zerotree in more without affecting the DC components.

For all images the PSNR obtained at each bitrate is found to be better than JPEG coders[9] and our contribution also brought an improvement in comparison to[12].

Generally speaking, the DCT decomposition based image coder has comparable PSNR performance with wavelet coder, yet with lower complexity.

CONCLUSION

We presented results from a comparative study of wavelet-based and DCT-based image compression systems using objective PSNR quality measures. These two methods use an embedded zerotree structure.

In this study, we show that the block-based DCT with proper organization and representation of its coefficients can have similar characteristics to wavelet transform. The experimental results show that the DCT-based embedded image coder presents a low complexity that is better than JPEG and almost similar than Shapiro's EZW coder. DCT is capable of delivering much better performance than JPEG, just as it is for the wavelet transform.

REFERENCES

1. Shapiro, J., 1993. Embedded image coding using zerotree of wavelet coefficients. IEEE Trans. Signal Processing, 41: 3445-3463.
2. Feig, E., 1990. A Fast Scaled DCT Algorithm Proc. SPIE Image ProcessingAlgorithms and Techniques, 1244: 02-13.
3. Xiong, Z., O. Guleryuz and M. Orchard, 1996. A DCT-based embedded image coder. IEEE Signal Processing Lett., 3: 289-290.
4. Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. Comm. Pure Applied Math., 41: 909-966.
5. Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intl., 11: 674-693.
6. Smith, M.J.T. and T.P. Barnwell, 1987. A new filter bank theory for time frequency representation. IEEE Trans. Acoust. Speech and Signal Process, 35: 314-327.
7. Antonini, M., P. Mathieu and I. Daubechies, 1992. Image coding using wavelet transform. IEEE Trans. Image Process., 1: 205-220.
8. Vetterli, M. and C. Herley, 1992. Wavelets and filter banks: Theory and design. IEEE Trans. Signal Process., 40: 2207-2232.
9. Pennebaker, W.B. and J.L. Michell, 1993. JPEG still Data Compression Standard. Van Norstand Reinhold.
10. Written, I., M. Neal and J. Cleary, 1987. Arithmetic coding for data compression. Commun. ACM, 30: 520-450.
11. Moffat, A., R. Neal and I.H. Witten, 1995. Arithmetic coding revisited. In Proc. Of IEEE Data Compression Conference, pp: 202-211.
12. Zhao, D., W. Gao and Y.K. Chan, 2002. Morphological Representation of DCT Coefficients for Image Compression. IEEE Trans. Circuits Syst. Video Tech., 12: 819-823.