Discrete Cosine Transform Image Compression Using Modified Set Partitioning in Hierarchical Trees

Shen-Chuan Tai
Institute of Electrical Engineering
National Cheng Kung University
sctai@mail.ncku.edu.tw

Yen-Yu Chen
Institute of Electrical Engineering
National Cheng Kung University
miscyy@tcts1.seed.net.tw

Wen-Chien Yan
Institute of Electrical Engineering
National Cheng Kung University
ywc@rose.ee.ncku.edu.tw

Abstract

Discrete cosine transform (DCT) is widely used in many practical image/video compression systems because of its compression performance and computational efficiency. This work adopts DCT, and modified the SPIHT algorithm that designed initially for encoding the discrete wavelet transform (DWT) coefficients in order to suit to encode DCT coefficients. The algorithm represents the DCT coefficients to concentrate signal energy and proposes combination and dictator to eliminate the correlation in the same level subband for encoding the DCT-based images. The coding complexity of the proposed algorithm for DCT coefficients is just close to JPEG but the performance is higher than JPEG2000. Experimental results indicate that the proposed technique improves the quality of the reconstructed image in terms of both PSNR and the perceptual results over JPEG2000 at the same bit rate.

***Keywords*: SPIHT, JPEG2000**

1. Introduction

The transform coding is an efficient compression method and the transform includes KLT [1], DCT [2], DWT [3], etc. The DWT is the popular one and many excellent algorithms, such as EZW [4], SPIHT [5], GTW [6], etc, are proposed to encode the transform coefficients. But the computation complexity of discrete wavelet transform is more than that of discrete cosine transform. In order to improve the bit rate and reduce the computation of transform, this work adopts the DCT to transform the original image and uses the encoding method for DWT coefficients to encode the DCT coefficients. The simulation results compare with the JPEG and JPEG2000 on some ISO 400 test images. The PSNR of proposed method outperforms than JPEG and JPEG2000 at the same bit rate.

2. DCT

Initially, transform coding became popular mainly due to the introduction of the DCT, an efficient approximation to the theoretically optimal but highly complex Karhunen–Loeve transform (KLT). DCT is widely used in many practical image/video compression systems because of its compression performance and computational efficiency. DCT has been successfully selected as the first step in many coding system, such as JPEG, MEPG and H.26x, due to its good property of energy distribution in the frequency domain.

One basic DCT function (1) and its inverse function (2) are shown below.

$$F(u,v) = 4C(u)C(v)/n \times n \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} f(j,k)\cos[(2j+1)u\pi/2n] \times [(2k+1)u\pi/2n] \quad (1)$$

$$f(j,k) = C(u)C(v)/n \times n \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} F(u,v)\cos[(2j+1)u\pi/2n] \times [(2k+1)u\pi/2n] \quad (2)$$

Where $c(w) = \begin{cases} 1/\sqrt{2} & , for\ w = 0 \\ 1 & , otherwise \end{cases}$.

The computation of full frame DCT in whole image is heavy so the image is usually divided into non-lapped sub-images (8x8) for furthermore processing in many DCT based compression algorithms.

3. SPIHT

The wavelet-based image encoding algorithms considerably improve the compression rate and the visual quality, therefore many researches proposes many different methods for encoding the wavelet-based images. The SPIHT is a famous wavelet-based image coding, so describe this algorithm as follows.

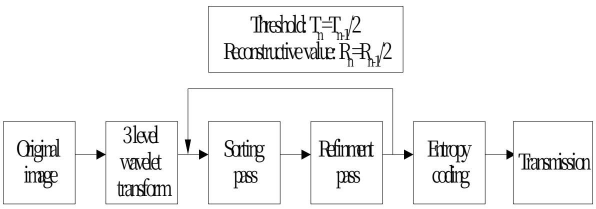

Fig. 1: Flowchart of SPIHT.

The flowchart of SPIHT is presented in Figure 1. First step, the original image is decomposed into ten subbands. Then, the method finds the maximum and the iteration number. Second step, the method puts the DWT coefficients into sorting pass that finds the significance coefficients in all coefficients and encodes the sign of these significance coefficients. Third step, the significance coefficients that be found in sorting pass are put into the refinement pass that use two bits to exact the reconstruct value for closing to real value. The front second and third steps are iterative, next iteration decreases the threshold ($T_n=T_{n-1}/2$) and the reconstructive value ($R_n=R_{n-1}/2$). Forth step, the encoding bits access entropy coding and then transmit.

4. Proposed algorithm

These algorithms for DWT coefficients, such as EZW, SPIHT, GTW, etc, are excellent in image compression, but the computing complexity of DWT is a common defect in these algorithms. Therefore, this algorithm adopts DCT, and modified the SPIHT algorithm that designed initially for encoding the DWT coefficients in order to suit to encode DCT coefficients.

(A) Representation

An input image is first partitioned into *nxn* blocks, where $n = 2^L, L > 0$, and each block is transformed into the DCT domain and can be taken as an L-scale tree of coefficients with $3xL+1$ subbands decomposition. After that, we represent DCT coefficients into a single DCT clustering entity. Fig. 2 gives an example of the representation of DCT coefficients on the Lena test image with DCT transform. Fig. 2(a) shows 8x8 DCT coefficients, and Fig. 2(b) shows the representation of 8x8 DCT coefficients.

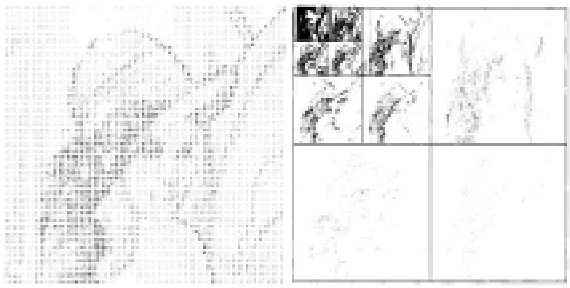

Fig. 2: (a) DCT coefficients as 8 x 8 blocks on Lena image (b) Reorganized 8x8 DCT coefficients into a single DCT clustering entity on Lena image

(B) Combination

For better compression ratio, the redundancy across subbands must be eliminated. Table 1 shows the correlation at the corresponding coordinate in LH_3, HL_3 and HH_3 in several types of ISO 400 images. The "same" condition implies that the coefficients at corresponding coordinate in LH_3, HL_3 and HH_3 have unimportant values. The "different" condition implies that the coefficients at corresponding coordinate in LH_3, HL_3 and HH_3 have at least one important value. In woman test image, this statistic shows that the percentage of insignificant coefficients in subbands (not include LL_3) is 98%. Large redundancies were hidden in these coefficients.

Table 1. Percentages of important coefficients at corresponding coordinate in LH_3, HL_3 and HH_3 for several kinds of images

Test images	Same condition	Different condition
Woman	98%	2%
Bike	89%	11%
Café	88%	12%

Table 2. Percentages of important coefficients in treenodes whose roots are at corresponding coordinate in LH_3, HL_3 and HH_3 in all recursions for several kinds of images

Test images	Same condition	Different condition
Woman	98%	2%
Bike	89%	11%
Café	87%	13%

Table 2 indicates the correlation of the corresponding coordinate in LH_1, HL_1, HH_1, LH_2, HL_2 and HH_2 in all recursions in test images. The "same" condition implies that the treenode's coefficients are unimportant on quad-trees whose roots are at corresponding coordinate in LH_3, HL_3 and HH_3. The "different" condition implies that the treenode's coefficients are at least important on a quad-tree whose roots are at corresponding coordinate in LH_3, HL_3 and HH_3. The statistic presented that the redundancy exist in the same level subbands. Therefore, The same level subband relationship is exploited here to reduce redundancy.

Based on the features that the signal energy is concentrated mostly into dc coefficients and small numbers of ac coefficients are related to the edges in spatial domain and the significant coefficients within subbands tend to be more clustered, in order to eliminate the correlation in the same level subbands to improve the compression rate, this work proposes the combination that removes the correlation in the same level subbands in all spatial orientation tree (SOT) roots. If the DCT-based and represented image is C(x,y), the all SOT tree roots in DCT-based and represented image are the coordinates in LH_3**,** HL_3**,** HH_3**.** The proposed method uses a set α to reduce the redundancy in these three subbands. The set α records which subband among LH_3, HL_3 and HH_3 has a significant coefficient. The proposed method adopts a set α to eliminate the correlation in the same level subbands.

$$\alpha=\{\,\alpha(x, y)|LH_3(x, y)\cup HL_3(x, y)\cup HH_3(x, y)\} \quad (3)$$

The set α must be sent to the decoder. If $\alpha(x, y)$=1, then the values in $LH_3(x, y)$, $HL_3(x, y)$ and $HH_3(x, y)$ would be send to the decoder. If $\alpha(x, y)$=0, nothing is sent to the decoder. Unlike by the original SPIHT, which sends all the bits of $LH_3\ (x, y)$, $HL_3\ (x, y)$ and $HH_3\ (x, y)$ to the decoder. Compared with SPIHT, The proposed method can reduce the bit rate about 0.1~ 0.2 bpp at a given PSNR (Pak signal to noise ratio).

(C) Dictator

The proposed methods applied the dictator to remove the correlation in the other subbands (in LH_2, HL_2, HH_2, LH_1, HL_1, HH_1) that are the leaf of the SOT tree. Those subbands include few significant coefficients, and the original

SPIHT algorithm suggests the use of one bit to represent whether the significant coefficient is in the quad-tree. The fact that a quad-tree includes at least one significant coefficient is represents as 1. That all of the nodes in the quad-tree are insignificant coefficients is presented as 0. The subbands originally neglected by the SPIHT algorithm neglected exhibits quite a large correlation among the same level subbands, and the proposed algorithm presents the dictator to solve this problem. According to the quad-tree concept, a correlation exists between LH_1 and LH_2. Equally the correlation exists between HL_1 and $HL_{2.}$ Equally the correction exists between HH_1 and HH_2. Therefore, $LH_{2,}$ $LH_{1,}$ $HL_{2,}$ $HL_{1,}$ HH_2 and HH_1 are divided into three partitions, P_t, $t=1, 2$, and 3.

$$Q_1=\{ LH_2 \cup LH_1 \} \quad (4)$$

$$Q_2=\{ HL_2 \cup HL_1 \} \quad (5)$$

$$Q_3=\{ HH_2 \cup HH_1 \} \quad (6)$$

The set $S_{u,}$ u=1, 2 and 3 is defined. The set S_u indicates whether that the subtree coefficients in Q_t are significant. S_1 is modified by the following conditions in the set Q_1.

$S_1(I, J)$= 1, if $LH_1(x, y)$=1, $I=\lfloor x/4 \rfloor$ and $J=\lfloor y/4 \rfloor$. (7)

$S_1(I, J)$= 1, if $LH_2(x, y)$=1, $I=\lfloor x/2 \rfloor$, and $J=\lfloor y/2 \rfloor$. (8)

$S_1(I,J)$=0,otherwise. (9)

Q_2 and Q_3 in the same steps result in S_2 and S_3. The correlation among the three sets (S_1, S_2, S_3) is greater, so the proposed algorithm creates the dictator that determines which subband has significant coefficients. The dictator d will decide what needs to be sent.

$$d=\{d(m, n) \mid S_1(m, n) \cup S_2(m, n) \cup S_3(m, n)\} \quad (10)$$

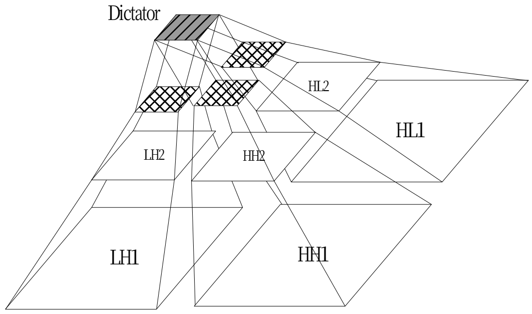

Fig. 3: The proposed algorithm uses the dictator concept and framework

Fig. 3 shows the concept and framework of the dictator. The oblique-line block is the set S_u, $u=1, 2$ *and* 3. This way saves the bits required to represent insignificant coefficients. From d, the subband with significant coefficients can be identified.

Fig. 4 presents the complete block diagram of the encoder for compressing still images. First, the test image passes through 8*8 block-DCT transform and represented into subband distribution image. Then, the proposed algorithm finds the maximum MAX that is sent to decoder, and calculates the number of recursions, RUN. Then, the proposed algorithm deals with subband LL_3 in a sorting pass that is the same as that of the SPIHT algorithm. The proposed algorithm sends A, which includes the bitmap and the sign information of significant coefficients. LH_3, $HL_{3,}$ and HH_3 subbands are handled by a combined function that reduces the interband redundancy, and then outputs B, which includes information indicates which subband has significant coefficients. The proposed algorithm deals with the other subbands using a dictator function and outputs C, which decides which should be sent. The proposed algorithm also uses the refinement pass and sends D that includes the bits to correct the reconstructed value. Finally, entropy coding is used to improve performance.

5. Simulation Results

The proposed algorithm for DCT coefficients is compared with JPEG and JPEG2000. The test images are Woman, Bike and Café, and the sizes of the others images are 2048*2560. Fig. 5(a) shows a part size of 300x500 in the Café test image that has more image context. Fig. 5(b) shows the test image decoded by JPEG at a bit rate of 0.99 bpp with a PSNR of 28.69 dB and Fig. 5(c) shows the test image decoded by JPEG2000 at a bit rate of 0.99 bpp with a PSNR of 31.96 dB. Fig. 5(d) shows the test image decoded by proposed algorithm at a bit rate of 0.99 bpp with a PSNR of 33.92 dB.

Table 3. PSNR values for JPEG standard and proposed algorithm at various bit rates in Woman test image

Bit rate (bpp)	Proposed algorithm	JPEG	JPEG2000	SPIHT
0.29	30.96	28.69	30.67	29.44
0.52	35.30	31.40	33.88	34.31
0.95	39.58	34.60	38.06	37.64

Table 4. PSNR values for JPEG standard and proposed algorithm at various bit rates in Bike test image

Bit rate (bpp)	Proposed algorithm	JPEG	JPEG2000	SPIHT
0.31	31.20	28.14	30.75	29.64
0.54	35.17	30.99	33.96	33.80
0.98	39.48	34.37	37.95	37.14

From Table 3 to Table 5, compare the PSNR of test images at various bit rates for JPEG, JPEG2000 and the proposed algorithm for DCT coefficients. At a given bit rate, Fig. 6 shows that the average rate distortion for test images is absolutely higher than the JPEG and JPEG2000 standard system.

Table 5. PSNR values for JPEG standard and proposed algorithm at various bit rates in Café test image

Bit rate (bpp)	Proposed algorithm	JPEG	JPEG2000	SPIHT
0.29	24.52	22.24	23.88	22.73
0.58	29.06	25.24	27.76	26.59
0.99	33.92	28.69	31.96	31.34

6. Conclusions

The proposed algorithm is similar to SPIHT, but the differences between the proposed algorithm and SPIHT are transform and the sorting pass. This method is to represent the DCT coefficients similarity to the subband coefficients and then encoding the DCT coefficients by our algorithm. This algorithm proposes combination and dictator to eliminate the correlation in the same level subband for encoding the DCT-based images. The coding complexity of the proposed algorithm for DCT coefficients is just close to JPEG but the performance is higher than JPEG2000. The proposed algorithm outperforms JPEG and JPEG2000 in terms of both PSNR and the perceptual results at the same bit rate.

Acknowledgement

This research is supported partially by the National Science Council of Taiwan under the contract number of NSC 92-2213-E-006-081

References

[1] *Dony, R.D., Haykin, S,* “Optimally adaptive transform coding,**”** *IEEE Transactions on Image Processing*, vol. 4 no. 10 pp. 1358

-1370, Oct. 1995.

[2] N. Ahmed, T. Natarajan and K. R. Rao, “Discrete cosine transform, ” *IEEE Transations Comput.*,” vol. C-23, pp. 90-93, Jan. 1974

[3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform,” *IEEE Transaction on Image Processing*,” vol. 1, no. 2, pp. 205-220, Apr. 1992

[4] J. M. Shapiro, “Embedded image coding using zerotrees of wavelet coefficients,” *IEEE Transaction on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[5] A. Said and W. A. Pearlman, “A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees,” *IEEE Transactions Circuits and System for Video Technology*, vol. 7,no. 3, pp. 243-250, June 1996.

[6] Edwin S. Hong and Richard E. Ladner, “Group Testing for Image Compression,” *IEEE Transactions On image processing*, vol. 11, no. 8, pp. 901-911, Aug. 2002

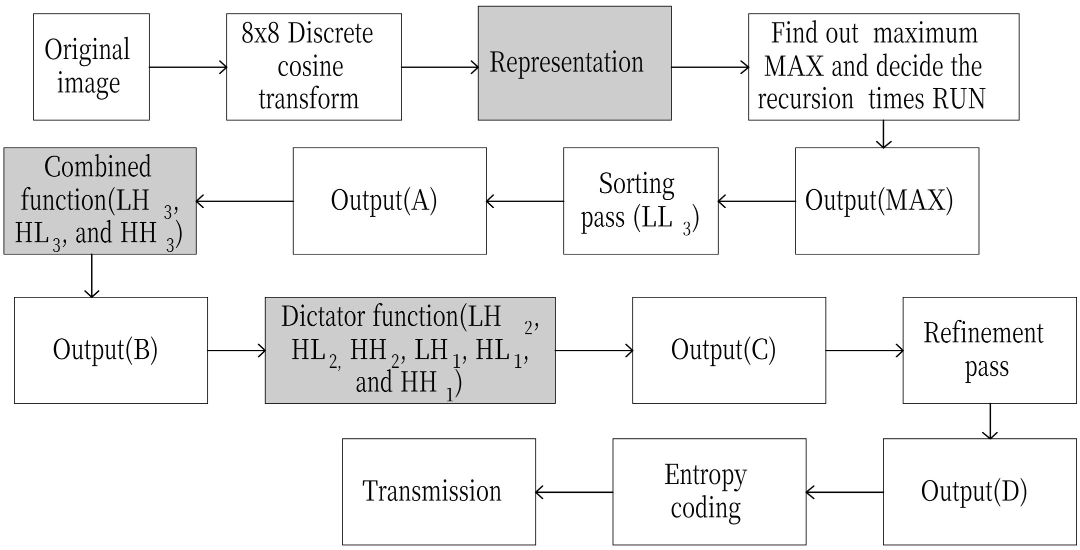

Fig. 4: Proposed algorithm for DCT coefficients flowchart

(a) (b)

(c) (d)

Fig.5: A part of café test image

(a) A part of original test image

(b) Compressed by JPEG, bit rate=0.99 bpp, PSNR =28.69 dB

(c) Compressed by JPEG2000, bit rate=0.99 bpp, PSNR =31.96 dB

(d) Compressed by proposed algorithm, bit rate=0.99 bpp, PSNR =33.92 dB

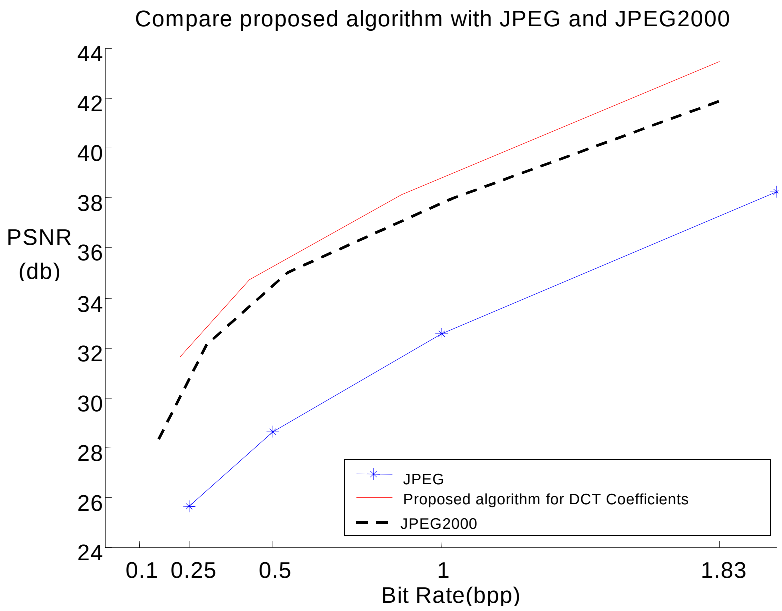

Fig. 6: Comparison proposed algorithm DCT coefficients with JPEG2000 and JPEG