PAPER

A Non-adaptive Optimal Transform Coding System

Cheng-Hsiung HSIEH[†a)], *Regular Member*

SUMMARY In this paper, a non-adaptive optimal transform (NAOT) coding system is proposed. Note that the energy-invariant property in an orthogonal transformation and that the mean squared error (MSE) of a reconstructed image is proportional to the total energy of transform coefficients discarded in the coding process. The NAOT coding system is developed and proved optimal in the sense of minimum average energy loss. Basically, the proposed coding system consists of the following steps. First, obtain the average energy image block from transform image blocks. Second, sort the average energy image block in the descending order by energy where the sorted indices are recorded. Third, specify the number of coefficients, M, to be retained in the coding process. Fourth, the first M sorted indices form a set denoted as $\mathbf{S}_M$ through which the problem of optimal feature selection in transform coding is solved. Fifth, find a fixed mask $\mathbf{A}_M$ from set $\mathbf{S}_M$ which is then used to select M significant transform coefficients in image blocks. Finally, the M selected coefficients are quantized and coded by the order as in $\mathbf{S}_M$. To verify the NAOT coding system, simulations are performed on several examples. In the simulation, the optimality and the optimal feature selection in the NAOT coding system are justified Also, the effectiveness of the proposed $\mathbf{S}_M$-based selection approach is compared with the zigzag scan used in the JPEG. For fair comparison, the JPEG is modified to code only M transform coefficients. Simulation results indicate that the performance of $\mathbf{S}_M$-based selection approach is superior or identical to the zigzag scan in terms of PSNR. Finally, the performance comparison between the NAOT coding system and the JPEG is made. It suggests that the proposed NAOT coding system is able to trade very little PSNR for significant bit rate reduction when compared with the JPEG. Or it can be said that the JPEG wastes much bit rate to improve very little PSNR on the reconstructed image, when compared with the NAOT coding system.

key words: *optimality, transform coding, image compression, coefficient selection, JPEG*

1. Introduction

The objective of image compression is to reduce required memory capacity while having acceptable visual quality in the reconstructed image. One of popular image compression schemes is called transform coding. The idea of transform coding is to transform image blocks, from spatial domain to transform domain, i.e., the information of image block is converted in its corresponding transform coefficients. Then transform coefficients of significant energies are retained and coded while the rest are set to zero. By this doing, it achieves the goal of image compression. Note that the mean squared error (MSE) in a reconstructed image is proportional to the total energy of transform coefficients discarded in the coding process. Consequently, a better image can be reconstructed through an effective approach to select significant transform coefficients.

In transform coding, it is clear that to reconstruct better image transform coefficients of large magnitudes should be selected and the others discarded. However, it generally requires large overhead indicating the coefficient selection on a block-to-block basis. Therefore, the selection of significant transform coefficients is still an active area in the field of transform coding. Up to present, several adaptive approaches to select significant transform coefficients have been proposed. Given a portion of total energy in the transform domain, Palau and Mirchandani [1] used several geometric shapes to search for the geometric zone, which contained the least number of coefficients having the specified portion of energy. Using equipotentials of energy in transform domain, Neto and Nascimento [2] proposed a modified zonal coding approach where the selection of transform coefficients was based a given signal-to-noise ratio. In [3], Crouse and Ramchandran applied the optimization technique in finding coefficient thresholds and optimal $\mathbf{Q}$-matrix used in the JPEG. In [4], Bi and et al. utilized the statistical property, cross-correlation, to choose transform coefficients. In [5], Tran and Safranek included the perceptual masking threshold model into the framework of image coding. The mask was then used in the selection of transform coefficients.

In this paper, a non-adaptive optimal transform (NAOT) coding system is proposed. The motivation of the NAOT coding system is based on the following two observations. First, the total energy of an image is invariant between spatial domain and transform domain if an orthogonal transformation is applied. Second, the MSE of a reconstructed image is proportional to the total energy of transform coefficients discarded in the coding process. By these observations, the NAOT coding system is developed. Basically, the proposed coding system consists of three stages. First, the average energy image block is obtained from transform image blocks. Second, indices of average energy image block are arranged in the descending order by energy. Then the first M elements in reordered indices form a set

Manuscript received October 26, 2001.
Manuscript revised September 10, 2002.

[†]The author is with the Department of Electronic Engineering in Chien Kuo Institute of Technology, Changhua, Taiwan 500, R.O.C.
a) E-mail: hsieh@ckit.edu.tw

denoted as $\mathbf{S}_M$ where M is the number of coefficients retained in the coding process. By set $\mathbf{S}_M$, a fixed mask $\mathbf{A}_M$ is found and used to select M significant transform coefficients in image blocks. Third, the M selected coefficients are quantized and coded by the order as in $\mathbf{S}_M$. The proposed NAOT coding system will be proved optimal in the sense of minimum average energy loss later in Sect. 3.

This paper is organized as follows: In Sect. 2, the NAOT coding system is described and an example to select significant transform coefficients is demonstrated. Besides, the advantages of the NAOT coding system and different aspects from the partition priority coding (PPC) [7] are described in this section. Next, the optimality of the NAOT coding system in the sense of minimum average energy loss is derived in Sect. 3. In Sect. 4, simulations are provided (i) to verify the theoretical results, (ii) to compare the effectiveness of the proposed selection approach with the zigzag scan used in the JPEG [8], and (iii) to compare the performance of the NAOT coding system with the JPEG. Finally, conclusions and further research are described in Sect. 5.

2. The NAOT Coding System

This section is divided into four subsections. In Sect. 2.1, the implementation steps of the NAOT coding system are described. Then an example is given in Sect. 2.2 to demonstrate how image blocks are manipulated and reconstructed in the NAOT coding system. Besides, the difference between the proposed masking approach and the quantization scheme used in the JPEG is described. Next, the advantages of the NAOT coding system are given in Sect. 2.3. Finally, different aspects of the NAOT coding system from the PPC scheme are described in Sect. 2.4.

2.1 Implementation Steps of the NAOT Coding System

When the number of transform coefficients, M, retained in the coding process is specified, the implementation steps of the NAOT coding system are described as follows.

Step 1: Input original $L \times L$ image $\mathbf{O}$.

Step 2: Divide image $\mathbf{O}$ as $\tau \times \tau$ image blocks $\{\mathbf{b}_i$, for $1 \le i \le N_b\}$ where L is a multiple of τ and $N_b = (L/\tau)^2$ is the total number of image blocks. Then $\mathbf{b}_i$ is 128-level shifted down, i.e., $\mathbf{b}_i \leftarrow \mathbf{b}_i - 128$.

Step 3: Obtain $\mathbf{B}_i = DCT\{\mathbf{b}_i\}$ where $DCT\{\cdot\}$ denotes the discrete cosine transform [8].

Step 4: Find the average energy image block $\bar{\mathbf{B}}$ as

$$\bar{\mathbf{B}} = \frac{1}{N_b}\sum_{i=1}^{N_b}\mathbf{B}_i^2 \tag{1}$$

where $\mathbf{B}_i^2 = \mathbf{B}_i . * \mathbf{B}_i$ and the operation $.*$ is the element-to-element multiplication.

Step 5: By a descending sorting, find M elements of most significant energies in $\bar{\mathbf{B}}$ and denote the indices of M selected elements as set $\mathbf{S}_M$.

Step 6: Obtain the corresponding mask of $\mathbf{S}_M$, $\mathbf{A}_M$, by setting $\bar{B}(k,l) = 1$ if $(k,l) \in \mathbf{S}_M$, and $\bar{B}(k,l) = 0$ otherwise, where $\bar{B}(k,l)$ is an element of $\bar{\mathbf{B}}$.

Step 7: By $\mathbf{A}_M$, $\mathbf{B}_i$ is modified as $\hat{\mathbf{B}}_i = \mathbf{A}_M . * \mathbf{B}_i$.

Step 8: Quantize $\hat{\mathbf{B}}_i$ as $\hat{\mathbf{B}}_i = \hat{\mathbf{B}}_i ./\mathbf{Q}$ where the operation $./$ is the element-to-element division and $\mathbf{Q}$ is a quantization matrix with appropriate dimension. Then code the selected coefficients $\hat{B}_i(k,l)$ of $\hat{\mathbf{B}}_i$ in the order of $\mathbf{S}_M$.

Step 9: Decode $\hat{B}_i(k,l)$ and reform $\hat{\mathbf{B}}_i$ as $\hat{\mathbf{B}}_i = \hat{\mathbf{B}}_i .* \mathbf{Q}$.

Step 10: By the inverse DCT (IDCT) of Step 3, find the reconstructed image block $\hat{\mathbf{b}}_i = IDCT\{\hat{\mathbf{B}}_i\} + 128$.

Step 11: Obtain the reconstructed image of $\mathbf{O}$, $\hat{\mathbf{O}}$, through $\hat{\mathbf{b}}_i$.

Step 12: Calculate the peak signal-to-noise ratio (PSNR) of $\hat{\mathbf{O}}$ as

$$\text{PSNR} = 10 \log \frac{255^2}{\text{MSE}} \tag{2}$$

where

$$\text{MSE} = \frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L}[O(i,j) - \hat{O}(i,j)]^2 \tag{3}$$

and $O(i,j)$, $\hat{O}(i,j)$ are elements of $\mathbf{O}$, $\hat{\mathbf{O}}$, respectively.

Step 13: Calculate the bit rate (bit/pixel), BR, for reconstructed image $\hat{\mathbf{O}}$ as

$$\text{BR} = \frac{B_u \times 8}{L \times L} \tag{4}$$

where B_u denotes the number of bytes used in the bit stream obtained in Step 8.

At least three points about the NAOT coding system should be pointed out. First, only several bytes are required to indicate the coefficient selection since mask $\mathbf{A}_M$ is fixed for all transform image blocks in the coding process. Therefore, the NAOT coding system is simple in the selection of transform coefficients. Second, note that $\mathbf{S}_M$ is an optimal set in the sense of minimum average energy loss, which will be proved in Sect. 3. Consequently, the reconstructed image obtained from the M selected elements in $\mathbf{S}_M$ is of minimum average energy loss. Third, since the elements of $\mathbf{S}_M$ are sorted in the descending order by energy, thus the order in $\mathbf{S}_M$ is also the significance order of the selected elements. In other words, the optimal feature selection problem is solved accordingly. This will be justified in Sect. 4. When K, for $1 \le K \le M-1$, transform coefficients need to be discarded further in the coding process, there is no need to find $\mathbf{S}_{M-K}$ but simply discard last K elements in $\mathbf{S}_M$. Note that set $\mathbf{S}_{M-K}$ is still optimal in the sense of minimum average energy loss. Consequently, the NAOT coding system is effective in the selection of transform coefficients.

2.2 Demonstration of the NAOT Coding System

The purpose of this subsection is to demonstrate the idea described in Sect. 2.1 and to clarify the difference between the proposed masking approach and the quantization scheme used in the JPEG. Suppose that the 4×4 average energy image block $\bar{\mathbf{B}}$ is found as

$$\bar{\mathbf{B}} = \begin{bmatrix} 402 & 11 & 80 & 33 \\ 31 & 45 & 3 & 59 \\ 108 & 28 & 40 & 48 \\ 28 & 123 & 7 & 157 \end{bmatrix} \tag{5}$$

If $M = 4$, it is obvious that elements (1,1), (4,4), (4,2), and (3,1) should be selected according to the energy criterion. Thus, set $\mathbf{S}_M = \mathbf{S}_4 = \{(1,1),(4,4),(4,2),(3,1)\}$ and its corresponding mask $\mathbf{A}_M = \mathbf{A}_4$ is found as

$$\mathbf{A}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \tag{6}$$

For image block $\mathbf{b}_\imath$, its quantized DCT-transformed image $\hat{\mathbf{B}}_\imath$ is given as

$$\begin{aligned} \hat{\mathbf{B}}_\imath &= \text{round}\{\mathbf{A}_4 .* \mathbf{B}_\imath ./\mathbf{Q}\} \\ &= \text{round}\left\{ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} .* \right. \\ & \begin{bmatrix} B_\imath(1,1) & B_\imath(1,2) & B_\imath(1,3) & B_\imath(1,4) \\ B_\imath(2,1) & B_\imath(2,2) & B_\imath(2,3) & B_\imath(2,4) \\ B_\imath(3,1) & B_\imath(3,2) & B_\imath(3,3) & B_\imath(3,4) \\ B_\imath(4,1) & B_\imath(4,2) & B_\imath(4,3) & B_\imath(4,4) \end{bmatrix} ./ \\ & \left. \begin{bmatrix} Q(1,1) & Q(1,2) & Q(1,3) & Q(1,4) \\ Q(2,1) & Q(2,2) & Q(2,3) & Q(2,4) \\ Q(3,1) & Q(3,2) & Q(3,3) & Q(3,4) \\ Q(4,1) & Q(4,2) & Q(4,3) & Q(4,4) \end{bmatrix} \right\} \\ &= \text{round}\left\{ \begin{bmatrix} \hat{B}_\imath(1,1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hat{B}_\imath(3,1) & 0 & 0 & 0 \\ 0 & \hat{B}_\imath(4,2) & 0 & \hat{B}_\imath(4,4) \end{bmatrix} \right\} \end{aligned} \tag{7}$$

where the matrix operator $.*$, $./$, and round$\{\cdot\}$ denote as the element-to-element multiplication, division, and the rounding function, respectively. That is, $\hat{B}_\imath(k,l) = A_4(k,l)B_\imath(k,l)/Q(k,l)$ where $A_4(k,l)$, $\hat{B}_\imath(k,l)$, and $Q(k,l)$ are elements of $\mathbf{A}_4$, $\hat{\mathbf{B}}_\imath$, and $\mathbf{Q}$, respectively. By set $\mathbf{S}_4$, quantized coefficients $\hat{B}_\imath(1,1)$, $\hat{B}_\imath(4,4)$, $\hat{B}_\imath(4,2)$, and $\hat{B}_\imath(3,1)$ are coded in order. After decoding coefficients $\hat{B}_\imath(k,l)$ and restoring $\hat{\mathbf{B}}_\imath$ through $\mathbf{S}_4$, the reconstructed image block $\hat{\mathbf{b}}_\imath$ is then found by $\hat{\mathbf{b}}_\imath = IDCT\{\hat{\mathbf{B}}_\imath .* \mathbf{Q}\}$. The reconstructed image $\hat{\mathbf{O}}$ is then found through $\hat{\mathbf{b}}_\imath$.

The main difference between the proposed selection approach and the quantization scheme used in the JPEG is described as follows. As shown in (6), the mask $\mathbf{A}_4$ indicates the four significant coefficients which are then quantized by $\mathbf{Q}$ and coded in order by set $\mathbf{S}_4$. That is, only the coefficients selected by mask $\mathbf{A}_4$ need to be coded. All other coefficients are discarded and therefore waste no operation and bit rate in the coding process. However, the quantized DCT-transformed image $\hat{\mathbf{B}}_\imath$ in the quantization scheme used in the JPEG is obtained as $\hat{\mathbf{B}}_\imath = \text{round}\{\mathbf{B}_\imath./\mathbf{Q}\}$. By the zigzag scan, elements of $\hat{\mathbf{B}}_\imath$ are coded. In general, the quantization scheme may waste some operations and bit rate on the insignificant coefficients in the coding process. Thus the quantization scheme in the JPEG is less efficient and effective than the proposed selection approach. This will be verified in Sect. 4.

2.3 Advantages of the NAOT Coding System

There are at least two advantages in the proposed coding approach. First, the NAOT coding system is non-adaptive. For an adaptive coding system, generally it is required to perform classification on image blocks. By the class index, transform coefficients are selected and coded. Reference [6] is an example. A vital problem in an adaptive coding approach is the large overhead for storing class indices. For example, an $L \times L$ image, with block size $\tau \times \tau$, classified into K classes requires $(L/\tau)^2 \times \lceil \log_2 K \rceil$ bits to indicate the class indices where $\lceil\cdot\rceil$ is the ceiling function. In the case of $L = 512$, $\tau = 8$, and $K = 4$, the overhead is $4{,}096 \times 2$ bits or 1,024 bytes. However, the proposed NAOT coding system requires only $M \times \lceil \log_2(\tau \times \tau) \rceil$ bits. In the case of $M = 24$ and $\tau = 8$, a 512×512 image needs only 24×6 bits or 18 bytes to indicate the selected coefficients which is far less than that in the adaptive coding system just described. The saving of overhead is the first advantage of the NAOT coding system over adaptive one.

Second, the proposed NAOT coding system is optimal in the sense of minimum average energy loss. In other words, transform coefficients can be optimally selected by the $\mathbf{S}_M$-based approach in the NAOT coding system. However, this is not possible for the JPEG which is also a non-adaptive coding scheme. The coefficient selection in the JPEG depends heavily on the property of energy compaction [8] from which the zigzag scan is developed. Since transform coefficients with lower frequencies generally have larger magnitudes than those with higher frequencies, it implies that most of energy is compacted in the lower frequencies. Thus the order of zigzag scan is from lower frequencies to higher frequencies. Though the zigzag scan performs well in most of cases, however for a given image it is not optimal in any sense and the scan order is not the significance order of coefficients in general. On the contrary, the $\mathbf{S}_M$-based selection approach in the NAOT coding system is optimal and retains the significance order of transform coefficients. By reordering the transform coefficients based on $\mathbf{S}_M$, the energy of coefficients is compacted in the descending order in an average sense. If the number of selected coefficients is assigned, the $\mathbf{S}_M$-based selection approach is able to optimally select the significant transform coefficients while the zigzag scan generally fails to. Consequently, it is expected that the NAOT coding system outperforms the JPEG with equal number of selected coefficients in terms of PSNR. Though the JPEG needs no overhead while the proposed NAOT coding system requires several bytes, yet it is worthy to trade the little overhead with the optimal coefficient selection. These advantages will be verified in Sect. 4.

2.4 Different Aspects of the NAOT Coding System from the PPC Scheme

The idea to reorder transform coefficients by magnitude and to code them in order is not new. Several approaches have been reported. One of them is the partition priority coding (PPC) [7]. Given an image of size $L \times L$, the coefficient selection in the proposed NAOT coding system is different from the PPC scheme at least in two aspects. First, the PPC scheme indicates the coefficients of whole transform image which is formed by transform image blocks. By a set of partitions, coefficients in the entire transform image are ordered by magnitude. The total number of coefficients, after PPC, is $L \times L$. However, the proposed selection approach in the NAOT coding system sorts the average energy image block obtained from transform image blocks. Then set $\mathbf{S}_M$ is found through which a fixed mask $\mathbf{A}_M$ is acquired and applied to select M coefficients in each transform image block. Consequently, the proposed selection approach is on a block-to-block basis whereas the PPC scheme is based on entire transform image. The total number of selected coefficients is $M \times N_b$ after the proposed selection approach where N_b is the number of image blocks. Second, the PPC scheme requires no overhead to indicate the locations of coefficients. To get rid of overhead, the order of coefficients in the PPC scheme is not in a descending order. That is, coefficients in a given partition are ordered in a natural order from left to right. Though the proposed selection approach in the NAOT coding system needs several bytes to indicate the selected coefficients, it, however, is worthy to trade little overhead with the optimal order of coefficients in the sense of average energy. This guarantees the reconstructed image obtained from the NAOT coding system is optimal in the sense of minimum average energy loss. The opitmality in the NAOT coding system is proved in the following.

3. Optimality in the NAOT Coding System

The optimality, in the sense of minimum average energy loss, in the NAOT coding system is derived here. Suppose the original image $\mathbf{O}$ is of size $L \times L$ and is partitioned into $\tau \times \tau$ image blocks, $\mathbf{b}_i$, where L is a multiple of τ. The derivation of optimality is given as follows. Note that an orthogonal transformation like DCT is of the energy-invariant property. The total energy of image $\mathbf{O}$ can be shown as

$$\sum_{i=1}^{N_b}\sum_{m=1}^{\tau}\sum_{n=1}^{\tau} b_i^2(m,n) = \sum_{i=1}^{N_b}\sum_{k=1}^{\tau}\sum_{l=1}^{\tau} B_i^2(k,l) \tag{8}$$

where N_b is the total number of image blocks, $b_i(m,n)$ is an element of $\mathbf{b}_i$ and $B_i(k,l)$ is an element of transformed $\mathbf{b}_i$, $\mathbf{B}_i$. Pre-dividing both sides of (8) by $1/N_b$, we have the average energy of $\mathbf{O}$, E, as

$$\begin{aligned} E &= \sum_{m=1}^{\tau}\sum_{n=1}^{\tau}\left[\frac{1}{N_b}\sum_{i=1}^{N_b} b_i^2(m,n)\right] \\ &= \sum_{k=1}^{\tau}\sum_{l=1}^{\tau}\left[\frac{1}{N_b}\sum_{i=1}^{N_b} B_i^2(k,l)\right] \\ &= \sum_{k=1}^{\tau}\sum_{l=1}^{\tau}\bar{B}(k,l) \end{aligned} \tag{9}$$

where

$$\bar{B}(k,l) = \frac{1}{N_b}\sum_{i=1}^{N_b} B_i^2(k,l) \tag{10}$$

In (10), $\bar{B}(k,l)$ is an element of $\bar{\mathbf{B}}$ which is the average energy of the (k,l) element in $\mathbf{B}_i$, for $1 \le i \le N_b$.

With elements $\bar{B}(k,l)$, the way to find the optimal set of elements in $\mathbf{B}_i$, in the sense of minimum average energy loss, is given in the following. Note that in (9) the average energy of original image $\mathbf{O}$ is related to the sum of $\bar{B}(k,l)$, for $1 \le k,l \le \tau$. When the number of transform coefficients kept, M, is specified, the M most significant elements in $\bar{\mathbf{B}}$ are selected in the NAOT coding system. Denote the indices of M selected elements as set $\mathbf{S}_M$ through which the mask $\mathbf{A}_M$ is found as described in Sect. 2.2. By set $\mathbf{A}_M$, the average energy of reconstructed image $\hat{\mathbf{O}}_M$, E_M, is then given as

$$\begin{aligned} E_M &= \sum_{k=1}^{\tau}\sum_{l=1}^{\tau}\left[\frac{1}{N_b}\sum_{i=1}^{N_b} A_M(k,l)B_i^2(k,l)\right] \\ &= \sum_{(k,l)\in\mathbf{S}_M}\bar{B}(k,l) \end{aligned} \tag{11}$$

where $A_M(k,l)$ is an element in $\mathbf{A}_M$. By (9) and (11), the average energy loss of $\mathbf{O}$, $\tilde{E}_M$, is then given as

$$\begin{aligned}\tilde{E}_M &= E - E_M \\ &= \sum_{(k,l)\notin \mathbf{S}_M} \bar{B}(k,l)\end{aligned} \quad (12)$$

Since E_M is of the M most significant energies in $\bar{\mathbf{B}}$, thus $\tilde{E}_M$ is of minimum average energy loss. In other words, the reconstructed image $\hat{\mathbf{O}}_M$ is optimal in the sense of minimum average energy loss.

4. Simulation Results

In this section, the proposed NAOT coding system is verified and compared with the JPEG. In Sect. 4.1, the optimality, the optimal feature selection, and the effect of parameter M in the NAOT coding system are justified and investigated. In Sect. 4.2, the effectiveness of coefficient selection in the NAOT coding system is demonstrated by comparing with the zigzag scan used in the JPEG where discussions on comparison results are given as well. In Sect. 4.3, the coding performances of the proposed system and the JPEG are compared by PSNR, bit rate, and subjective image quality. In the following simulation, four test images of size 512×512 are used. They are Lena, Baboon, Jet, and House, which are shown in Figs. 1, 2, 3, and 4, respectively. Moreover, all images are divided into 8×8 image blocks for all simulations.

4.1 Justification of the NAOT Coding System

This subsection consists of three parts. First, the optimality in the NAOT coding system is verified. Then the optimal feature selection, which comes along with optimal set $\mathbf{S}_M$, is demonstrated. Finally, the effect of parameter M on the NAOT coding system is investigated. In order to distinguish results of optimality test and the optimal feature selection, Step 8 and Step 9 described in Sect. 2.1 are skipped on purpose. The NAOT coding system without quantization and coding is depicted in Fig. 5. Based on Fig. 5, the optimality and the optimal feature selection in the NAOT coding system are justified. For the convenience of presentation, the 8×8 two-dimensional index is converted to one-dimensional index as given in Fig. 6.

4.1.1 Optimality Test

According to Fig. 5, all images are separately put into

Fig. 1 Original Lena

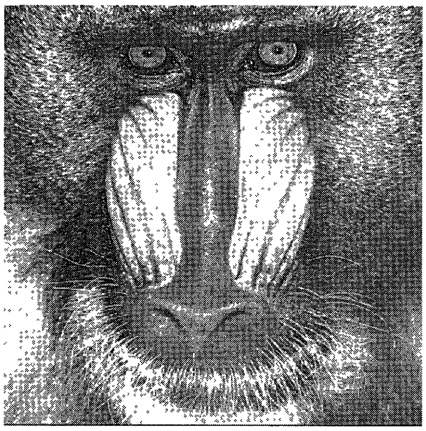


Fig. 2 Original Baboon.

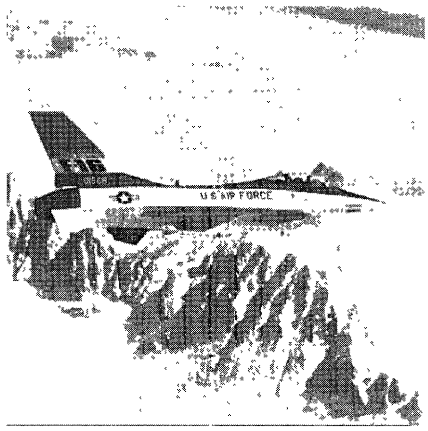


Fig. 3 Original Jet.

Fig. 4 Original House.

the NAOT coding system without quantization and coding. In the case of $M = 16$, the PSNR for Lena, Baboon, Jet, and House, are 36.0367 dB, 24.6300 dB, 34.7504 dB, and 30.7984 dB, respectively. The optimal sets $\mathbf{S}_{16}^L$, $\mathbf{S}_{16}^B$, $\mathbf{S}_{16}^J$, and $\mathbf{S}_{16}^H$, for images Lena, Baboon, Jet, and House are respectively recorded in Table 1, where the superscripts L, B, J, and H in $\mathbf{S}_{16}$ are for images Lena, Baboon, Jet, and House, respectively.

Note that the average energy loss is inverse proportional to PSNR in a reconstructed image, i.e., low average energy loss means high PSNR and vice versa. Thus the reconstructed image with optimal set $\mathbf{S}_M$ is of highest PSNR since it is of minimum average energy loss. In other words, the PSNR for a reconstructed image with set other than $\mathbf{S}_M$ is less than that with $\mathbf{S}_M$. If this is the case, the optimality in the NAOT coding system is verified. With the idea described, the optimality test on the NAOT coding system is performed. For $M = 16$, the results of optimality test for images Lena, Baboon, Jet, and House, are given in Table 2 where ξ_o denotes an element in $\mathbf{S}_{16}$ and ξ_s denotes a substitute element not in $\mathbf{S}_{16}$. For each image, eight experiments are performed where only one element is replaced in each case. Table 2 shows that all PSNR are less than that with optimal set $\mathbf{S}_{16}$. For example, all PSNR in Table 2 are less than 36.0367 dB for image Lena. This is also true for other images. Therefore, the opitmality in the NAOT coding system is justified.

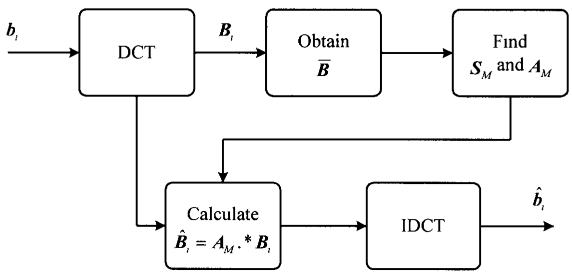


Fig. 5 The NAOT coding system without quantization and coding.

$$\begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) & (1,7) & (1,8) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & (2,7) & (2,8) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & (3,7) & (3,8) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & (4,7) & (4,8) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & (5,7) & (5,8) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & (6,7) & (6,8) \\ (7,1) & (7,2) & (7,3) & (7,4) & (7,5) & (7,6) & (7,7) & (7,8) \\ (8,1) & (8,2) & (8,3) & (8,4) & (8,5) & (8,6) & (8,7) & (8,8) \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 25 & 26 & 27 & 28 & 29 & 30 & 31 & 32 \\ 33 & 34 & 35 & 36 & 37 & 38 & 39 & 40 \\ 41 & 42 & 43 & 44 & 45 & 46 & 47 & 48 \\ 49 & 50 & 51 & 52 & 53 & 54 & 55 & 56 \\ 57 & 58 & 59 & 60 & 61 & 62 & 63 & 64 \end{bmatrix}$$

Fig. 6 2D-to-1D index conversion.

4.1.2 Optimal Feature Selection

The optimal feature selection, which comes along with $\mathbf{S}_M$, is demonstrated here. Note that the order of elements in Table 1 is in the descending order by energy and that energy loss is inverse proportional to PSNR. Consequently, the order in Table 1 is also the order of significance for the selected elements, when the PSNR is concerned. For example, in $\mathbf{S}_{16}^L$ element 1 affects PSNR more than element 2, which then has more effect on PSNR than element 9. To see the order of significance in the elements of $\mathbf{S}_M$, an element in Table 1 is taken out at a time in the simulation. That is, only $M - 1$ elements are used to obtain a reconstructed image. For $M = 16$, the simulation results for images Lena, Baboon, Jet, and House, are given in Table 3 where ξ_d denotes the element deleted in $\mathbf{S}_M$ and the superscripts L, B, J, and H are for images Lena, Baboon, Jet, and House, respectively. Table 3 indicates that the element with higher ranking index affects PSNR more than that with lower ranking index for all cases. For example, the PSNR loss is 19.8604 dB when element 1 (ranking index = 1) is taken out of $\mathbf{S}_{16}^L$ while the PSNR loss is 8.007 dB when element 2 (ranking index = 2) is deleted. Since deleting element 1 in $\mathbf{S}_{16}^L$ results in higher PSNR loss than deleting element 2, it is clear that element 1 has

Table 1 Elements of $\mathbf{S}_{16}^L$, $\mathbf{S}_{16}^B$, $\mathbf{S}_{16}^J$, $\mathbf{S}_{16}^H$ in the descending order by average energy.

Ranking index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\mathbf{S}_{16}^L$	1	2	9	3	10	11	4	17	18	19	12	5	20	25	26	27
$\mathbf{S}_{16}^B$	1	9	2	17	10	25	18	33	3	26	11	19	41	34	27	49
$\mathbf{S}_{16}^J$	1	9	2	17	3	10	25	11	33	18	4	19	41	5	12	26
$\mathbf{S}_{16}^H$	1	9	2	17	10	25	18	3	26	33	11	4	19	34	41	27

Table 2 Results for optimality test.

Lena			Baboon			Jet			House		
ξ_o	ξ_s	PSNR	ξ_o	ξ_s	PSNR	ξ_o	ξ_s	PSNR	ξ_o	ξ_s	PSNR
5	6	35.6889	3	4	24.3872	5	6	34.3156	4	5	30.4591
12	13	35.5221	11	12	24.4593	12	13	34.4558	3	12	29.5310
20	21	35.7837	19	20	24.4779	11	20	33.5650	11	13	30.0828
19	28	35.2706	41	42	24.5228	19	28	34.1790	19	20	30.3383
27	36	35.7970	34	35	24.5085	18	27	33.7037	27	28	30.5151
26	34	35.7696	27	28	24.5354	26	34	34.4251	26	35	30.0327
25	33	35.7440	33	43	24.2226	33	42	33.5478	34	43	30.4451
18	35	34.8860	49	50	24.5407	41	49	34.3469	41	42	30.5778

Table 3 Test results for optimal feature selection

ξ_d^L	PSNR	ξ_d^B	PSNR	ξ_d^J	PSNR	ξ_d^H	PSNR
1	16.1763	1	16 9343	1	12 1365	1	13 2508
2	28.0360	9	23.3245	9	27.2637	9	24.0679
9	31.1388	2	23.4403	2	27.8965	2	26.6660
3	32.7623	17	23.7780	17	30 7199	17	27.3803
10	33.0560	10	23.9618	3	31.7973	10	28.2683
11	34 4433	25	23.9764	10	31 8167	25	29 0564
4	34.7677	18	24.1289	25	32 5529	18	29.2314
17	34.7789	33	24.1520	11	33.2739	3	29.2684
18	34.7868	3	24.1574	33	33.4588	26	29.8031
19	35.0482	26	24.2328	18	33.4592	33	29.8824
12	35 1893	11	24.2443	4	33 4611	11	29.9339
5	35.3858	19	24.2910	19	34 0116	4	30.1228
20	35.5185	41	24.2948	41	34.0343	19	30.1288
25	35.6045	34	24.2998	5	34.0432	34	30.2990
26	35.6364	27	24.3656	12	34.1529	41	30.3500
27	35.6711	49	24.3887	26	34 1732	27	30 3711

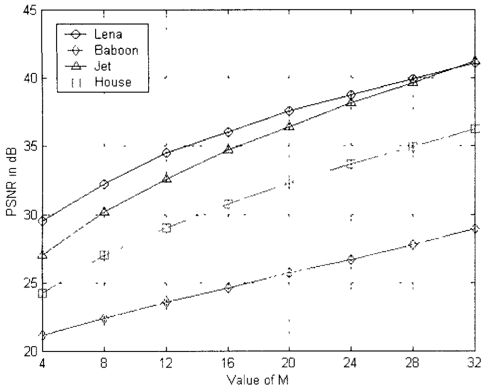


Fig. 7 Effect of parameter M without quantization and coding.

more significance than element 2. Note that the PSNR loss in Table 3 for each test image is in a descending order. Therefore, the optimal feature selection in $\mathbf{S}_M$ is verified.

4.1.3 Effect of Parameter M

The NAOT coding system shown in Fig. 5 has one parameter M which denotes the number of selected transform coefficients. The effect of parameter M on the coding performance without quantization and coding is investigated. The PSNR with various M for test images are found. The simulation results are depicted in Fig. 7. The differences of PSNR between $M = 4$ and $M = 32$ are 11.5964 dB, 7.7680 dB, 14.1708 dB, and 12.0551 dB for images Lena, Baboon, Jet, and House, respectively. When parameter M increases by 4, the PSNR are increased, on average, by 1.6566 dB, 1.1097 dB, 2.0244 dB, and 1.7222 dB for images Lena, Baboon, Jet, and House, respectively.

4.2 Effectiveness Comparison for ZZS and SBS

In this subsection, the effectiveness of coefficient selection in the NAOT coding system is demonstrated

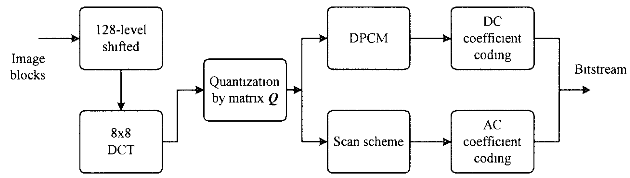


Fig. 8 The coding system for effectiveness comparison between ZZS and SBS

by comparing with the zigzag scan (ZZS) used in the JPEG. In the following discussion, the $\mathbf{S}_M$-based scan used in the NAOT coding system is denoted as SBS. In Sect. 4.2.1, a modified JPEG is introduced for fair comparison between SBS and ZZS. Next, the effectiveness of SBS is demonstrated by comparing with the modified JPEG in terms of PSNR. Finally, bit rates obtained from SBS and ZZS schemes are compared where discussions are given as well.

4.2.1 A Modified JPEG

Basically the JPEG consists of two stages in the coding process: (i) to quantize transform coefficients and (ii) to code quantized transform coefficients in the order of ZZS. With the 1-D index in Fig. 6, the ZZS order is 1, 2, 9, 17, 10, 3, 4, 11, 18, 25, and so forth. In the JPEG, there is no coefficient selection scheme applied and therefore different numbers of transform coefficients are coded for different image blocks in general. Therefore, it is not appropriate to compare ZZS and SBS directly since only M coefficients are selected in SBS. To make the effectiveness comparison for ZZS and SBS possible, in the simulation we modify the JPEG implementation in the following way. For each image block, only M transform coefficients are selected, quantized, and coded in the order of ZZS. This modification ensures that the effectiveness comparison for ZZS and SBS is made under same ground.

The coding system used to compare the effectiveness of SBS with ZZS is shown in Fig. 8. Note that

Table 4 Comparison results on PSNR for ZZS and SBS.

Value of M	Lena		Baboon		Jet		House	
	ZZS	SBS	ZZS	SBS	ZZS	SBS	ZZS	SBS
4	28.0255	28.7821	20.9326	20.9326	26.0713	26.0713	24.2117	24.2117
8	31.5424	31.5707	21.8509	22.1831	29.0938	29.2375	26.0780	26.9717
12	32.6627	33.7038	23.2180	23.3032	31.3858	31.5177	28.8781	28.8781
16	34.4988	34.9689	23.8157	24.2938	33.1688	33.3279	29.9644	30.4227
20	35.6150	35.7904	24.6818	25.2366	34.3248	34.5512	31.1134	31.4837
24	35.6973	36.0854	25.8493	26.0046	35.2322	35.3554	32.2362	32.2362
28	36.2173	36.3278	26.3234	26.6117	35.7608	35.7608	32.6015	32.6455
32	36.3556	36.3603	26.5090	27.2102	35.8307	35.8640	32.7031	32.7763

Fig. 9 Reconstructed Lena ($M = 24$).

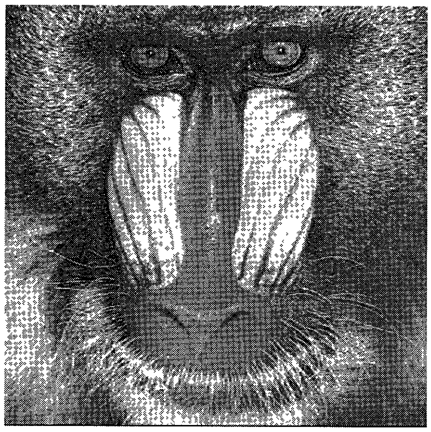


Fig. 10 Reconstructed Baboon ($M = 24$).

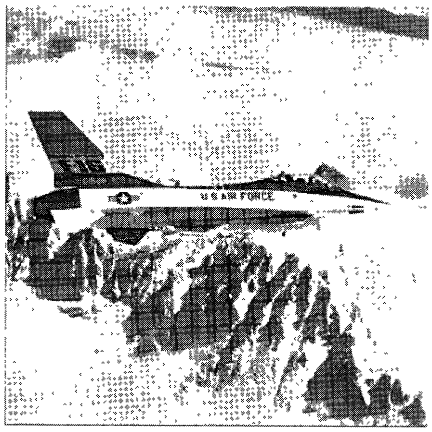


Fig. 11 Reconstructed Jet ($M = 24$).

the coding system in Fig. 8 is the JPEG when the scan scheme is ZZS and $M = 64$. In Fig. 8, the DCPM stands for differential pulse code modulation [8]. Here, the default settings in the JPEG [8] are used for quantization matrix **Q**, DC coefficient coding, and AC coefficient coding. In the simulation, the quantization matrix **Q** is given as

$$\mathbf{Q} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \quad (13)$$

The category table and Huffman code table for DC coefficient coding are found in Tables 8.3 and 8.4 in [8]. For AC coefficient coding, the category table and the variable length coding (VLC) table are obtained in Tables 8.5 and C.1 in Appendix C in [8]. As for the block of scan scheme in Fig. 8, it can be either ZZS or SBS.

4.2.2 Comparison Results on PSNR

Based on the coding system given in Fig. 8, the comparison results on PSNR, with various M, are given in Table 4. For the case of $M = 24$, the reconstructed images of Lena, Baboon, Jet, and House obtained from the NAOT coding system are shown in Figs. 9 to 12, respectively. Consider the case of $M = 4$ where only 4 coefficients are selected by ZZS or SBS in each image block. For ZZS, the selected coefficients in transform image blocks for all images are elements 1, 2, 9, and 17 as shown in Fig. 6. However, for SBS elements 1, 2, 9, and 3 are selected in image Lena while elements 1, 9, 2, and 17 are selected for other images as given in Table 1. Note that the elements selected by ZZS and SBS for images Baboon, Jet, and House, are same except in different order. Therefore, the PSNR are expected to be identical. This can be verified from Table 4. Since the order to code the selected coefficients is different from the JPEG, thus the number of bytes used in the coded image may be varied in general. This explains why different bit rate is obtained in Table 5 for the case of $M = 4$ and other similar cases. Nevertheless, it is expected to have better PSNR with SBS than that with ZZS since SBS is optimal in the sense of minimum average energy loss. As expected, Table 4 shows that the PSNR obtained from SBS is no less than that from ZZS for all cases. The maximum improvement on PSNR for images Lena, Baboon, Jet, and House, are 1.0373 dB ($M = 12$), 0.7012 dB ($M = 32$), 0.2264 dB ($M = 20$), and 0.8937 dB ($M = 8$), respectively. To sum up, simulation results indicate that SBS is more effective than ZZS in terms of PSNR.

4.2.3 Comparison Results on BR

In this subsection, the comparison of bit rate for SBS and ZZS is made. The comparison results on bit rate, with various M, are given in Table 5. It indicates that for image Lena the bit rate for SBS is less than that for ZZS for all M except $M = 28$. Besides, the bit rate for SBS is slightly higher than that for ZZS in images Baboon, Jet, and House except the case $M = 8$ for image Baboon. The higher bit rates shown in Table 5 does not necessarily imply that the NAOT coding system trades bit rate with better PSNR shown in Table 4. One possible reason for higher bit rates is that the ZZS is generally advantageous to the VLC since large amount of coefficients, after quantization, reduce to zero. This benefits the use of VLC on which AC coefficient coding in Fig. 8 is based. However, the VLC is not suitable for SBS in general. Since coefficients are coded in the descending order by average energy, it generally reduces the possibility to have a sequence of zeros in the coding process and therefore the application of VLC in SBS is not as good as in ZZS. In others words, it may cause a higher bit rate in most of cases. Consequently, simulation results in Table 5 suggest that an appropriate coefficient coding approach should be sought such that the bit rate can be reduced in the NAOT coding system.

Fig. 12 Reconstructed House ($M = 24$).

4.3 Comparison between the NAOT Coding System and the JPEG

In Sect. 4.2, the effectiveness of SBS and ZZS is compared in terms of PSNR and bit rate where a modified JPEG is applied. In this subsection, the performance of the NAOT coding system and the JPEG is compared by PSNR, bit rate, and subjective image quality.

4.3.1 Comparison Results on PSNR and BR

By setting the scan scheme as ZZS and $M = 64$ in Fig. 8, the PSNR and bit rate of the JPEG for test images are obtained. The PSNR are 36.4003 dB, 28.1197 dB, 35.8956 dB, and 32.9965 dB, for images Lena, Baboon, Jet, and House, respectively. The bit rate for images Lena, Baboon, Jet, and House, are 1.0641, 1.6479, 1.1296, and 1.4021, respectively. The comparison results on PSNR for the NAOT coding system and the JPEG, up to ΔPSNR less than 0.1 dB, are given in Table 6 where $\Delta PSNR = PSNR_{JPEG} - PSNR_{NAOT,M}$. Notations $PSNR_{JPEG}$ and $PSNR_{NAOT,M}$ stand for the PSNR for the JPEG and the NAOT coding system with parameter M, respectively. In the case of $M = 32$ for image Lena, $\Delta PSNR = 36.4003 - 36.3603 = 0.0400$ dB which means the NAOT coding system is worse than the JPEG by 0.0400 dB in PSNR. The comparison results on bit rate, corresponding up to ΔPSNR less than 0.1 dB, are shown in Table 7 for the two systems where $\Delta BR = BR_{JPEG} - BR_{NAOT,M}$ and BR_{JPEG}, $BR_{NAOT,M}$ denote the bit rate resulted from the JPEG and the NAOT coding system with param-

Table 5 Comparison results on BR for ZZS and SBS.

Value of M	Lena		Baboon		Jet		House	
	ZZS	SBS	ZZS	SBS	ZZS	SBS	ZZS	SBS
4	0.2611	0 2602	0.3053	0.3055	0.2606	0 2647	0.3147	0.3225
8	0 4032	0.3881	0.5376	0.5372	0 4048	0.4184	0 5184	0 5490
12	0 4719	0.4690	0.7464	0.7615	0 5112	0.5225	0 7039	0 7158
16	0.5316	0.5218	0 8985	0.9098	0 5799	0.5962	0.8174	0.8359
20	0 6085	0.6064	1 0303	1 0627	0.6721	0 6913	0.9079	0.9287
24	0 6614	0.6530	1.1453	1.1625	0.7276	0 7570	0.9845	0.9914
28	0.6953	0 6957	1.2263	1 2381	0.7693	0.7912	1.0331	1 0563
32	0 7205	0 7202	1.2625	1.3309	0.7860	0.8134	1.0574	1.0813

Table 6 Difference of PSNR between the JPEG and the NAOT coding system

Value	Lena		Baboon		Jet		House	
of M	$\mathrm{PSNR}_{\mathrm{NAOT,M}}$	ΔPSNR	$\mathrm{PSNR}_{\mathrm{NAOT,M}}$	ΔPSNR	$\mathrm{PSNR}_{\mathrm{NAOT,M}}$	ΔPSNR	$\mathrm{PSNR}_{\mathrm{NAOT,M}}$	ΔPSNR
4	28.7821	7.6182	20.9326	7.1871	26.0713	9.8243	24.2117	8.7848
8	31.5707	4.8296	22.1831	5.9366	29.2375	6.6581	26.9717	6.0248
12	33.7038	2.6965	23.3032	4.8165	31.5177	4.3779	28.8781	4.1184
16	34.9689	1.4314	24.2938	3.8259	33.3279	2.5677	30.4227	2.5738
20	35.7904	0.6099	25.2366	2.8831	34.5512	1.3444	31.4837	1.5128
24	36.0854	0.3149	26.0046	2.1151	35.3554	0.5402	32.2362	0.7603
28	36.3278	0.0725	26.6117	1.5080	35.7608	0.1348	32.6455	0.3510
32	36.3603	0.0400	27.2102	0.9095	35.8640	0.0316	32.7763	0.2202
36	–	–	27.5262	0.5941	–	–	32.8801	0.1164
40	–	–	27.8286	0.2911	–	–	32.9317	0.0648
44	–	–	27.9093	0.2104	–	–	–	–
48	–	–	28.0046	0.1151	–	–	–	–
52	–	–	28.0450	0.0747	–	–	–	–

Table 7 Difference of BR between the JPEG and the NAOT coding system.

Value	Lena		Baboon		Jet		House	
of M	$\mathrm{BR}_{\mathrm{NAOT,M}}$	ΔBR	$\mathrm{BR}_{\mathrm{NAOT,M}}$	ΔBR	$\mathrm{BR}_{\mathrm{NAOT,M}}$	ΔBR	$\mathrm{BR}_{\mathrm{NAOT,M}}$	ΔBR
4	0.2602	0.8039	0.3055	1.3424	0.2647	0.8649	0.3225	1.0796
8	0.3881	0.6760	0.5372	1.1107	0.4184	0.7112	0.5490	0.8531
12	0.4690	0.5951	0.7615	0.8864	0.5225	0.6071	0.7158	0.6863
16	0.5218	0.5423	0.9098	0.7381	0.5962	0.5334	0.8359	0.5662
20	0.6064	0.4577	1.0627	0.5852	0.6913	0.4383	0.9287	0.4734
24	0.6530	0.4111	1.1625	0.4854	0.7570	0.3726	0.9914	0.4107
28	0.6957	0.3684	1.2381	0.4098	0.7912	0.3384	1.0563	0.3458
32	0.7202	0.3439	1.3309	0.3170	0.8134	0.3162	1.0813	0.3208
36	–	–	1.3707	0.2772	–	–	1.1229	0.2792
40	–	–	1.4237	0.2242	–	–	1.1694	0.2327
44	–	–	1.4591	0.1888	–	–	–	–
48	–	–	1.5142	0.1337	–	–	–	–
52	–	–	1.5580	0.0889	–	–	–	–

eter M, respectively. In the case of $M = 32$ for image Lena, $\Delta\mathrm{BR} = 1.0641 - 0.7202 = 0.3439$. This means the JPEG takes 0.3439 bit rate more than the NAOT coding system. By Tables 6 and 7, it suggests that the JPEG pays ΔBR bit rate to have ΔPSNR dB improvement on PSNR when compared with the NAOT coding system. From the other viewpoint, it can be said that the NAOT coding system trades ΔPSNR dB in PSNR for ΔBR bit rate reduction. Take an example. For image Lena, the JPEG uses 0.3439 bit rate to obtain 0.0400 dB improvement on PSNR when compared with the NAOT coding system with $M = 32$. Also, it can be said that the NAOT coding system pays 0.0400 dB in PSNR to exchange 0.3439 bit rate reduction. Similarly, the bit rate reduction for images Baboon, Jet, and House, are 0.0889 ($M = 52$), 0.3162 ($M = 32$), and 0.2327 ($M = 40$), at the cost of 0.0747 dB, 0.0316 dB, and 0.0648 dB, degradation in PSNR, respectively. If higher PSNR degradation is acceptable, i.e., smaller M, the bit rate can be reduced further in the NAOT coding system. The simulation results indicate that it is possible for the proposed NAOT coding system to trades very little PSNR for significant bit rate reduction when compared with the JPEG.

4.3.2 Comparison of Subjective Image Quality

Previously, the coding performance of the NAOT coding system and the JPEG is evaluated by PSNR, an assessment of objective image quality. However, it is known that better PSNR does not mean better subjective image quality. Therefore, the comparison of subjective image quality between the two systems is under consideration here. From a perception viewpoint, transform coefficients have different significance or visual sensitivities in general. The low-frequency coefficients usually are far more sensitive than the high-frequency coefficients. This property of human visual system (HVS) can be utilized in the coding process through weighting transform coefficients. An HVS weighting matrix for 8×8 DCT coefficients can be found in [9]. In the JPEG, the HVS is reflected in quantization matrix $\mathbf{Q}$ as given in (13) where the low-frequency coefficients are finely quantized and the high-frequency coefficients coarsely quantized. Note that all non-zero quantized coefficients contribute to the subjective image quality when the HVS is applied in the coding process. Since some non-zero quantized coefficients are discarded in the NAOT coding system, thus the subjective image quality of the JPEG is better than the NAOT coding system, except $M = 64$. However, with appropriate M the subjective image quality is generally not degraded too much in the NAOT coding system since the discarded coefficients are of minimum average energy loss. For example, the visual difference between the reconstructed Lena shown in Fig. 9 and the one obtained from the JPEG should be very little since the difference of energy loss between the JPEG and the NAOT coding system is quite small. Consequently, the subjective image quality in the JPEG is slightly better, except $M = 64$, than the NAOT coding system with appropriate M in general.

5. Conclusions and Further Research

In this paper, a non-adaptive optimal transform (NAOT) coding system is proposed. Based on the following observations, the NAOT coding system is developed. First, under an orthogonal transformation the total energy of an image is invariant between spatial domain and transform domain. Second, the MSE of a given reconstructed image is proportional to the total energy of transform coefficients discarded in the coding process. By the energy-invariant property, the NAOT coding system is proved optimal in the sense of minimum average energy loss. Basically, the proposed NAOT coding system can be divided into three stages. First, obtain the average energy image block from transform image blocks. Second, record the indices of average energy image block in the descending order by energy. When the number of selected coefficients, M, is specified, the first M elements in reordered indices form a set denoted as $\mathbf{S}_M$. By set $\mathbf{S}_M$, a fixed mask $\mathbf{A}_M$ is found and used to select M significant coefficients in transform image blocks. Third, quantize and code the M selected coefficients by the order of $\mathbf{S}_M$. The NAOT coding system requires several bytes to indicate the optimal coefficient selection where the significance order of coefficients is obtained as well. The NAOT coding system is then verified through simulations where images Lena, Baboon, Jet, and House, are used as examples.

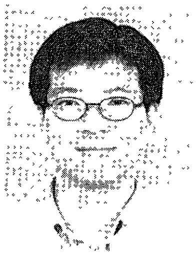
The simulation is divided into two parts. In the first part, the optimality, the optimal feature selection, and the effect of parameter M in the NAOT coding system are justified and investigated, respectively, where no quantization and coding is performed. Simulation results agree with the theoretical results in the NAOT coding system. In the second part, the effectiveness of $\mathbf{S}_M$-based scan (SBS) in the NAOT coding system is compared with the zigzag scan (ZZS) used in the JPEG. To be fair, the JPEG is modified to code only M transform coefficients in the order of ZZS. Simulation results show that SBS is more effective than ZZS in terms of PSNR. The improvements on PSNR are as high as 1.0373 dB, 0.7012 dB, 0.2264 dB, and 0.8937 dB for images Lena, Baboon, Jet, and House, respectively. However, the bit rate for images Baboon, Jet, and House, obtained from SBS is a little bit higher than that from ZZS for most of cases. One possible reason for higher bit rates may be the coefficient coding approach in the JPEG is not appropriate for the NAOT coding system. Consequently, an appropriate coefficient coding approach for SBS will be sought in further research such that lower bit rate can be obtained in the NAOT coding system. Besides, the performance of the NAOT coding system and the JPEG is also compared in the simulation, by PSNR, bit rate, and subjective image quality. The simulation results suggest that the proposed NAOT coding system is able to trade very little PSNR for significant bit rate reduction when compared with the JPEG. And the subjective image quality in the JPEG is generally slightly better, except $M = 64$, than the NAOT coding system with appropriate M.

Acknowledgement

This work was supported by the National Science Council of the Republic of China under grants NSC 91-2213-E-270-002 and NSC 92-2213-E-270-006.

References

[1] A. Palau and G. Mırchandani, "Image coding with discrete cosine transforms using efficient energy-based adaptive zonal filterıng," IEEE Int. Conf Acoust , Speech Signal Process., pp.337–340, 1994.
[2] S. do Nascımento Neto and F.A de Oliverira Nascimento, "Hybrid adaptive ımage coding transform coding and vectorial classification approach for bit allocation," 38th Midwest Symposium on Circuits and Systems, pp.846–849, Aug. 1995.
[3] M. Crouse and K Ramchandran, "Joint thresholding and quantizer selection for transform image coding: Entropy-constrained analysıs and applicatıons to baseline JPEG," IEEE Trans. Image Process., vol 6, no.2, pp.285–297, Feb. 1997.
[4] M. Bi, S.H. Ong, and Y.H. Ang, "Coefficıents grouping method for shape-adaptıve DCT," Electron. Lett , vol.32, no.3, pp.201–202, Feb. 1996.
[5] T D Tran and R. Safranek, "A locally adaptive perceptual masking threshold model for image coding," IEEE Int. Conf. Acoust., Speech Signal Process., pp.1882–1885, 1996.
[6] J.Y. Nam and K R Rao, "Image codıng usıng a classified DCT/VQ based on two-channel conjugate vector quantization," IEEE Trans. Circuıts Syst Video Technol., vol.1, no.4, pp 325–336, 1991.
[7] Y. Huang, H.M. Dreizen, and N.P. Galatsanos, "Prioritized DCT for compression and progressıve transmıssıon of images," IEEE Trans. Image Process., vol.1, no.4, pp.477–487, Oct. 1992.
[8] K R Rao and J.J. Hwang, Standards for Image, Vıdeo, and Audio Coding, Prentice Hall, 1996.
[9] Chitprasert and K.R Rao, "Human vısual weighted progressive image transmıssion," IEEE Trans. Commun., vol.38, no.7, pp.1040–1044, July 1990



Cheng-Hsiung Hsieh received the B.S. degree in Electronic Engineering from National Taiwan Institute of Technology, Taiwan, in 1989. In 1995, he earned the M.S. degree from the Department of Electrical Engineering of Tennessee Technological University, USA. He obtained his Ph.D. degree in Electrical Engineering from the University of Texas at Arlington, USA, in 1997. Currently, he is an associate professor at Department of Electronic Engineering in Chien Kuo Institute of Technology, Changhua, Taiwan, Republic of China. His research interests include grey systems, artificial neural networks, digital image processing, statistical signal processing, and digital communications.