VQ-BASED MASKS FOR TRANSFORM CODING

Cheng-Hsiung Hsieh

Department of Electronic Engineering
Chien Kuo Institute of Technology, Changhua, Taiwan, R.O.C.
E-mail:hsieh@cc.ckit.edu.tw

ABSTRACT

In this paper, an approach based on VQ to find masks for transform image coding is presented. In the approach, the key role of VQ is to find representative image blocks and use them to obtain masks for different classes of image blocks. Then an adaptive image coding system is proposed where VQ-based masks are employed to select significant transform coefficients. Two examples are provided to investigate the coding performance of proposed system. In the simulation, a self-organizing mapping (SOM) neural network, which is considered as neural VQ, is used because of its good topology-preserving property. Simulation results indicate that the proposed coding system is feasible and promising.

1. INTRODUCTION

In order to utilize the bandwidth of transmission channel efficiently and effectively, image compression schemes are sought. One of popular image compression approaches is transform coding [1]. In transform coding, the mean-square error (MSE) of reconstructed image is proportional to the total energy of transform coefficients discarded in the coding process. Therefore, effectively selecting significant transform coefficients implies better coding performance. In other words, an appropriate masking method is the key to better performance.

Up to present, several adaptive approaches to select significant transform coefficients have been proposed. Chen and Smith [2] first divided image blocks into classes based on the energy of ac-coefficients and then adaptively encoded coefficients in different classes. Given a portion of total energy in the transform domain, Palau and Mirchandani [3] used several geometric shapes to search for the geometric zone, which contained the least number of coefficients having the specified portion of energy. Pan [4] employed VQ in transform domain to achieve better coding performance while Amerijckx and at el. [5] used similar technique in coding process in which neural VQ was applied. Using equipotentials of energy in transform domain, Neto and Nascimento [6] proposed a modified zonal coding approach where the selection of transform coefficients was based a given signal-to-noise ratio. Crouse and Ramchandran [7] applied the optimization technique in finding coefficient thresholds and optimal Q-matrix used in JPEG. Ong and Ang [8] utilized the statistical property, cross-correlation, to choose transform coefficients. Tran and Safranek [9] included the perceptual masking threshold model into the framework of image coding. The mask was used in the selection of transform coefficients. Dony and Haykin [10] proposed an approach to optimally adaptive transform coding where modular architecture was applied and stands for a group of transformations. Among several coefficient vectors obtained from different transformations, the coefficient vector with maximum norm was chosen.

In this paper, a VQ-based masking approach for transform coding is proposed. In Section 2, the concept of VQ [11], on which the masking approach is based, is described. Then the transform coding system based on masks from VQ is proposed in Section 3. Simulation results to justify the proposed coding system are given in Section 4 where the results are discussed as well. Finally, conclusions are made in Section 5.

2. VQ AND MASK

The idea to find masks through VQ scheme is motivated by following two observations. First, the VQ scheme classifies image blocks based on a similarity measure. This implies image blocks within the same class, in general, have similar distribution of transform coefficients. Second, image blocks having same class index are represented by their corresponding representative image block found in VQ approach. This implies the distribution of transform coefficients in representative image block is a good estimate of that for image blocks within the class. Based on these observations, the masks required in adaptive transform coding are found. Here the key role of representative image blocks is to find the masks by which significant transform coefficients are located.

To find the mask for a given class, the representative image block associated with that class is first transformed. Next, transform coefficients of the representative image block are rearranged in the descending order according to their energies and a map, before and after reordering, is constructed. When the total number of coefficients to be kept, M, is specified, the addresses of the M most significant coefficients are recorded as the mask for the given class.

To make the idea clear, an example is given in the following. Assume the 4×4 matrix given in the left of Figure 1 is the representative image block for some class.

Then it is transformed by the discrete cosine transform (DCT) [1]. The matrix of transform coefficients is shown in the right of Figure 1. If $M=4$, obviously elements (1,1), (4,4), (4,2), and (3,1) should be selected according to energy criteria. The mask for the given class is formed by the addresses of selected elements in the 4×4 representative image block, i.e., addresses (1,1), (4,4), (4,2), and (3,1).

3. THE PROPOSED CODING SYSTEM

Basically, the proposed coding system consists of four steps as follows. First, utilize VQ to find *K* representative image blocks and then to obtain *K* masks for *K* classes as described in the previous section. Second, classify image blocks into classes. Third, perform DCT on image blocks. Fourth, for each image block one mask, according to the class index, is chosen and applies to keep significant transform coefficients and to discard others. In the coding system, a neural VQ, self-organizing mapping (SOM) network [12], is used to find representative image blocks and to perform classification on image blocks, since SOM network is of good topology-preserving property. As for transformation, the DCT is employed since it is widely used in the image coding community and the JPEG standard [13].

The block diagram for the proposed coding system is depicted in Figure 2. Blocks (1), (2), (3), and (4) are related to the first step described in last paragraph. Blocks (1) and (5) are associated with the second step. Blocks (6) and (7) are for the third step and the fourth step, respectively. The rest of blocks are for decoding and performance evaluation. In the coding system, there are two important parameters *K* and *M*. Parameter *K* is the total number of classes used and Parameter *M* is the total number of transform coefficients selected in the coding system. How these two parameters affect the coding performance is discussed in the following section.

4. SIMULATION AND DISCUSSIONS

In this section, two examples are provided to investigate the performance of the proposed coding system. In the simulation, there is no bit allocation scheme applied. In the first example, the 512×512 image Lena shown in Figure 3 is used. According to the proposed coding system given in Figure 2, the image Lena is first divided into non-overlapping image blocks of size 8×8. Then SOM VQ is employed to find 256 representative image blocks. That is, parameter *K* is equal to 256. Next, 256 masks are found as described in Section 2. Each mask is for one class of image blocks. In the case of $K=256$ and $M=16$, the reconstructed image is shown in Figure 4 whose PSNR (Peak Signal-to-Noise Ratio) is 35.39 dB. Next, to see how parameter *K* affects the coding performance, several values of *K* are used while *M* is fixed. The simulation results plotted in Figure 5 is under the condition $M=16$. Similarly, the effect of parameter *M* on the coding performance is given in Figure 6 where $K=256$.

To have more general understanding of the coding system, the 512×512 image Harbor shown in Figure 7 is used as the second example. The block size is 8×8. With $K=256$ and $M=16$, the reconstructed image is given in Figure 8 which is of PSNR 27.74 dB. As in the first example, the effects of parameters *K* and *M* on coding performance are also shown in Figure 5 and Figure 6, respectively.

From Figure 5, it indicates that parameter *K* does not affect the coding performance very much. When $M=16$, the difference in PSNR between cases $K=4$ and $K=256$ is 3.05 dB for image Lena and only 1.07 dB for image Harbor. When *K* is reduced to $K/2$, the PSNR is degraded by 0.7 dB maximum for image Lean and by 0.54 dB maximum for image Harbor. On the other hand, Figure 6 indicates that parameter *M* has much effect on the coding performance. Under the condition $K=256$, the difference in PSNR between cases $M=4$ and $M=32$ is 8.54 dB for image Lena and 6.39 dB for image Harbor. In both examples, the worst case in PSNR reduction happens from $M=8$ to $M=4$. The difference is 2.64 dB for image Lena and 1.55 dB for image Harbor.

5. CONCLUSIONS

In this paper, the VQ scheme is extended to find masks for image transform coding. An image coding system is proposed where VQ-based masks are applied. Two images, Lena and Harbor, are given to examine the proposed coding system. Simulation results show that the proposed approach is feasible and promising. Besides, two important parameters *K* and *M* in the coding system are investigated in the effect on the coding performance. It turns out parameter *M* affects more than parameter *K* on PSNR. The result gives useful information in the application of the proposed coding system.

ACKNOWLEDGEMENT

This work was funded by the National Science Council of the Republic of China under grant NSC 89-2213-E-270-013.

REFERENCES

[1] K. R. Rao and J. J. Hwang, *Techniques and Standards for Image, Video, and Audio Coding*, Pretice Hall, 1996.

[2] W. H. Chen and C. H. Smith, “Adaptive Coding of Monochrome and Color Images,” *IEEE Trans. on Communications*, Vol. 25, pp. 1285-1292, Nov. 1977.

[3] A. Palau and G. Mirchandani, “Image Coding with Discrete Cosine Transforms Using Efficient Energy-Based Adaptive Zonal Filtering,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 337-340, 1994.

[4] J. Pan, “A Classification Method for Adaptive Transform Image Coding,” *International Conference on Image Processing*, pp. 286-289, 1995.

[5] C. Amerijckx, M. Verleysen, and P. Thissen, and J.-D. Legat, “Image Compression by Self-Organized Kohonen Map,” *IEEE Trans. on Neural Networks*, Vol. 9, No. 3, May, 1998.

[6] Sebastião do Nascimento Neto and Fransisco Assis de Oliverira Nascimento, “Hybrid Adaptive Image Coding：Transform Coding and Vectorial Classification Approach for Bit Allocation,” *The 38th Midwest Symposium on Circuits and Systems*, pp. 846-849, Aug. 1995.

[7] M. Crouse and K. Ramchandran, “Joint Thresholding and Quantizer Selection for Transform Image Coding: Entropy-Constrained Analysis and Applications to Baseline JPEG,” *IEEE Trans. on Image Processing*, Vol. 6, No. 2, pp. 285-297, Feb. 1997.

[8] M. Bi, S. H. Ong and Y. H. Ang, “Coefficients Grouping Method for Shape-Adaptive DCT,” *Electronics Letters*, Vol. 32, No. 3, pp. 201-202, Feb. 1996.

[9] T. D. Tran and R. Safranek, “A Locally Adaptive Perceptual Masking Threshold Model for Image Coding,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1882-1885, 1996.

[10] R. D. Dony and S. Haykin, “Optimally Adaptive Transform Coding,” *IEEE Trans. on Image Processing*, Vol. 4, No. 10, pp. 1358-1370, Oct. 1995.

[11] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992.

[12] S. Haykin, *Neural Networks – A Comprehensive Foundation*, Macmillan, 1994.

[13] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*, New York, NY: Van Nostrand Reinhold, 1993.

$$\begin{bmatrix} 50 & 70 & 30 & 90 \\ 80 & 150 & 50 & 250 \\ 250 & 40 & 150 & 50 \\ 70 & 100 & 55 & 125 \end{bmatrix} \xrightarrow{DCT} \begin{bmatrix} 402.50 & -11.08 & 80.00 & -33.29 \\ -30.52 & -44.51 & 2.71 & -59.43 \\ -107.50 & -27.98 & -40.00 & -47.94 \\ -27.95 & 123.07 & -6.53 & 157.01 \end{bmatrix}$$

Figure 1. An example to find mask for a class

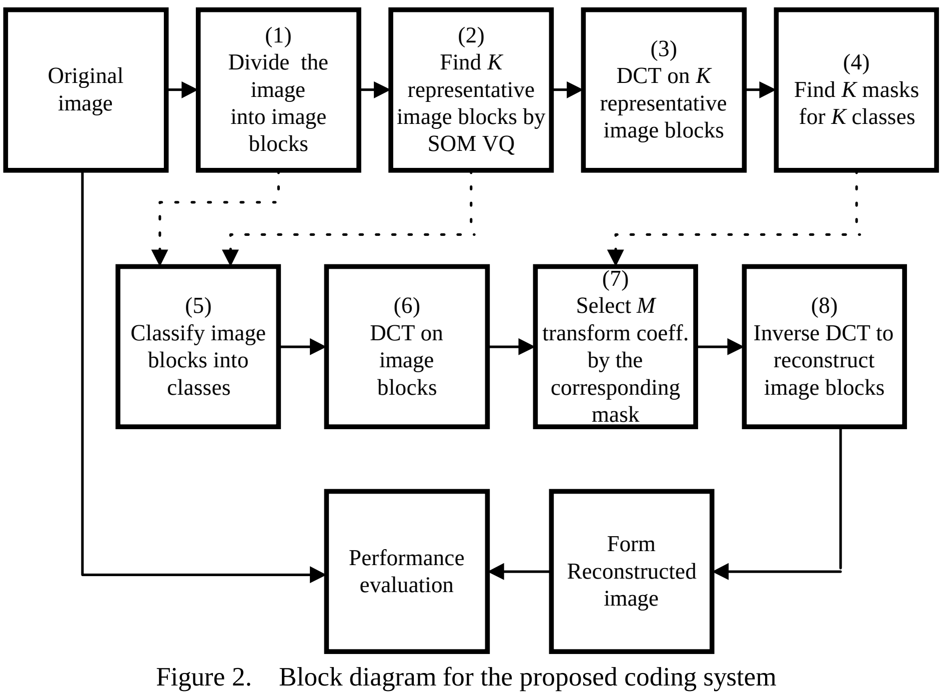

Figure 2. Block diagram for the proposed coding system

Figure 3. Original Lena

Figure 4. Reconstructed Lena

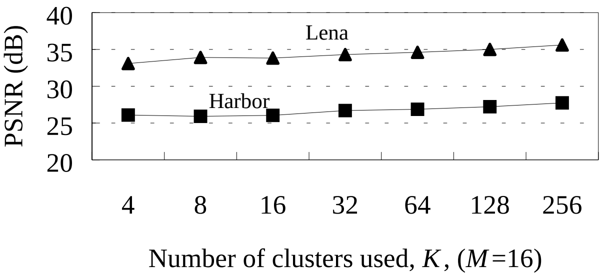

Figure 5. Effect of parameter K on coding performance

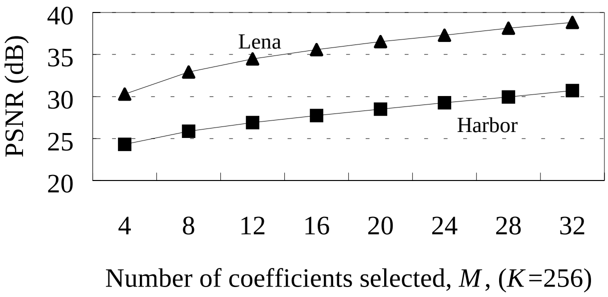

Figure 6. Effect of parameter M on coding performance

Figure 7. Original Harbor

Figure 8. Reconstructed Harbor