1998 International Computer Symposium
Workshop on Image Processing and Character Recognition
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

CLASSIFIED VECTOR QUANTIZATION USING DIRECTION CLASSIFICATION

Chou-Chen Wang, Chin-Hsing Chen, Kun-Yu Wang, and Yih-Daw Tsai

Department of Electrical Engineering,
National Cheng Kung University, Tainan, Taiwan, R. O. C.
E-mail: chench@eembox.ncku.edu.tw

ABSTRACT

In this paper, an effective classifier based on the direction of vectors for classified vector quantization is proposed. The new classifier uses dimensionality-independent distortion measure based on inner product to determine the similarity between vectors. The distortion measure is simple and enough to classify various edge types other than single and straight line types, which limit the size of image block to a rather small size. Simulation results show that the technique can achieve good perceptual quality and edge integrity at low bit rates with reduced coding complexity.

1. INTRODUCTION

Classified vector quantization (CVQ) has proved to be an efficient technique for lossy image compression at low bit rates [1-4]. CVQ preserves the perceptual features, such as edges, while maintaining a simple distortion measure and reducing the computational complexity of the traditional vector quantization (VQ) [5]. It is based on a composite source model, in which an input vector is classified and coded with codebook especially designed for the class to which it belongs. Transform domain classification methods require the use of features that are extracted from the transformed data, such as the discrete cosine transform (DCT) or Walsh Hadamard transform (WHT) [6]. Spatial domain classification techniques are complicate which involve preprocessing and the employment of threshold parameters.

In general, the classifier of CVQ is implemented in two steps: an edge enhancement step, followed by a complicated decision step which extracts the edge description from the enhanced version. The enhancement is to enhance the edge components in the blocks. The gradients of neighboring pixels are computed and compared with some threshold parameters to decide if the neighboring pixels have edge components. The threshold parameters are heuristically defined and largely dependent on the training data. The decision algorithm approximates an edge by a straight line. Only vertical, horizontal, and two diagonal orientations are recognized, and the classes are defined heuristically. Hence, the algorithm gives meaningful results for small blocks. As the block dimension increases, the edges become more complicated. The straight lines and the monotonous classes can not approximate the complicated edges very well, and the threshold parameters for decision are difficult to define. Different parameters are required for different training data. The classifier proposed by Kim and Lee [3-4] used four 2-D 4 × 4 DCT coefficients as the features for classification. The class patterns are generated from training data by means of iterative "learning algorithm". Then the distortions between the feature of each input transformed vector and those of all classes are computed to decide which class the input vector belongs to. When the block dimension is small, the edges are simple and monotonous. The energy in the transform domain is compacted into a few dominant frequencies. The coefficients of these frequencies can be treated as the feature of the block. Hence, this algorithm based on feature coefficients is very effective while the block vector is small or the edge pattern is simple. However, as the block dimension increases, the location and orientations of edges are much more complicated and the energy of the blocks would not be compacted into a few dominant frequencies. The coefficients of the unitary transform, like DCT, are nearly mutually uncorrelated and any coefficients can not be expected from others. Then the feature coefficients would not represent the active edges well. A similar method of classification in the WHT domain was reported [6] but it is neither not suitable for large block vectors.

In this paper, a new classifier scheme for CVQ based on the direction of image vectors, which suitable for large blocks is proposed to solve the above problems. The method can obtain better performance using smaller subcodebooks.

2. CLASSIFICATION USING THE DIRECTION CLASSIFIER

For CVQ, it is desirable to use a classification rule that is sensitive to the perceptual features of the input blocks. Therefore, if two vectors belong to the same class in CVQ, they must look "alike". In order to preserve the edges of images, the vectors which are alike are used to train one subcodebook. Usually the mean square error (*MSE*) is used as the distortion measure due to its simplicity and analytical tractability. In other words, the distance is used to decide if two vectors are alike. However, the *MSE* does not have any edge preserving property.

Intuitively, if two vectors have similar edges and belong to the same class, they must be located in similar direction and the angle between them is very small. This can be explained in the two-dimensional space shown in Fig. 1. The distance between two vectors B and C is smaller than the distance between B and A. But, the edge pattern represented by vector B is more like that by A since the angle between two vectors B and A is smaller than the angle between B and C. This is the basic idea of classification in this paper.

In this scheme, a direction-based classification method is proposed to design the classbook and classifies the vectors. The class to which a vector is assigned to is the one having the minimum angle with the vector. The distortion measure based on the direction is defined as

$$d(X_i, X_j) = \theta^2 \tag{1}$$

where X_i and X_j are any two training vectors and θ is the angle between X_i and X_j. Applying the Taylor's series of $\cos\theta$ about 0, we obtain

$$\cos\theta = 1 - \frac{1}{2}\theta^2 + \frac{1}{24}\theta^4 - \frac{1}{720}\theta^6 + \cdots \tag{2}$$

The first two terms in Eq. (2) are dominant, thus

$$\theta^2 \cong 2(1-\cos\theta). \tag{3}$$

Therefore, the distortion measure can be redefined as

$$d(X_i, X_j) = 1 - \cos\theta. \tag{4}$$

The value of the distortion is in the range [0,1] and is minimum 0 when the two vectors are in the same direction. As the angle increases, the distortion measure gets larger. The distortion measure given by Eq. (4) is easy to calculate in the spatial domain by

$$d(X_i, X_j) = 1 - x_i^T x_j = 1 - \sum_{l=0}^{k-1} x_{il} x_{jl} \tag{5}$$

where k is the dimension of the vector, x_i and x_j are the normalized vectors of X_i and X_j. The vectors are normalized before classification since it is convenient to compute the distortion. As the two vectors are more similar, the distortion is less and they more likely belong to the same class.

With M-classes, our classifier uses a set of M initial perceptually vectors $\{D_i : i = 1, \ldots, M\}$ of the same dimensionality as the input vectors. The proceed of the classbook training is as follows:

Step (1): Initialize the M unit classvectors

$$\{d_1^{(0)}, d_2^{(0)}, \ldots, d_M^{(0)}\}$$

Step (2): For each training input unit vector x

$$x \in d_i^n \quad if \quad (1 - x^T d_i^{(n)}) = \min_{j=1,\ldots,M} (1 - x^T d_j^{(n)}) \tag{6}$$

Step (3): Update the vector $d_i^{(n)}$ using the centroid of the training vectors for $x \in d_i^n$

$$d_i^{(n+1)} = \frac{1}{m} \sum_{x_j \in d_i^{(n)}}^{m} x_j \tag{7}$$

where n is the iteration index and m is the number of the vectors $x \in d_i^n$.

Step (4): Repeat for each training vector until the classvectors converge.

In order to obtain the classbook more efficiently, the shade blocks which contain no obvious edges are excluded from the classbook training. To classify each input block, we first decide whether it is a shade or active block according to the difference of the high-mean and low-mean of the block. The high mean and low mean represent the mean of pixels whose gray values are equal to or greater than the block mean and less than the block mean, respectively. If the difference is less than a pre-selected threshold shade value t, the block is regarded as a shade block. Otherwise, it is active block and the direction-based classification algorithm is applied. It calculates the input vector with each of the classvectors and chose the index of the classvector that has the minimum angle as the class index.

The direction space of vectors is partitioned into one direction subspace for shade vectors and M direction subspaces for active vectors. Every subspace is represented by a classvector from the classbook. In our method, the classification is based on the vector distribution in the space domain and the parameters used have nothing to do with the block dimension. This is an attractive property for classification. As the block dimension increases, our method remains effective by partitioning the vectors according to direction similarity. The magnified 8×8 classvectors of 32 classes are shown in Fig. 2. There are several classes containing complicated edges, e.g., two edges in the diagonal, horizontal and vertical direction. The above task is not easy for the classification method in [1, 4, 6].

3. EXPERIMENT RESULTS

The codebook training sets of the classes are created by running the direction classification algorithm on a set of four 512×512, 8 bits images of different nature. The test images "Lena" and "F-16" outside the training set are shown in Fig. 3. The quality of the encoded images is evaluated by the peak signal to noise ratio (*PSNR*) which is defined as

$$PSNR = 10 \cdot \log_{10}(\frac{255^2}{MSE}) \quad dB \tag{8}$$

where *MSE* is the mean square error between the reconstructed and the original image.

The number of $(M+1)$ classes was equal to 17 and 33. A shade threshold $t = 20$ was chosen. A codebook with N_i codewords is designed based on [1] for each class such that the total number of codewords in all the classes is equal to N. The performance for Lena and F-16 images are summarized in Table 1. In Table 2, we compare the coding results with those of other CVQ-based techniques that use different block dimension. It clearly reveals the success of the new method. When the dimension increasing, the performance of our method far outperform others. The decoded images with good visual quality are shown in Fig. 4. Edges are reproduced faithfully and their jaggedness is greatly reduced. Moreover, the classifier is simple than that of [1,3,4].

4. CONCLUSIONS

In this paper, a new classifier uses dimensionality-independent distortion measure is presented. The classifier is simpler than that of other CVQ-based classifiers. The proposed method can achieve good perceptual quality and preserve edges very well. The bits can be further reduced by removing the redundancy between blocks, especially the high interblock correction in the shade class and the edge patters in the active classes.

REFERENCES

[1] B. Ramamurthi and A. Gersho, "Classified Vector Quantization of Images," IEEE Trans. Commun., Vol. COM-34, no. 11, pp. 1105-1115, Nov. 1986.

[2] M. K. Quweider and J. B. Farison, "Classified Vector Quantization Using Principal Components," Electronics Letters, Vol. 34, no. 6, pp. 538-540, March 1998.

[3] D. S. Kim and S. U. Lee, "Image Vector Quantizer Based on a Classification in the DCT Domain," IEEE Trans. Commun., Vol. 39, no. 4, pp. 549-556, Apr.1991.

[4] J. W. Kim and S. U. Lee, "A Transform Domain Classified Vector Quantizer for Image Coding," IEEE Trans. Circuits Syst. Video Technol., Vol. 2, No. 1, pp. 3-14, Mar. 1992

[5] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, New York: Kluwer-Academic, 1991.

[6] L. M. Po and C. K. Chan, "Directionally Classified Image Vector Quantization Using Walsh Hadamard Subspace Distortion," Electronics Letters, Vol. 27, no. 21, pp. 1964-1967, Oct. 1991.

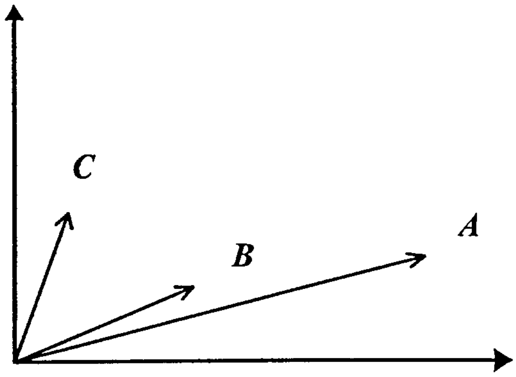


Fig. 1. Illustration of similar vectors.

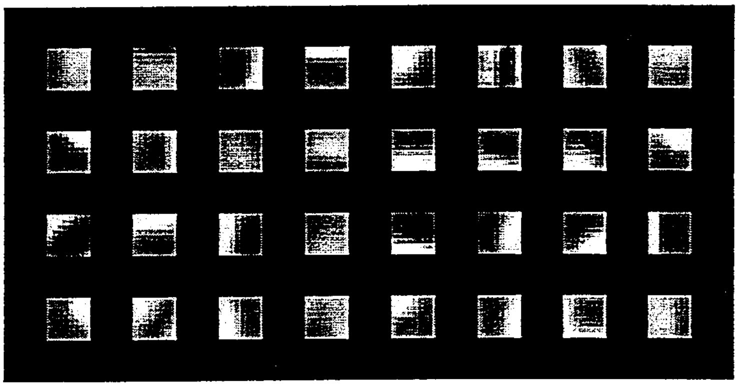


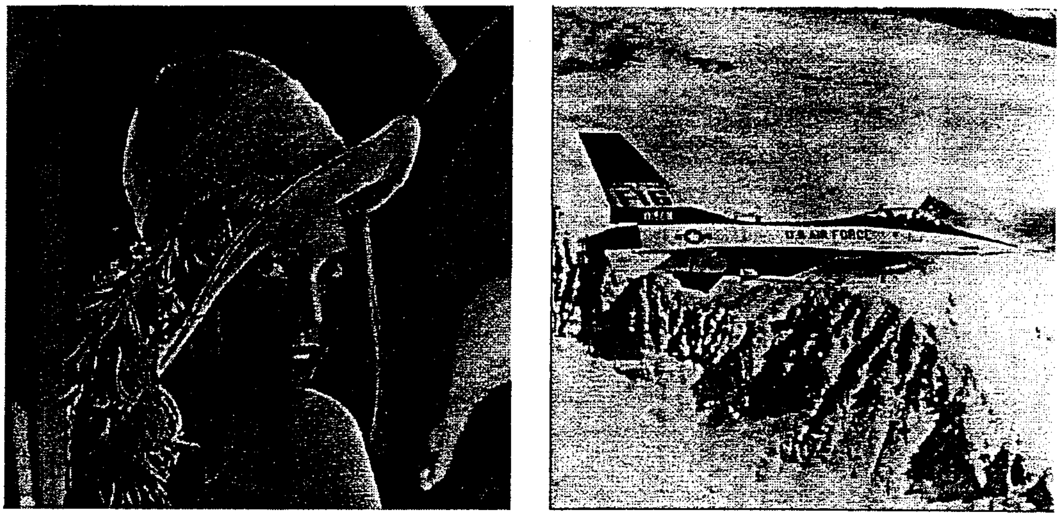
Fig. 2. Magnified classvectors of 32 classes for active vectors ($k = 64$).

(a)

(b)

Fig 3. The original images. (a) Lena. (b) F-16.

1998 International Computer Symposium
Workshop on Image Processing and Character Recognition
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.



(a) (b)

Fig. 4 The reconstructed images with 17 classes. (a) Lena image at 0.618 *bpp* and *PSNR* = 32.74 dB. (b) F-16 image at 0.631 *bpp* and *PSNR*=31.49 dB.

Table 1 The direction-based CVQ coding results using different classvectors.

Image / (M+1) classvectors	Lena		F-16	
	Bits/pixel (*bpp*)	*PSNR* (dB)	Bits/pixe (*bpp*)	*PSNR* (dB)
17	0.568	31.92	0.584	31.16
	0.618	32.74	0.631	31.49
	0.680	33.12	0.694	31.96
33	0.581	32.01	0.598	31.88
	0.631	32.99	0.661	32.44
	0.693	33.56	0.709	32.92

Table 2 Comparison the coding results between our proposed method and other CVQ-based techniques for different block dimension.

Image / Methods	Lena		F-16	
	k (dimension)	*PSNR* (dB)	k (dimension)	*PSNR* (dB)
Proposed	16	32.74	16	31.29
	64	28.88	64	27.96
CVQ in [1]	16	29.79	16	29.31
	64	25.57	64	25.10
DCT-CVQ in [3]	16	31.01	16	30.32
	64	26.12	64	25.98