EUSIPCO-88 Sept. 1988

EUSIPCO-88

TRANSFORM TRELLIS CODING OF IMAGES AT LOW RATES WITH BLOCKING EFFECT REMOVAL*

S. Farkash[1], W.A. Pearlman[2], and D. Malah[1]

ABSTRACT

A transform trellis coding scheme is applied to encode images at rates below 1 bit per pel, with high SNR values and very good subjective performance. A discrete cosine transform (DCT) is used to decorrelate the data samples, and a trellis diagram is used to further encode the transformed coefficients. To overcome image non-stationarity a clustering algorithm is used to segment the transformed image into clusters and a separate trellis diagram is then constructed for each cluster. To further improve the image quality, and particularly to reduce the blocking-effect, a scalar quantizer is used to quantize large magnitude coefficients of the error image in the transform domain. the performance of the proposed scheme was found to be very good, and is better then other recently reported schemes.

I. INTRODUCTION

The transform trellis coding (TTC) technique was proven by Mazor and Pearlman [1] to be asymptotically optimal for stationary Gaussian sources and the squared error distortion measure. In [2] Mazor and Pearlman applied this technique to encode speech, and the same basic ideas were used in [3] to encode images. The main problem encountered in [3] was the "blocking effect" which is usually accompanying transform coding schemes at low rates. This paper presents an improved transform trellis coding scheme, which encodes images at rates below 1 bit per pel and alleviates the "blocking effect" problem.

The proposed TTC coding scheme performs two passes on the source image. On the first pass the image is divided into s blocks. Each block undergoes a two dimensional discrete cosine transform (DCT), which is considered to be the best suboptimal transformation method from the viewpoint of data compactness and ease of implementation [4,5]. The produced transformed coefficients are less redundant and are almost uncorrelated (the coefficient are fully uncorrelated in the case of a finite Markov source in the asymptotical case when the block size goes to infinity [6]).

The s transformed blocks are classified by a clustering algorithm to c classes (clusters). For each class an estimated spectrum is produced and used to construct the trellis diagram for that class. The same estimated spectrum is sent to the receiver as side information to enable the construction of the decoder.

On the second pass each transformed block is reordered as a vector. The encoder then searches the trellis, which is populated by a random selection of Gaussian code vectors, for a code vector which minimizes the squared-error between the source vector and the code vectors. The corresponding path-map on the diagram is forwarded to the receiver.

To remove the blocking-effect which is typical to transform coding methods, and was also observed in [3], an error vector is constructed by subtracting the code vector, composed of code words residing along the above path-map, from the source vector.

A scalar quantizer is used to quantize large magnitude terms in the error vector, which are the main contributors to this undesired effect. When the encoding process is completed, the values and locations of the error vector coefficients that were quantized (by the scalar quantizer), are coded and sent to the decoder as additional side-information.

The decoder which constructs the same trellis diagram, using the estimated spectrum side-information, uses the received path-map to uniquely determine the code vectors. The quantized error coefficients are added to the code vectors, at the proper locations, to produce the reconstructed vectors. These vectors are reordered into two-dimensional blocks and inverse transformed to obtain the reproduction image.

The remainder of this paper describes in more details the proposed coding scheme and the results achieved. Section II introduces the basic transform trellis coding scheme. Section III describes practical implementation considerations and the additional means used to reduce the blocking-effect. Simulation results are contained in section IV, and finally, a summary and conclusions are presented in section V.

II. TRANSFORM TRELLIS CODING

Fig. 1 presents the basic transform trellis coding scheme implemented on images. The first stage of the scheme, the transformation stage, is composed, in principle, of a two-dimensional optimal transformation (the Kahrounen-Loeve Transform - KLT) which is applied to a block Z of the image to produce a block U of uncorrelated elements. The transformed block U is reordered as a vector $\underline{u}$ which serves as an input data vector to the second stage of the scheme - the coding/decoding stage. This stage is based on a trellis diagram shown in Fig. 2. The diagram is composed of nodes and branches which are populated with code words. The coding process is performed as follows:

For each vector $\underline{u}$ the encoder searches the trellis for the optimal path, i.e the path along the trellis which minimizes the squared-error between the source words, composing the vector $\underline{u}$, and the code words residing along paths through the trellis. The path-map Q corresponding to the chosen path is sent through the channel (assumed here to be noiseless) to the decoder. The decoder, which stores the same trellis diagram, uses the received path-map to get the reconstructed vector.

The basis of this coding scheme is established in Mazor and Pearlman works [1,2], and the steps concerning its construction can be found in [3].

---

* This work was supported in part by the Ministry of Industry and Commerce (050-0520).

[1] S. Farkash and D. Malah are with the Electrical Eng. Dept., Technion - I.I.T., Haifa 32000, Israel.

[2] W.A. Pearlman is with the Electrical, Computer and System Eng. Dept., RPI Troy, N.Y. 12180.

III. PRACTICAL IMPLEMENTATION CONSIDERATIONS

The implementation of the transform trellis coder presented in the previous section suffers from several shortcomings. First, the computation of the optimal KLT is regarded as a too complicated and expensive task, as it requires computation of the covariance function and its eigenfunctions and eigenvalues. Another weakness of the scheme is that it does not take into account the non-stationary nature of the images we usually have to code. This non-stationarity prevents the transform trellis coder from obtaining good performance for every block in the image. The blocking-effect problem which is usually encountered in transform coding at low bit rates is also a problem in the above scheme. This section presents our solutions to these difficulties.

A. SYSTEM DESCRIPTION

The system we implemented (Fig. 3) performs two passes on the input image. In the first pass, the image of size $[P' \times P']$ (total of P pels) is divided into s blocks of size $[N' \times N']$ (total of N pels). Each block undergoes a two dimensional Discrete Cosine Transform (DCT) and produces an equal size block of transform coefficients. The DCT was used because it overcomes the main disadvantages of KLT, i.e its signal dependence and its complexity. Furthermore, the DCT is regarded as the best sub-optimal transform [5,6]. The algorithm we applied for computing the DCT is the one presented in [4]. The s transformed blocks are then classified into c clusters according to an algorithm which we will discuss later on. For each of the clusters that were created a cluster spectrum $\hat{S}_i$, $i=1$ to c , is estimated, and used for for the code construction. The estimated spectrum, along with the DC coefficient of each transformed block, are sent as side information to the decoder.

In the second pass, the DCT coefficients are reordered as a vector, and each vector is encoded by the trellis which corresponds to the cluster it belongs to. Along with the trellis encoding process, an error vector is constructed and those of its coefficients having large magnitude are quantized, coded, and sent to the receiver. At the receiver the above operations are reversed as depicated in fig 3b. A more detailed discussion of the system can be found in [3].

B. CLUSTERING

In the discussion of the optimal coding system, we have assumed the source to be stationary and Gaussian. This assumption does not hold for most of the natural images. To solve this problem we divide the image blocks into several clusters considering similarities between their spectrum (without the DC term). Then, we estimate a spectrum for each cluster and use it to construct a trellis diagram for that cluster. Assuming stationarity among blocks belonging to the same cluster, the trellis diagram is considered optimal.

The clustering problem deals with the partitioning of a set of s vectors $W=\{\underline{w}_i \mid i=1,\ldots,s\}$ into c clusters $C_k,\ k=1,\ldots,c$, each represented by its mean value $\underline{m}_k$:

$$\underline{m}_k \triangleq \frac{1}{\gamma_k}\sum_{i=1}^{\gamma_k} \underline{w}_i \ , \underline{w}_i \in C_k \qquad (1)$$

where γ_k is the cardinality of the k-th cluster.

The problem can be stated as finding, among all possible partitions, the partition which minimizes the cost function

$$J(C)=\sum_{k=1}^{c}\sum_{i=1}^{\gamma_k}\delta(\underline{w}_i,\underline{m}_k)=\sum_{k=1}^{c} J_k(C) \ , \underline{w}_i \in C_k \qquad (2)$$

where $\delta(\underline{w}_i,\underline{m}_k)$ is a measure of distance between a vector $\underline{w}_i$ and a cluster representative $\underline{m}_k$.

The algorithm used in our scheme is an improved version of the dynamic clustering algorithm that was introduced in [7 ch 11]. The algorithm which was employed circumvents the need for setting an initial partition by using the splitting method proposed in [13]. This method was originally developed to construct codebooks for vector quantizers (VQ). The result of the clustering process is a segmented image where each block belongs to a cluster of blocks having similar spectra. Fig. 5c demonstrate the segmented image for $c=16$ clusters

C. CODE ADAPTATION

The clustering process segmented the image into c classes with different spectral characteristics. If the same rate will be allocated to all the clusters, the scheme would not be optimal, because clusters containing high energy blocks need larger rate then clusters with smooth blocks. Therefore, we had to perform rate allocation among the clusters, such that for a given total rate we will get the best performance (i.e the smallest MSE). The average rate (in bits/pel) allocated to every block in cluster i, $i=1,\ldots,c$, is given by:

$$R_i=\frac{1}{N}\sum_{k=1}^{N}\max\left\{0,\frac{1}{2}\log\left[\frac{\hat{S}_i^k}{\theta}\right]\right\} \qquad (3)$$

where N is the number of pels in a block, θ is the parameter which was determined such that the total rate would be the desired one, and $\hat{S}_i^k$ is the kth coefficient of the estimated spectrum vector of cluster i.

D. SEARCH ALGORITHM

Trellis encoding requires a search of the trellis for the path-map, which the code vectors that populate it provide an adequate match to the input sequence of source vectors. The algorithm used to perform the search is the M - Algorithm which is a sub-optimal version of the Vitterbi algorithm [12].

E. BLOCKING-EFFECT REMOVAL

The above coding scheme, with the clustering module, is expected to achieve good results if each cluster contains blocks which are all from the same distribution, i.e the cluster is stationary. However, this assumption is usually not satisfied. There are blocks in the image, particularly those which contains edges between regions, that are not stationary. The effect of this non-stationarity on the reconstructed image is a blocking-effect which is usually encountered in transform coding at low bit rates.

To overcome this phenomenon an error vector is constructed by subtracting the code vector, which comprises of code words residing along the chosen path-map on the trellis diagram, from the source vector. A scalar quantizer is used to quantize the large magnitude terms in the error vector. The magnitudes and locations are coded efficiently by Huffman codes.

The correction of the large magnitude terms in the error vector significantly reduces the blocking-effect. The additional complexity is very small, because the decoding process needed to create the reconstructed image is a by-product of the search in the trellis, and the reconstructed vectors exist in the encoder anyway. The complexity accompanying the Huffman coder is negligible.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed coding scheme we simulated it on a VAX/750 computer with a Gould IP8500 image display system. We use the woman's head and shoulder image of size 256×256 pels shown in Fig. 4a, as test image #1, and one quarter (the lower right corner) of the same image at a higher resolution (512×512) shown in Fig. 5a. Fig. 4(b,c,d) shows reconstructed images of test image #1 at three different rates, Fig. 5b show a reconstructed image of test image #2. The subjective quality of the reconstructed images in Fig. 4 are rated according to our best judgement, from very good quality - for the image in Fig. 4b, to good quality - for the image in Fig. 4c, to more than fair quality - for the image in Fig. 4d. The SNR ($10\log(255^2/MSE)$) achieved in our simulations for test image #1 are presented in table 1, along with the performance of other coding systems on the same test image which were recently published. The table shows that the performance of our coder is better then the performance obtained in [9,10,11]. The SNR of the reconstructed image resulting from the more redundant image (Fig. 5b) was $37.4\,dB$ at rate of $R = 0.41\ bit/pel$ The subjective quality of this image is very high, and one can not see any difference between the original and the reconstructed images. The superiority of the more redundant image is expected due to its much larger redundancy.

V. SUMMARY AND CONCLUSIONS

The transform trellis coding scheme described in this work encodes cosine transform coefficients on a trellis diagram. The image is classified by a clustering algorithm into several clusters. Each cluster is coded by a different trellis diagram, constructed by using an estimated spectrum of that cluster. To reduce the blocking-effect encountered in transform coding schemes, a scalar quantizer is used to encode large magnitude coefficients of the error image in the transform domain.

The trellis coder is used in our work due to its optimality for Gaussian-Markov signals, which are considered to properly model real images, and because its implementation is simpler than tree and block coders (for the same performance). The deviation of the image from the assumed model is compensated by the clustering process and the scalar quantizer used. The performance obtained with the proposed scheme is superior to those reported in the literature with other coders [9,10,11]. The complexity of the scheme is very high. However, most of the complexity is in the encoder part of the scheme. This fact enables employment of the proposed scheme in image storing and retrieving systems, where the decoder simplicity is a vital property. With the continued development of VLSI technology, this scheme could in the future be implemented also in real-time systems.

REFERENCES

[1] B. Mazor & W. A. Pearlman, "A Trellis Code Construction & Coding Theorem for Stationary Gaussian Sources", IEEE Trans on Information Theory Vol IT-20 No. 3, May 1983.

[2] B. Mazor & W. A. Pearlman, "An Optimal Transform Trellis Code With Application To Speech ", IEEE Trans on Comm. Vol. COM-33 No. 10, October 1985.

[3] S. Farkash "Transform Trellis Coding of Images" Proceedings IEEE Tel-Aviv, April 1987, pp. 2.5.4/1-2.5.4/4.

[4] J. Makhoul "A Fast Cosine Transform in One & Two Dimensions", IEEE Trans. on Acoustics & Signal Processing, Vol. ASSP-28, No. 1, pp. 27-34, Feb 1980.

[5] M. Hamidi & J. Pearl, "Comparison of the Cosine and Fourier Transforms of Markov-1 Signals",IEEE Trans. on Acoustics & Signal Processing, pp. 428-429, Oct 1976.

[6] Y. Yemini & J. Pearl, "Asymptotic Properties of Discrete Unitary Transforms", IEEE Trans. on Pattern Analysis & Machine Intelligence, Vol. PAMI-1 No. 1, pp. 366-371, Oct. 1978.

[7] P. A. Devijuer & J. Kittler "Pattern Recognition A Statistical Approach", Prentice-Hall London, 1982.

[8] W. K. Pratt, Digital Image Processing New York: Wiley Interscience 1978.

[9] J. W. Woods & S. D. O'Neil, "Subband Coding of Images", IEEE Tran. On Acoustics and Signal Processing, Vol. ASSP-34, Nov 1986.

[10] H. Hang & J. W. Woods, "Predictive Vector Quantization of Images", IEEE Trans on Comm. Vol, COM-33, No. 11, Nov. 1985.

[11] W. A. Pearlman, M. M. Leung & P. Jakatdar, "Adaptive Transform Tree Coding of Images" ICASSP 85 pp. 145-148, Tampa, Florida.

[12] N. S. Jayant & P. Noll "Digital Coding of Waveforms", Engwood Cliffs N.J., Prentice Hall, 1971.

[13] Y. Linde, A. Buzo & R.M. Gray, "An Algorithm for Vector Quantization", IEEE Tran. on Comm., Vol. COM-28, No. 1, Jan. 1980, pp. 84-95.

Table 1 - Performance Comparisons

Reference		Rate [bits/pel]	SNR [dB]
[9]		0.67	30.9
		1	32.5
[11]		0.74	32.4
		1	32.8
[10]		0.5	27.5
		1	30.9
proposed coder	(Fig 4b)	0.92	35.2
	(Fig 4c)	0.76	33.6
	(Fig. 4d)	0.61	32.6

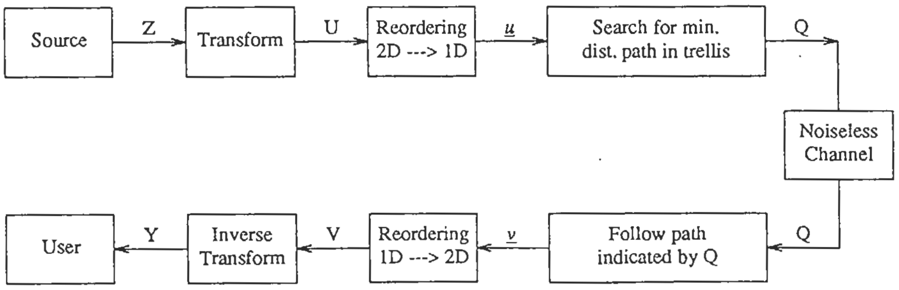

Fig.1 - TTC Basic Scheme

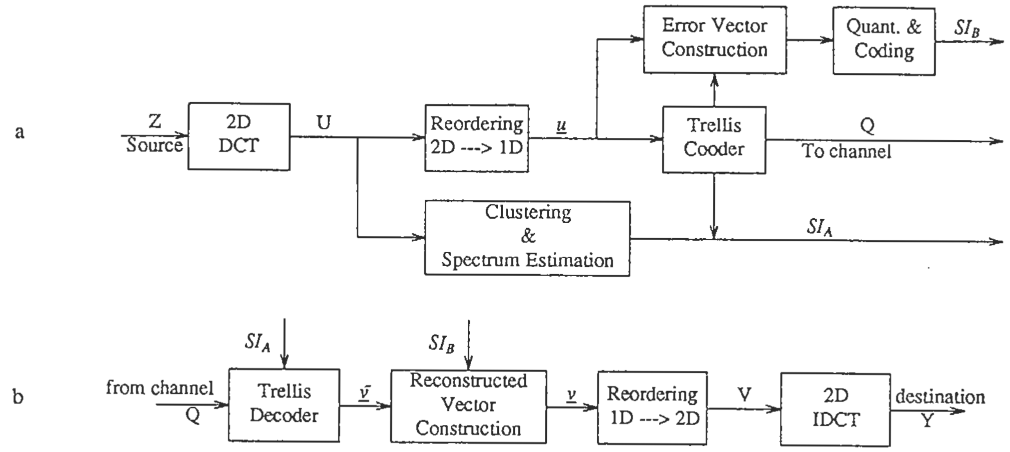

Fig. 3 - (a) Transmitter (b) Receiver

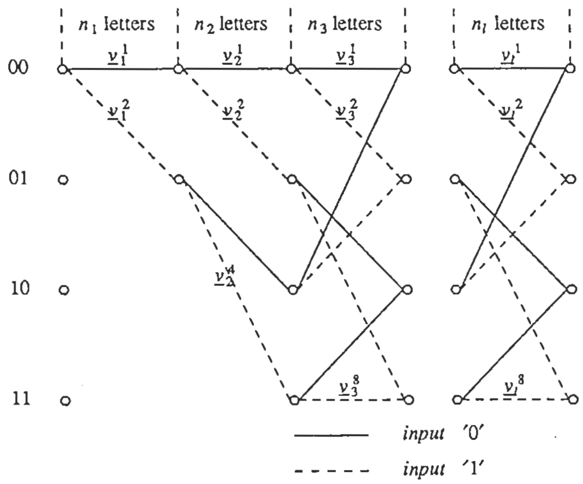

Fig. 2 - Trellis Diagram

Fig. 4 - Results for 256×256 pels image size.

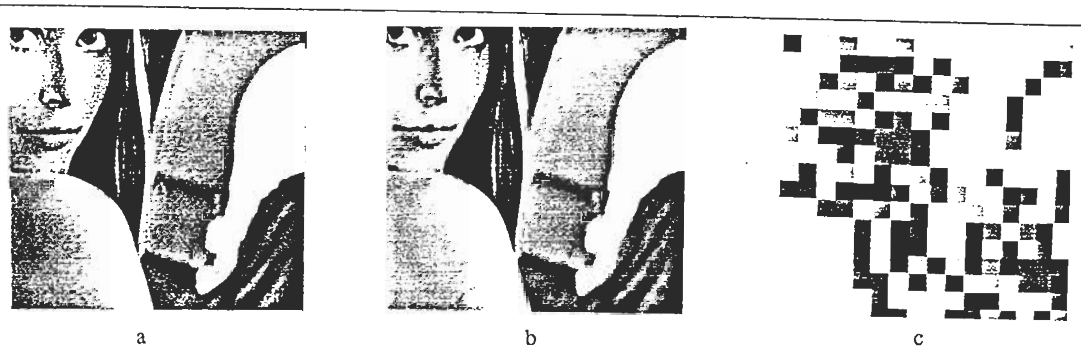

Fig. 5 - Results for 512×512 pels image size.