NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

Transform Image Coding by Vector Quantization

PAUL BOUCHER* MORRIS GOLDBERG*

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS - LABORATOIRE IMAGE
* Both presently on leave from the Electrical Engineering Departement University of Ottawa, OTTAWA, CANADA.

RESUME

Une nouvelle méthode de compression de données images par transformation et quantifiquation de vecteurs est presentée. Elle offre un taux de compression élevé de même qu'un décodage très simple. Pour le codage de la parole, la quantifiquation vectorielle a été démontrée comme étant plus efficace que la méthode classique de quantifiquation scalaire. Des experiences antérieures menées sur des données images multi-spectral (MSS) et à niveaux de gris ont aussi demontré son efficacité.

La méthode classique de compression d'images par transformation dans un domaine spectral , utilise la quantifiquation scalaire des différents coefficients; ceux-ci étant quantifiés selon une série de règles compliquées. La quantifiquation vectorielle traite un bloc de P par P pixels, ou coefficients de transformation, comme un vecteur de dimension P^2 . Ces vecteurs sont ensuite classifiés (clustering) selon une segmentation de l'espace vectorielle de dimension P^2 en 2^n nuées (cluster). Chaque vecteur est ensuite remplacé par le vecteur moyen de chaque nuée, ou par d'autres données statistiques. Ainsi seule une étiquette de N bits est nécessaire au codage de chaque vecteur. Le décodeur utilise cette étiquette comme l'index d'un tableau contenant les données statistiques de chaque nuée, pour regénérer chaque vecteur.

On présente ici la quantification vectorielle de coefficients de transformation. On comparera le codage d'intensités lumineuses avec le codage de coefficients de transformation. Cette nouvelle technique de codage répond bien aux besoins des systèmes d'archivage et de transmission de données images, tel le videotex.

SUMMARY

A transform image compression technique, combining high compression rate, good reconstruction subjective quality and simple decoding, based upon vector quantization is investigated.

Vector quantization has been shown to yield better results than classical scalar quantization techniques, in speech coding. Investigations on Multi-Spectral-Scanner and Black-White imagery have also demonstrated its effectiveness for compression. Here requantization was performed on the pixels intensity value.

The classical transform coding technique employs scalar quantization, where the coefficients are individually quantized or simply ignored using some complicated set of rules. Vector quantization, instead processes blocks of P by P pixels, or transform coefficients, treated as a P^2 dimensional vector. The vectors are clustered; this effectively segments the P^2 dimensional vector space into 2^n clusters. Each vector is then represented by the cluster mean, or some other statistics. The decoder uses this label to address a look-up table, containing the 2^n clusters statistics, to regenerate each vector.

Results will be presented comparing different vector quantization coding strategies. This technique has potential in image retrieval systems were simple and fast image reconstruction is a requirement, as in videotex systems.

INTRODUCTION

With the introduction of digital techniques in image processing, there has been great interest in image coding for data reduction. Archiving and transmission of imagery would both benefit tremendously from such coding, as a typical image requires 10^6 to 10^8 bits. Many efficient coding schemes have evolved <1>, two well known are Adaptive Block Transform <2> and predictive coding (DPCM) <3>. Recently a new strategy has been proposed: vector quantization by clustering techniques. Hilbert <4> and Lowitz <5> studied extensively its use for Multi-Scanner-Spectral (MSS) data and Abut and Gray <6> for speech.

In this paper we present a new adaptation of vector quantization specifically adapted for archiving and transmission systems. It offers fast, inexpensive and progressive image reconstruction. These latter advantages are provided by the table look-up decoding feature of vector quantization and by the S-Transform <7> hierarchical image descriptors. Bit rates below 1. 8 bits per pixel combined with good subjective quality are obtained with this encouraging new coding scheme.

In the first and second section of this paper a description of the S-Transform and of vector quantization are presented. In the third and fourth section our scheme is described and compared to Gersho's <8> simple and efficient vector quantization technique. Finally from the encouraging results future research possibilies are presented.

S-TRANSFORM IMAGE CODING

The S-Transform is an hierarchical approach to description of images, originally developed for archiving. A derivative of the Hadamard transform, it decomposes an image matrix, stepwise into a coarse picture matrix and an ordered set of detail information matrices. The Hadamard transform is applied to non-overlapping cells of two by two pixels. This results in a sum coefficient (corresponding to the luminosity mean of the cell) and coefficients for the vertical, horizontal and diagonal differences. The combination of the sum coefficients represent exactly the original picture at half resolution.

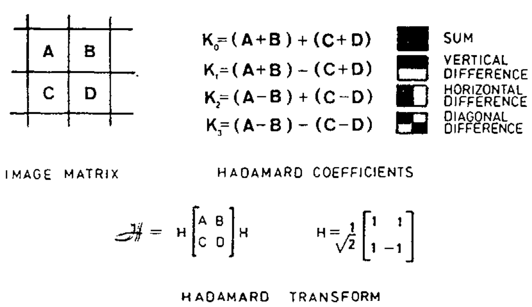

Figure 1 - S-Transform Image Coding : The basic Hadamard transform on a sub-picture of 2 by 2 elements is repeated successively on resulting sum coefficients.

In subsequent steps the matrix of sum coefficients is transformed using the same algorithm. As for the difference components of each cell, these are simply quantized and stored. At the last step a single sum coefficient remains, corresponding to the overall intensity mean of the picture.

The reconstruction process is the reverse. It is started by first retrieving the coarse picture information and then the corresponding transform difference coefficients. The four coefficients of each subpicture are inversely transformed to obtain a finer resolution picture. The algorithm is repeated until a full resolution image is obtained. Thus a coarse image representation is instantaneously received with gradually detail level. This allows for browsing rapidily through stacks of high resolution pictures.

This technique is best adapted to both the data flow needs of the human observer and low resolution image display terminals. Furthermore only add, subtract and shift operations are required for decoding.

VECTOR QUANTIZATION OF IMAGERY

Vector quantization is a simple and powerful coding scheme which can be applied to imagery in many different ways. A technique developed by Gersho <8> is given as example. Basically a single channel (black and white) image is partioned into cells of P by P pixels. Each cell is regarded as a P^2 dimensional vector. The vectors are clustered, which effectively segments the P^2 dimensional space into 2^n regions or clusters. Each vector comprised in a cluster is coded by a N bit label. Each cluster is then coded by its statistics such as its mean vector. Reconstruction is obtained by first transmitting the clusters statistics codebook, followed by the vector labels used as indices to access the codebook.

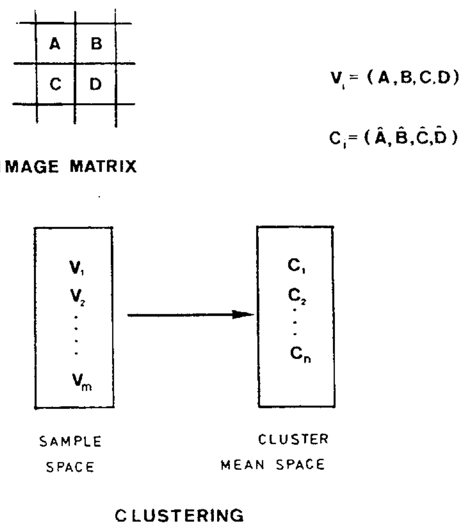

Figure 2 - Image Coding Using Vector Quantization. At the first step, sub-pictures, vector V_i, are clustered. Each vector is then replaced by the cluster mean, C_i.

Thus only a small receiver frame buffer is required, since table look-up decoding can be performed at video rates. Gersho obtained subjectively acceptable pictures below one bit per pixel, hence a reduction in frame buffer size by more than a factor of eight.

PROPOSED NEW CODING SCHEME

By vector quantizing the difference components at each step of the S-Transform, a powerful image compression scheme results. The difference coefficients of each cell at one step are regarded as a three- dimensional vector. These are clustered using a modified version of the classical iterative K-Means algorithm <9>. Hilbert <4> has shown that a more efficient clustering technique, such as ISODATA <9>, isn't worth the extra computational load.

One important requirement of the K-Means algorithm is the initial cluster center seeding. It was found that by preprocessing the data by a Graph-Theoritic mode and cluster seeking algorithm <10>, fast convergence is assured. This extra computational load is minimal: approximately one K-Means iteration. The Euclidian metric was found to be the best compromise between processing speed and subjective results. A simple cluster difference measurement between the last and present iteration, is used for measuring cluster convergence. This convergence criteria, depending on both the distribution and dynamic range of the data, is determined by simple heuristics.

Although K-Means clustering still involves considerable computational costs, it is performed only once per image and is easily amortized over a large number of receivers, these requiring a small frame buffer and on the fly look-up table decoding.

As a first study of this new coding technique, the Hadamard difference coefficients, contained in the codebook, are simply linearly quantized to their full dynamic range.

EXPERIMENTAL RESULTS

The first test picture is of type head and shoulders (figure number 3 (a)) choseen because of its prevalent use in image coding technique testing. It is a 256 by 224 pixels picture, each pixel PCM coded with eight bits.

Figure number 3 (b) is coded using Gersho's scheme, described previously. For this picture a two by two pixel cell is used, a choice resulting from a compromise between subjective quality and computational load. By grouping the vectors into 64 clusters, an effective bit rate of 1. 54 bits per pixel is obtained, for a signal to noise ratio (SNR) of 29. 6 dB.

Figure number 3 (c) is coded using the S-Transform vector quantization method. The difference coefficient vectors at each step are coded with 16, 64, 128 and 128 clusters respectively for the first, second, third and fourth S-Transform step. This particular coding sequence was found to be the best overall choice. Although three more steps of transformation are still possible they wont contribute to anymore data reduction. An effective bit rate of 1. 8 bits per pixel is obtained, for a corresponding SNR of 29. 9 dB. This bit rate could be reduced by quantizing each difference coefficient down to eight bits, since the Hadamard transform is very robust to round-off errors <11>. Further bit rate reduction could be gained by cascaded clustering <4> of the cluster representative vectors (mean vector for our case) since high redundance exists between codebooks at each step. By combining both of these latter strategy 0. 2 bit per pixel could be gained. Thus a lower limit of 1. 6 bits per pixel, which is only 0. 05 bit per pixel more than Gersho's technique, could be achieved for the added advantage of progressive decoding.

These pictures should be juged at a certain distance, since the eye tends to smooth out spatial information when observing the pictures on a monitor screen. Because of this and poor photographic reproduction only a approximate idea of the results can be shown. To better represent the quantization effects of each method, a false colour version of each coded image is shown. Here a false colour is assign in a pseudo-random manner to each grey level intensity. For Gersho's technique it is seen that large regions are smoothed out, thus producing false contouring, as under the eyes. Apart from this, overall reproduction is good.

For the S-Transform vector quantization technique no false contouring occurs, but a noisy texture results. The objectionable nature of each artifact is dependent on the observer.

In figure number 4 a graphic type picture is coded using the same schemes. Again note the same type of behaviour.

CONCLUSION

This paper is only a fragment of an ongoing investigation on the use of vector quantization to imagery. A number of modifications for improving both the subjective quality and processing speed, for this particular scheme, are now being explored and will be reported in a later publication.

In this paper we have established the viability of vector quantization of transform coefficients as a novel technique in progressive coding of images. Its strength lies in its ability to fully exploit the spatial redundancy present in images. We have shown that a bit rate of 1. 8 bits per pixel, with only limited degradation, is easily achievable.

ACKNOWLEDGEMENT

We wish to thank the Canadian Communication Research Center for their assistance and particularly Dr. Seymour Shlien. We also wish to thank the Ecole Nationale Superieure des Telecommunications of Paris for receiving us during the course of this research.

REFERENCES

<1> Habibi A.,'Survey of Adaptive Image Coding Techniques'. IEEE Trans. on Communications, Vol. COM-25, No. 11, november 1977. pp. 1275-84.
<2> Chen W. H. and Smith H.,'Adaptive Coding of Monochrome and Color Images', IEEE Trans. on Communications, Vol. COM-25, No. 11, november 1977, pp. 1285-92.
<3> O'Neal J. B.,'Predictive Quantizing System (DPCM) for the Transmission of Television Signals'. BSTJ. Vol number 45, pp. 689-721. May-june 1966.
<4> Hilbert E. E.,'Cluster Compression Algorithm a Joint Clustering/Data Compression Concept'. Jet Propulsion Laboratory Publication 77-43. 143 pages.
<5>Lowitz G. E.,'Image Data Reduction'. Proc. International Conf. on Spacecraft On-Board Data Management. Nice, Oct. 78, pp. 291-4
<6>Abut H., Gray R. M. Rebolledo G.,'Vector Quantization of Speech and Speech-like Waveforms'. IEEE Trans. on ASSP. Vol. ASSP-30, June 1982.
<7>Wendler T. H., Meyer-Ebrecht D.,'Proposed Standard for Variable Format Picture Processing and a Codec Approach to Match Diverse Imaging devices'. SPIE Vol. 318 (part 1), Picture Archiving and Communications Systems (PACS) for Medical Applications. 1982, pp. 298-305.
<8>Gersho A., Ramamurthi B., Image Coding using Vector Quantization'. IEEE ICASSP. May 1982, pp. 428-431.
<9> Tou J. T., Gonzalez R. C.,'Pattern Recognition Principles', Addison-Wesley Publishing Company, 1974. pp. 94-109.
<10> Narendra P. M. and Goldberg M. "A Non-Parametric Clustering Scheme for Landsat". Pattern Recognition. Vol number 9 1977, pp. 207-215.
<11> Knab J. J., "effects of Round-off Noise on Hadamard Transformed Imagery", IEEE Trans. on Communications. Vol COM-25, No. 11, pp. 1292-1294, Nov 1977.

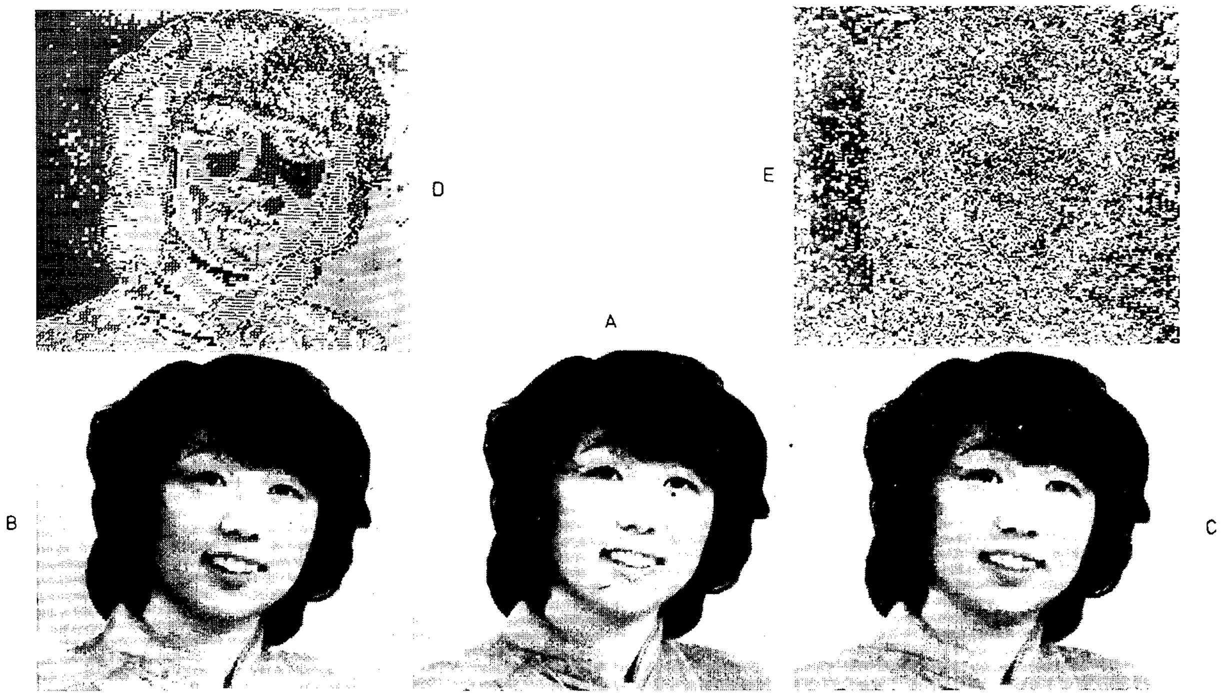

Figure 3 - Simulation results for compression (see text) (a) Original.
(b) Gersho's technique, coded at 1.54 bits par pixel, SNR = 29.6 dB
(c) S- Transform -Vector Quantization scheme coded at 1.8 bits per pixel , SNR = 29,9 dB.
(d) False colour version of (b)
(e) False colour version of (c)

Figure 4 - Simulation results for compression (see text) (a) Original.
(b) Gersho's technique, coded at 1.54 bits par pixel, SNR = 22.335
(c) S - Transform - Vector Quantization scheme, coded at 1.8 bits per pixel, SNR = 21.295
(d) False colour version of (b)
(e) False colour version of (c)