Tree-Structured Lattice Vector Quantization

Vincent Ricordel and Claude Labit

IRISA/INRIA Rennes, Campus de Beaulieu
35042 Rennes Cedex, France
e-mail: ricordel@irisa.fr, labit@irisa.fr

ABSTRACT

In [14][15] we introduced a new vector quantizer (VQ) for the compression of digital image sequences. Our approach unifies both efficient coding methods : a fast lattice encoding [3] and an unbalanced tree-structured codebook design according to a distortion vs. rate trade-off [2][16]. This tree-structured lattice VQ (TSLVQ) is based on the hierarchical packing of embedded truncated lattices. Now we investigate the determination of the most efficient lattice respectively to this method. We also describe a fast test which permits to detect the input vectors whose norm is above than the maximum allowed by the TSLVQ. Finally we analyse experimental results applied to image sequence with our VQ taking place in a region-based coding scheme for a videophone application.

1 Introduction

This paper deals with the design of a VQ for the compression of digital image sequences and we consider a differential or hybrid image vector source (i.e : vectors of (transformed) motion-compensated prediction errors). Such a source, whose distribution is commonly modelized by a multivariate generalized gaussian function, is always a nonstationary signal. So a temporal updating of the VQ codebook [8][11] has to be designed in order to fit the spatiotemporal statistics of the image source. A training procedure, for the VQ codebook design from source representative training sequences, performs this temporal replenishment. But the codebook design and the corresponding encoding-decoding algorithm have to be very fast.
The computation complexity of usual classification methods [9][7] presents a limitation on their applicability. A lattice VQ [6][1], for which encoding is fastest, is adapted only for a stationary signal whose distribution permits to truncate the lattice.

2 Summary of the previous study

To overcome these drawbacks, we proposed a new lattice VQ based on the hierarchical packing of embedded lattices and, considering the lattices for which Conway and Sloane determined fast quantizing and decoding algorithms [3] (i.e : $Z^n/n \geq 1, D_n/n \geq 2, E_8, \Lambda_{16}$), we described in [14][15] :
- the lattice truncating method in order to embed it (by contracting it) in its voronoï cell ;
- the hierarchical set organization of truncated lattices, by suitably shifting their scales in order that a lower scale lattice is embedded into the next higher scale one ;
- the corresponding multistage quantization procedure ;
- the tree pruning or a greedy tree growing approach used in order to design, by training, an unbalanced tree-structured codebook according to a distorsion vs. rate criterion [2][16].

However the greedy tree growing approach is more suitable when the vector dimension, and consequently the tree airies number, is high : for each iterative process loop, only a single splitted leaf is added to the tree (only one new truncated lattice is embedded). But this later method, which explored the short term effect of extending the tree, is suboptimal considering a tree pruning approach where, after a complete tree design, some branches are successively removed.
In practice the lattice truncation energy is chosen minimal in order to restrict the number of embedded lattice points. So the TSLVQ codebook design achieves a progressive vector space partition involving an efficient entropy coding of the tree leaves.

3 Optimal lattice determination

This optimal lattice is choosen again among the lattices for which Conway and Sloane determined fast quantizing and decoding algorithms [3] (note that D_4, E_8, and Λ_{16} are the best quantizing lattices [4] respectively to their dimension and considering the high-resolution theory).
Figure 1 shows experimental codebook entropy vs. distorsion curves and we compare Z^4 with D_4, Z^8 with E_8 : it appears that Z^n performs better than the others : lower distorsion and rate are obtained. Of course, a low rate is achieved with higher vector dimensions.

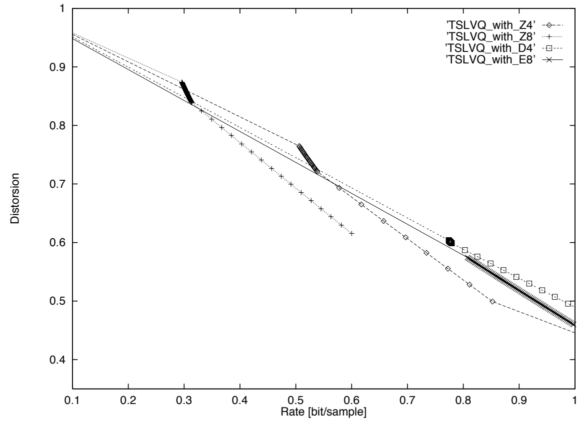

Figure 1: *TSLVQ using a tree growing approach : experimental distorsion vs. rate curves for an i.i.d synthetic gaussian source (source variance :* $\sigma^2 = 1$)*. The training sequence size is adapted to the vector dimension in order to reach a training ratio higher than 150.*

Z^n is optimal for the TSLVQ because it is the only lattice which realises an optimal packing : the cubic truncated lattice, when we embed it (by contracting it), recovers exactly the cubic voronoï cell.
Figure 2 illustrates the optimal packing advantage : suppose that the source vectors are uniformly displayed into the shaded region. For the truncated cubic lattice, the decrease in distorsion and the increase in rate are equably shared between all the points. For the truncated hexagonal lattice, the six peripheral points are less probable, it induces an increase in rate and a less decrease in distorsion.
Using Z^n implies that the resulting TSLVQ hasn't a space-filling advantage [10], but in practice this gain is very low [10]. On the other hand, Z^n is the least dense lattice, so the truncated lattice points number (the tree airies number) is reduced and the vector space partition is more progressive.

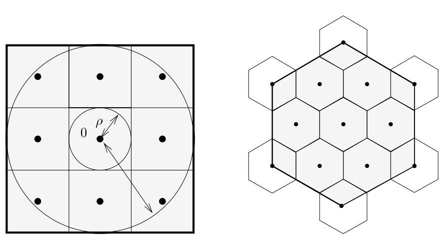

Figure 2: *An optimal packing with the cubic lattice, a nearly optimal packing with the hexagonal lattice.*

4 Input vectors norm test

A training procedure is used in order to design the codebook. Since the input vectors, to be quantized, don't come exactly from the training sequence (TS), it's necessary to verify that their norm is below the maximal allowed. We hence introduce a fast test whose aim is to sort the input vectors.
Let F be the scaling factor used in order to project the source vectors into the first truncated lattice [14] [15] :

$$F = \frac{3 \times \rho}{\sqrt{\mathcal{E}_{max}}}$$

with

$$\mathcal{E}_{max} = \max_{\mathcal{E}} \left\{ \mathcal{E}(\mathbf{x}) = L_2(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 \ / \ \mathbf{x} \in \text{TS} \right\}$$

where ρ is the lattice packing radius [4] ($\rho = 1/2$ for Z^n) and $\mathcal{E}_{max}$ is the maximum energy for the source to be coded. The Z^n voronoï cells totally or partially within the multidimensional sphere of radius $(3 \times \rho)$ constitute the basic truncated lattice conserved for the packing, because :

$$\sum_{i=1}^{n} (F \times x_i)^2 = \frac{(3 \times \rho)^2}{\mathcal{E}_{max}} \times \mathcal{E}(\mathbf{x}) \leq (3 \times \rho)^2 \ / \ \mathbf{x} \in \text{TS}$$

This truncated sub-space is a cube and a vector $\mathbf{u}$, which lies into it, is such as (figure 2 illustrates it) :

$$L_\infty(\mathbf{u}) = \max_{i=1,\dots,n} |u_i| \ \leq \ (3 \times \rho)$$

Consequently, the source vectors which can be quantized by the TSLVQ have to verify (before projection) :

$$L_\infty(\mathbf{x}) = \max_{i=1,\dots,n} |x_i| \ \leq \ \sqrt{\mathcal{E}_{max}}$$

This test using the L_∞ metric is obviously faster than one with the L_2 metric.

5 Experimental results

Figure 3 shows the region-based coder where the TSLVQ takes place. We describe the modules of this coding scheme :

- the motion information (module E), performed directly using the original image sequence, is obtained in two steps. A first motion estimation is achieved by a coarse-to-fine multiresolution technique [12]. This primilary motion information (segmentation maps, affine-model motion parameters) is processed by a minimisation algorithm based on a minimum description length criterion in order to reduce the contour segments cost [13]. The resulting regions have polygonal shapes (see figure 7) and the average cost is about 500 bits per segmentation map with 900 control points ;
- for a given input image, the prediction image is made by module C which achieves a motion compensation of the previous decoded image using the corresponding motion information (in order to initialize the experimental coding process, the first sequence image is not coded) ;

• a dyadic wavelet transform (module T) is applied to the prediction error images (a two levels decomposition using Daubechies orthonormal compact wavelets [5]) and then wavelet coefficients are vector quantized (module TSLVQ). In fact a multiresolution codebook is designed, figure 4 displays the vector dimension and the maximal codebook entropy allowed for each subband such as the maximal multiresolution codebook entropy is about 0.27 [bit/sample or bpp] (the training sequences are obtained by the open loop coder). The non-coded source vectors, whose norm are too high, are directly duplicated into the quantized subband.

The simulations are made using the QCIF image sequence "miss america" (108 images, image size 176 by 144, 8 bpp). The computer is a SPARCStation 20 (75 Mhz).

Images extracted from the sequences are shown on figure 7. They illustrate how the TSLVQ approach is adapted to hybrid image source coding since, for a given rate, the high-density vector space region (associated with the homogeneous areas of the differential image) is coarsely quantized in order to permit a finer coding of the low-density vector space region (associated with the inhomogeneous areas).
The table displays global numerical results corresponding to the multiresolution codebook design and to the image sequence coding (note that the final multiresolution codebook entropy is about 0.16 bpp). The cpu time for an image encoding is about 0.7 s.
Figures 5 and 6 show results for images 60 up to 90. The gains formulae (in [dB]) are given by :

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\frac{1}{N}\Sigma e_i^2} + 10\log_{10}\frac{\Sigma e_i^2}{\Sigma(e_{iq}-e_i)^2} = \mathrm{Gp} + \mathrm{Gq}$$

where N is the image size, e_i the prediction errors and e_{iq} the vector quantized prediction errors.

6 Conclusion

The experimental results, obtained with this region-based coding scheme, demonstrate the TSLVQ efficiency for a very low bit rate compression of videophone image sequences.
For high vector dimension the training sequence, required for the codebook design, becomes considerable and the set of code-vectors (the tree leaves) used to quantize the images is not stable. The solution could consist in designing the TSLVQ codebook in two parts : a tree-structure "stump" using several image sequences, the branches added with respect to the sequence to be coded.

Acknowledgement
The authors wish to thank S. Pateux, V. Nzomigni and E. Nguyen for their help.

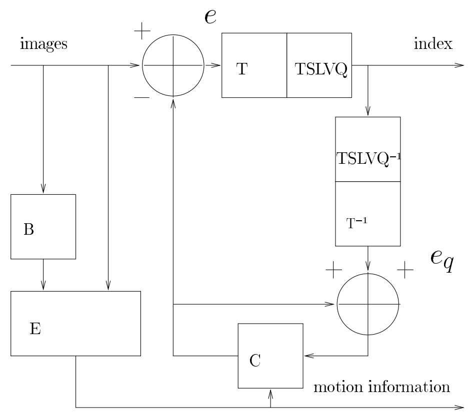

Figure 3: *Predictive coder scheme for image sequences compression (E : motion estimation, C : motion compensation, T : wavelet transform, B : buffer, e : prediction errors image, e_q : vector quantized prediction errors image).*

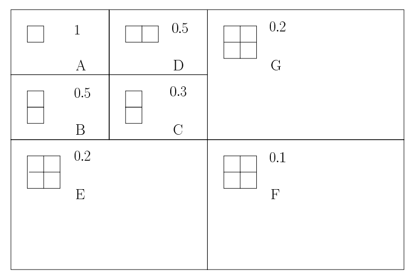

Figure 4: *Multiresolution codebook for the QCIF images sequence coding : vector shape and maximal codebook entropy [bpp] with respect to the subbands. Each subband is labelled by a letter.*

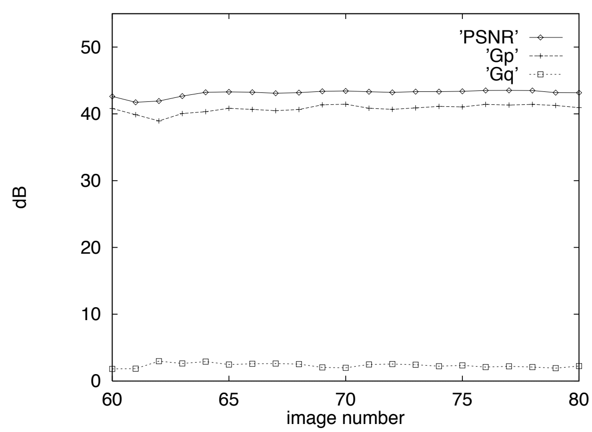

Figure 5: *Image sequence coding, gain results for the images 60 up to 80.*

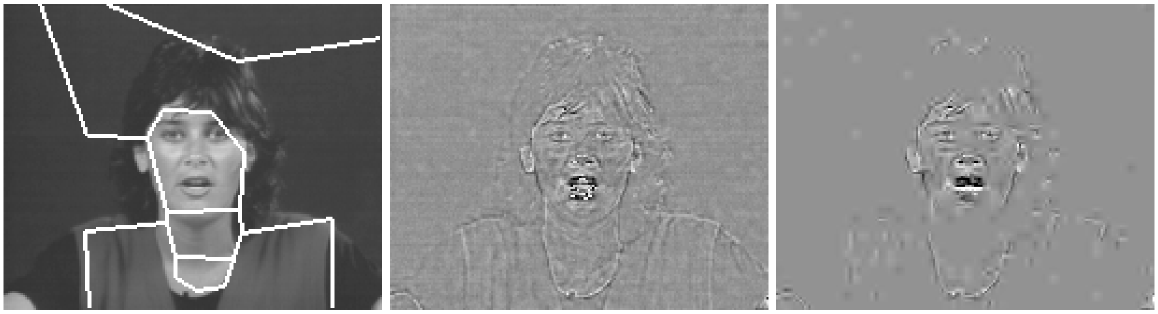

Figure 7: *Images extracted from the sequences : the original image and its segmentation map, the corresponding prediction errors image and the vector quantized prediction errors image (differential images are centered and scaled 5 times).*

	codebook design			sequence coding			
subband label	cpu time	number of code-vectors	codebook entropy	training ratio	number of used code-vectors	number of non-coded source vectors	entropy
A	0.30 s	9	0.56 bpp	352	9	32	0.60 bpp
B	0.47 s	23	0.10 bpp	206	23	28	0.19 bpp
C	0.59 s	28	0.11 bpp	311	28	31	0.22 bpp
D	0.70 s	33	0.16 bpp	264	33	43	0.30 bpp
E	16.64 s	803	0.07 bpp	140	616	0	0.10 bpp
F	2.94 s	52	0.01 bpp	639	52	1	0.02 bpp
G	21.49 s	1179	0.10 bpp	122	857	0	0.18 bpp

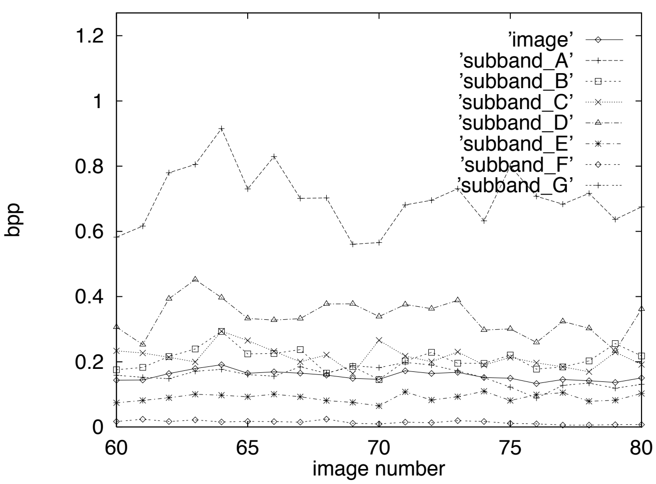

Figure 6: *Image sequence coding, entropy results for the images 60 up to 80.*

References

[1] M. Barlaud, P. Solé, T. Gaidon, M. Antonini, and P. Mathieu. – Pyramidal lattice vector quantization for multiscale image coding. – *IEEE Trans. Image Processing*, 3(4):367–381, July 1994.

[2] L. Breiman, J.H. Friedman, R.A Olshen, and C.J. Stone. – *Classification and regression Trees*. – Wadsworth, Belmont, 1984.

[3] J.H. Conway and Sloane N.J.A. – Fast quantizing and decoding algorithms for lattice quantizers and codes. – *IEEE Trans. Inform. Theory*, 28(2):227–232, March 1982.

[4] J.H. Conway and Sloane N.J.A. – *Sphere Packings, Lattices and Groups*. – Springer-Verlag, New York, 1993.

[5] I. Daubechies. – Orthonormal bases of compactly supported wavelets. – *Comm. Pure Appl. Math*, 41:909–996, 1988.

[6] T.R. Fisher. – A pyramid vector quantizer. – *IEEE Trans. Inform. Theory*, 32(4):568–583, July 1986.

[7] A. Gersho and R.M. Gray. – *Vector Quantization and Signal Compression*. – Kluwer, Boston, 1992.

[8] M. Goldberg, P.R. Boucher, and S. Schlien. – Image compression using adaptative vector quantization. – *IEEE Trans. Commun.*, 34(2):180–187, February 1986.

[9] Y. Linde, A. Buzo, and R.M. Gray. – An algorithm for vector quantizer design. – *IEEE Trans. Commun.*, 28:84–95, 1980.

[10] T.D. Lookabaugh and R.M. Gray. – High-resolution quantization theory and the vector quantizer advantage. – *IEEE Trans. Inform. Theory*, 35(5):1020–1033, September 1989.

[11] P. Monet and C. Labit. – Codebook replenishment in classified pruned tree-structured vector quantization of image sequences. – In *Proc. of ICASSP*, pages 2285–2288, 1990.

[12] V. Nzomigni. – *Compression sans pertes de séquences d'images biomédicales*. – PhD thesis, Université de Rennes I, December 1995.

[13] S. Pateux and C. Labit. – Codage efficace de carte de segmentation pour la compression orientée-régions de séquences d'images. – Technical report, INRIA, 1996.

[14] V. Ricordel and C Labit. – Vector quantization by hierarchical packing of embedded truncated lattices. – In *Proc. of SPIE VCIP*. Taiwan, China, May 1995.

[15] V. Ricordel and C Labit. – Vector quantization by packing of embedded truncated lattices. – In *Proc. of IEEE ICIP*. Washington DC, USA, October 1995.

[16] E.A. Riskin and R.M Gray. – A greedy tree growing algorithm for the design of variable rate vector quantizers. – *IEEE Trans. Signal Processing*, 39(11):2500–2507, November 1991.