# ADDRESS VECTOR QUANTIZATION WITH TOPOLOGY-PRESERVING CODEBOOK ORDERING

**Elisabetta Cammarota, Giovanni Poggi**

**Università di Napoli, Dipartimento di Ingegneria Elettronica**
**Via Claudio 21, I-80125 Napoli, Italy**

RÉSUMÉ

Dans cet article est proposée une nouvelle méthode de codage d'image, basée sur la quantification vectorielle, qui exploite la depéndance statistique entre blocs voisines. Dans une première étape il est élaboré un codebook ordonné, de sorte que les mots de codage qui ayant proche addresses dans le codebook sont proche dans le espace vectoriel aussi. En conséquence les addresses correspondant à blocs voisines exhibont une remarquable corrélation et peut étre comprimé au moyen de quelconque méthode de codage.

ABSTRACT

A new image coding scheme, based on vector quantization (VQ), is presented, which effectively exploits the inter-block dependence. It is based upon generating an ordered VQ codebook such that codewords having close addresses represent image blocks which are close in the K-dimensional input space. As a result, spatially close image blocks are represented by VQ addresses which are highly correlated, and which can be, therefore, succesfully compressed by any conventional coding technique.

## 1 Introduction

Vector quantization is a lossy compression technique which extends the concepts of scalar quantization to a higher dimensional space [1,2]. It consists of grouping blocks of input samples in vectors and representing each of them by means of a suitable template vector (codeword) chosen from a reference codebook so as to minimize a given distortion measure.

The rationale for vector quantization is that the high statistical dependency among neighboring image pels can be better exploited by quantizing them jointly rather than individually. Accordingly, as predicted by information theory, VQ performance improves with growing vector dimension.

The major drawback of VQ is its encoding complexity, which grows exponentially with the vector dimension for a given compression rate, severely limiting, hence, the block sizes one can reasonably deal with. In addition, using too large a block size would entail poor reproduction of edges and, in general, "blocky" decoded images. As a consequence, only small block-sizes are usually considered for the vector quantization of images, such as $2 \times 2$ and $4 \times 4$ pels, so only the dependence among pels very close to each other can be exploited.

In order to overcome this fundamental limitation, several sub-optimal techniques have been proposed; in particular, Finite State VQ [3] and Predictive VQ [4] try to exploit in different ways the dependency among neighboring blocks of pels to be encoded. Recently, a new coding scheme, referred to as Address Vector Quantization (AVQ), has been considered at this end [5]; it works by performing a conventional VQ on the original image and then entropy encoding blocks of VQ addresses corresponding to spatially close blocks of pels.

In this paper the idea of address coding is further developed by resorting to such a VQ algorithm, as to obtain addresses which exhibit a high degree of correlation. The algorithm is based upon rearranging the codebook so as to reduce the euclidean distance between codevectors having close addresses. This allows one to encode addresses either by resorting to simpler noiseless methods or even by adopting lossy compression techniques, without significantly affecting the image quality.

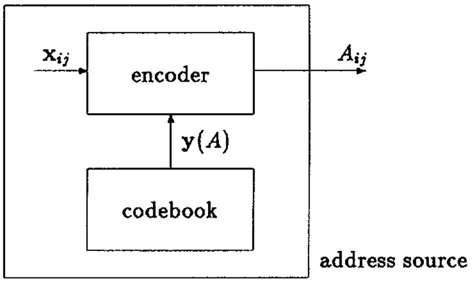
Next Section describes in more detail the idea of address coding, while Section 3 deals with the problem of codebook ordering. In Section 4 and 5 two coding schemes are proposed and their performances are analyzed. Finally, in Section 6, conclusions are drawn and future developments are outlined.

## 2 Address Coding

A vector quantizer $\mathcal{Q}$ maps each input vector $\mathbf{x} \in \mathcal{R}^K$ into one of a finite set of codevectors $\mathbf{y}(A) \in \mathcal{R}^K$, with $A \in I_N \equiv \{1, \ldots, N\}$, chosen in such a way that a given distortion measure $d(\cdot, \cdot)$ is minimized over the whole codebook; in formulas:

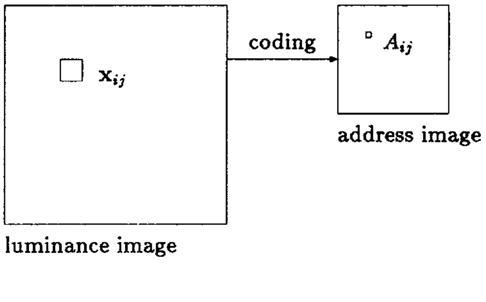
$$\mathcal{Q}(\mathbf{x}) = \mathbf{y}(A^*) \Leftrightarrow d(\mathbf{x}, \mathbf{y}(A^*)) \leq d(\mathbf{x}, \mathbf{y}(A)) \quad \forall A \in I_N \quad (1)$$

As far as the transmitting end is concerned, a vector quantizer can be viewed as a mapping of each input image block $\mathbf{x}_{ij}$, $i$ and $j$ being its spatial coordinates, into a suitable codebook address $A_{ij} \in I_N$, which is then transmitted over the channel.



The transmitter itself, then, can be viewed as a source of addresses, which is not memoryless since addresses corresponding to spatially close blocks are dependent on each other, reflecting the dependency existing among the blocks they are associated to. It is sensible, then, to exploit this dependency by applying some encoding technique to such a source.

In particular, if addresses are regarded as luminance values and suitably arranged in a bidimensional array, they can be considered to form a virtual second-layer image, to which conventional image coding techniques can be applied.



If no care is taken, however, this address-image will be highly nonstationary (see Fig.1b) and, hence, it will be difficult to exploit its statistical redundancy. In fact, even if spatially close blocks are likely to be similar, it is not said that similar codevectors are close in the codebook, nor there is, more in general, a simple relationship between the euclidean distance of two codevectors and the distance of their addresses. In [5] blocks of addresses are entropy encoded whenever they appear in a reference codebook, otherwise they are sent individually. In order for the technique to be effective, however, the reference codebook must be sufficiently large compared to the number of possible address combinations, thus calling for an heavy computational burden and a large memory requirement.

Better results can be obtained if the codebook is properly arranged so that codevectors having close indexes are close also in the vector space. This processing, while leaving unchanged the entropy of the address image and the quality of the decoded image, can considerably increase the correlation among spatially close addresses and stationarize the image, at least on a local scale (see Fig.1c). As a consequence, it is possible to effectively encode the address image by means of very simple noiseless techniques, such as DPCM, and even to resort to more efficient noisy techniques, since it is now guaranteed that small errors on addresses will entail correspondingly small errors on the reproduced blocks.

## 3 Codebook Ordering

It is not a simple task to find out such a codebook ordering procedure as to optimize subsequent encoding of address images. Actually, given a set of points in a $K$-dimensional space, with $K > 1$, it is not obvious how to link them in a chain in order to keep as small as possible the average distance between successive points. The task is even more challenging if also the distances between points being several steps apart in the chain are taken into account.

An advantageous approach is to resort to such a VQ algorithm that the codevectors are automatically ordered as the codebook is generated, and no further processing is required. This is accomplished by the Self-Organization algorithm, proposed by Kohonen in the neural network field [6], and already applied by several researchers to the image coding task [7].

In such a technique, for each input vector $\mathbf{x}$, belonging to a given training set TS , the nearest codevector $\mathbf{y}(A^*)$ is selected and displaced in order to reduce its distance from the input and, hence, the average coding distortion. In addition, codevectors whose addresses are near to the selected one, are accordingly modified, and displaced towards the input vector:

$$\mathbf{y}(A) = \mathbf{y}(A) + \alpha(t)(\mathbf{x} - \mathbf{y}(A)) \quad \forall\ |A - A^*| \leq \mathcal{N}(t) \qquad (2)$$

where $\mathcal{N}(t)$ represents the width, decreasing with time, of a neighborhood centered around the selected address. Loosely speaking, one can imagine the codevectors as linked by springs: the codevector nearest to the current input is attracted towards it, trailing the neighboring ones in the same direction. This process tends to match the codewords to the input patterns in such a way that those corresponding to close indexes represent similar blocks of pels.

The rate of adaptation to the training vectors $\alpha(t)$ is rather high at the beginning, so as the codebook can adapt to the distribution of the input vectors; then it decreases in order to reach a steady state situation:

$$\alpha(t) = \alpha_0 e^{-t/T_\alpha} \qquad (3)$$

The ordering effect, instead, is essentially controlled by the width of the neighborhood which states how dependent on each other the codevectors are:

$$\mathcal{N}(t) = \mathcal{N}_0 + \Delta_{\mathcal{N}} e^{-t/T_{\mathcal{N}}} \qquad (4)$$

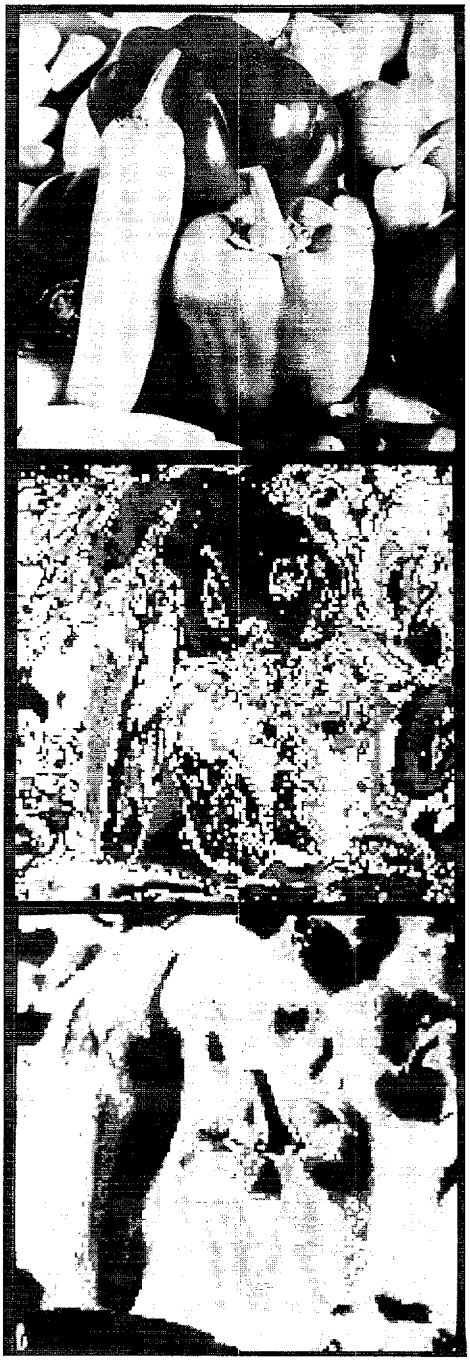


**Fig.1** Image "Peppers" (out of training-data); (a) original; (b) magnified address-image without ordering (code $C_A$); (c) magnified address-image with ordering (code $C_B$).

By choosing a slowly shrinking neighborhood, a very ordered codebook is obtained, with positive effects on the efficiency of address encoding; in this case, however, each codevector is influenced by many training vectors not belonging to its quantization cell, resulting, often, in a poor choice of the codeword. On the opposite side, a neighborhood whose width quickly drops, results in a codebook showing little or no organization. To sum up, the parameter $T_{\mathcal{N}}$ allows one to trade-off distortion for compression and deserves, thence, great attention.

In the following sections, experiments are conducted with reference to three codebooks (256 codewords, blocks of 4 × 4 pels) having a different degree of ordering (see Fig.2), depending on the parameter $S_{\mathcal{N}}$, related to $T_{\mathcal{N}}$, which represents the shrinking of the neighborhood between a presentation of the training set and the successive one. All codebooks have been realized using four gray-level images of size 512 × 512 pels, quantized at 8 bit/pel, as training data.

In Table 1, some basic characteristics of the codebooks are reported, i.e., the encoding mean square error (MSE) with respect to the training set

$$\mathrm{MSE_{TS}} = \frac{1}{N_{\mathrm{TS}}} \sum_{\mathbf{x} \in \mathrm{TS}} \frac{1}{K} \parallel \mathbf{x} - \mathcal{Q}(\mathbf{x}) \parallel^2 \qquad (5)$$

and the average distance per component between successive codevectors

$$D_1 \triangleq \left[ \frac{1}{N-1} \sum_{A=1}^{N-1} \frac{1}{K} \parallel \mathbf{y}(A) - \mathbf{y}(A+1) \parallel^2 \right]^{1/2} \qquad (6)$$

As long as just a limited ordering effect is required, as for code $\mathcal{C}_B$ having $D_1 = 12.7$, the MSE does not significantly increase with respect to the case of code $\mathcal{C}_A$ where no codeword arrangement is performed. If better ordering is required, however, like for code $\mathcal{C}_C$ where $D_1 = 3.8$, the codebook lacks of sharp edges, and a certain increase in distortion must be accounted for.

| code | $\mathcal{N}_0$ | $\Delta_{\mathcal{N}}$ | $S_{\mathcal{N}}$ | $\mathrm{MSE_{TS}}$ | $D_1$ |
|---|---|---|---|---|---|
| $\mathcal{C}_A$ | 1 | 0 | ~ | 66.4 | 22.7 |
| $\mathcal{C}_B$ | 1 | 64 | 0.925 | 70.8 | 12.7 |
| $\mathcal{C}_C$ | 1 | 64 | 0.950 | 99.8 | 3.8 |

**Tab.1**

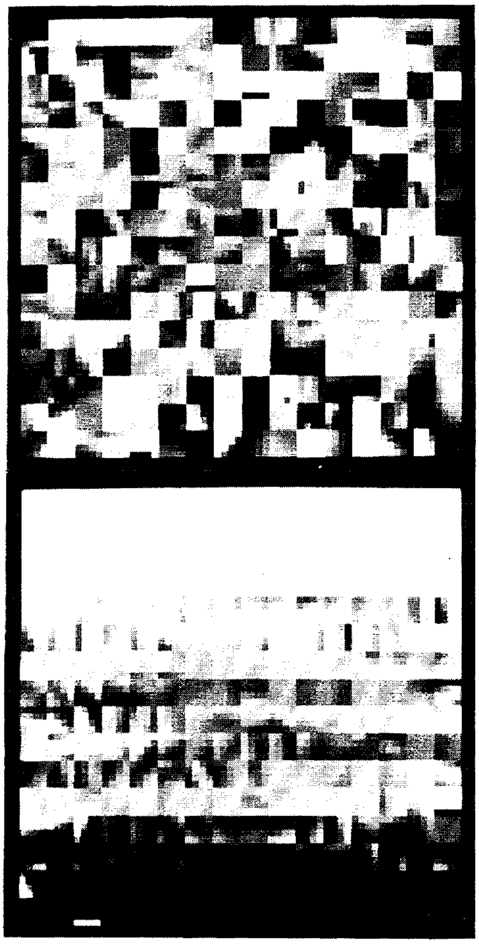


**Fig.2** Two codebooks adopted for vector quantization: (a) code $\mathcal{C}_A$, non ordered and (b) code $\mathcal{C}_B$, ordered; blocks are magnified and displayed on successive rows in order of increasing index.

## 4 Address Prediction

Using an ordered codebook, vector quantization yields address images which exhibit a high degree of correlation among neighboring pels and appear like a rough copy of the original images (see Fig.1c); thus, they can be encoded by means of a simple adaptive DPCM technique [8], reducing the bit-rate without affecting the image quality.

The current address, say $A_{i,j}$, is estimated as the previous address lying on either the same row $A_{i,j-1}$ or the same column $A_{i-1,j}$. The choice is made by considering the previous column and row prediction errors, $E_c = \mid A_{i,j-1} - A_{i-1,j-1} \mid$ and $E_r = \mid A_{i-1,j} - A_{i-1,j-1} \mid$, and performing the prediction along the direction which exhibits the smaller error. This simple adaptation rule reduces the problems arisen by straight edges, generally improving the performance.

The training images have been vector quantized, and the resulting addresses have been considered to get the DPCM statistics for generating an appropriate Huffman code. Fig.3 plots the probability of occurrence of the various prediction errors for two different codebooks: it can be seen, as expected, that when an ordered codebook ($\mathcal{C}_B$) is used, prediction errors are much more concentrated around zero;

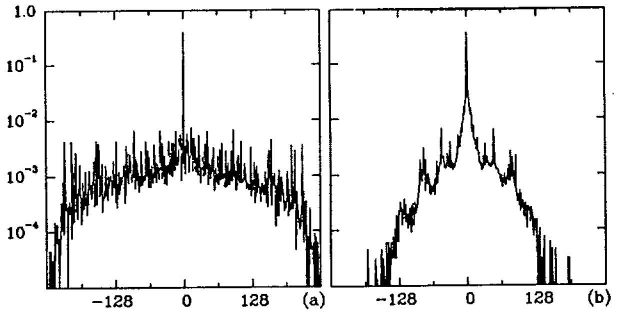


**Fig.3** Probability of occurrence in the training address-images of the various prediction errors; (a) codebook $\mathcal{C}_A$: entropy = 5.93 bit/error; (b) codebook $\mathcal{C}_B$: entropy = 5.10 bit/error.

Let us note that, even when a standard VQ is adopted, it is worthwhile using some address prediction technique, at least for small codebooks. In such a case, in fact, only a few codewords are available, and it often happens that adjacent blocks are quantized by the same codeword, as testified by the large probability of a zero prediction error. At higher bit-rates, however, this is no longer true, and the ordering significantly improves the performance [9].

In Table 2, coding results, relative to two out-of-training images, are summarized in terms of bit-rate in bit/pel and peak signal-to-noise ratio PSNR $\triangleq 256^2/\mathrm{MSE}$ in dB when codebooks having different degree of ordering are employed.

|  | Einstein |  | Peppers |  |
|---|---|---|---|---|
| code | bit-rate | PSNR | bit-rate | PSNR |
| $\mathcal{C}_A$ | 0.41 | 29.56 | 0.38 | 30.71 |
| $\mathcal{C}_B$ | 0.37 | 29.58 | 0.33 | 30.67 |
| $\mathcal{C}_C$ | 0.34 | 28.94 | 0.30 | 29.36 |

**Tab.2**

Results show that address prediction allows some bit saving in all cases, even when no codebook rearrangement is performed; using the fairly ordered codebook $\mathcal{C}_B$, however, the compression performance increase without loss in image quality (Peppers is encoded at 0.33 bit/pel with PSNR=30.67 dB). Further compression can be obtained by resorting to code $\mathcal{C}_C$, which is more ordered, at the price of some additional distortion (on the average a loss of about 1 dB in PSNR).

## 5 Address Vector Quantization

One major advantage obtained by resorting to an ordered VQ codebook is the opportunity of encoding addresses by means of a lossy technique and, hence, reach rather large compression factors. In fact, codevectors having close addresses are close, thanks to ordering, also in the input space $\mathcal{R}^K$, so that small errors on addresses correspond usually, even if not always, to small errors on the reproduced blocks. Much attention must be payed, however, to the distortion introduced in this step, since even a small number of large errors on addresses gives rise to a quite annoying block-effect. It is then necessary to provide an additional check, in order to single out these situations and make up for them by transmitting further information.

In this work, experiments have been conducted using VQ also in the second step; addresses have been grouped in blocks of size 2 × 2 and vector quantized using a 14-bit mean/shape codebook. This rather large codebook is necessary in order to keep small the number of large errors on addresses; many shapes, however, are only occasionally utilized, so the actual entropy of the codebook is about 6 bits. For each block, the increase in distortion

$$\Delta_{ij} \triangleq \frac{1}{K}\left(\| \mathbf{x}_{ij} - \mathbf{y}(\mathcal{Q}_A(A_{ij})) \|^2 - \| \mathbf{x}_{ij} - \mathbf{y}(A_{ij})) \|^2\right) \qquad (7)$$

due to the address quantization $\mathcal{Q}_A$, has, then, been evaluated, and, whenever exceeding a given threshold $\Delta_{\text{MAX}}$, further information has been transmitted.

In Tables 3 and 4, coding results are reported for the two ordered codebooks considered, as function of the threshold $\Delta_{\text{MAX}}$; for comparison purpose the numbers obtained when no address quantization is performed are also shown;

| | | Einstein | | Peppers | |
|---|---|---|---|---|---|
| $\mathcal{Q}_A$ | $\Delta_{\text{MAX}}$ | bit-rate | PSNR | bit-rate | PSNR |
| yes | 25 | 0.46 | 29.44 | 0.47 | 30.55 |
| yes | 50 | 0.41 | 29.19 | 0.43 | 30.21 |
| yes | 100 | 0.36 | 28.68 | 0.38 | 29.51 |
| yes | 200 | 0.31 | 27.89 | 0.33 | 28.48 |
| no | ~ | 0.50 | 29.58 | 0.50 | 30.67 |

**Tab.3:** code $\mathcal{C}_B$

| | | Einstein | | Peppers | |
|---|---|---|---|---|---|
| $\mathcal{Q}_A$ | $\Delta_{\text{MAX}}$ | bit-rate | PSNR | bit-rate | PSNR |
| yes | 25 | 0.37 | 28.74 | 0.40 | 29.18 |
| yes | 50 | 0.33 | 28.57 | 0.35 | 28.97 |
| yes | 100 | 0.30 | 28.33 | 0.31 | 28.61 |
| yes | 200 | 0.27 | 28.00 | 0.29 | 28.20 |
| no | ~ | 0.50 | 28.94 | 0.50 | 29.36 |

**Tab.4:** code $\mathcal{C}_C$

Numerical results confirm the compression ability of such a coding scheme: rather low bit-rates can be achieved by raising the distortion threshold and allowing so largest error on addresses to be tolerated and more blocks of addresses to be vector quantized; of course, as the threshold level increases, the corresponding PSNR decreases, however good compromises can be found. It is worth saying that images encoded by means of a very ordered codebook appear rather smoothed, whatever the distortion threshold, due to the lack of sharp edges; this smoothing effect is rather annoying under a subjective point of view, so this choice is suitable only when very low bit-rates are requested and, hence, some additional distortion is acceptable. In any case, interesting bit-rates can be achieved (~0.3 bit/pel) with a loss of just 2 dB in PSNR with respect to conventional VQ.

## 6 Conclusions

In this paper a new image coding scheme has been presented which exploits the inter-block correlation by jointly encoding VQ addresses corresponding to spatially close blocks. The potentialities of such an approach are fully exploited by rearranging the codebook so that codevectors having close indexes are close also in the vector space; this processing, in fact, greatly increases the linear dependence (correlation) among addresses corresponding to spatially close blocks, allowing one to resort to simpler techniques for encoding addresses. In addition, also lossy techniques can be adopted, since it is guaranteed that small errors on addresses entail small errors on the corresponding codevectors; this leads to very low bit-rates with limited increase in distortion.

Numerical experiments, based on Kohonen algorithm, have confirmed the potentialities of such an approach: much has been gained, under the complexity respect, utilizing adaptive DPCM for address coding, whereas very low bit-rates have been reached, when mean/shape VQ has been considered for this purpose. Experiments have also pointed out that a high degree of ordering entails, as a drawback, lack of sharp edges in the codebook and, hence, some smoothing effect on reproduced images.

On the basis of these considerations, future investigation will concern two main points: i) the search for a better codebook ordering algorithm which optimize subsequent encoding steps without increasing the distortion; ii) the assessment of the performances of various techniques, especially lossy ones, for the address coding task.

## References

[1] N.M.Nasrabadi, R.A.King, "Image coding using vector quantization: a review", IEEE Trans. Commun., vol COM-36, pp.957-971, Aug. 1988.

[2] A.Gersho, R.M.Gray, "Vector quantization and signal compression", Kluwer Academic Press, Massachusetts, 1990.

[3] J.Foster, R.M.Gray, M.O.Dunham, "Finite-state vector quantization for waveform coding", IEEE Trans. Inform. Theory, vol.IT-31, pp.348-359, May 1985.

[4] V.Cuperman, A.Gersho, "Vector predictive coding of speech at 16 Kb/s", IEEE Trans. Commun., vol COM-33, pp.685-696, July 1985.

[5] N.M.Nasrabadi, Y.Feng, "A multilayer address vector quantization technique", IEEE Trans. Circuits Syst., vol.37, pp.912-921, July 1990.

[6] T.Kohonen, "Self-organization and associative memory", 2nd Ed. Springer-Verlag, New York, 1988.

[7] N.M.Nasrabadi, Y.Feng, "Vector quantization of images using Kohonen self-organization", Proc. IEEE Conf. on Neural Networks, vol.I, pp.101-108, July 1988.

[8] R.E.Graham, "Predictive quantizing of television signals" IRE Wescon Convention Record, Vol.2, pt.4, pp.147-157, 1958.

[9] E.Cammarota, G.Poggi, "Address predictive vector quantization of images", Proc. AEI symposium on Image Processing: Applications and Trends, Genova (I), June 1991.