A weighted distortion measure for vector quantization in transform domain

Iole Moccagatta and **Murat Kunt**

Signal Processing Laboratory
Swiss Federal Institute of Technology
CH-1015 Lausanne Switzerland
email: iole@ltssg1.epfl.ch

Abstract

In this paper wavelet transform and vector quantization (VQ) are combined for color image compression. A multiresolution and frequency oriented codebook structure is used to code the wavelet coefficients. Furthermore, a vector forming is proposed to exploit intra and inter bands correlations. To take into account the different importance of the quantization noise in subband coefficients, a weighted distortion measure is adopted in the codebook design. Therefore the generated code vectors are forced to reflect the relative significance of the transform coefficients, resulting in an improved performance of the coding system.

1 Introduction

The goal of image coding is to reduce data rate for transmission or storage, while maintaining an acceptable image quality. Among the different compression strategies, the use of subband/wavelet coding is well known, both in term of energy compaction and perceptual tolerability to error coding. Data compression is then achieved by transmitting a properly quantized subset of the coefficients previously generated. In this paper, the problems of exploiting the inter/intra bands correlations and appropriately considering quantization noise have been addressed.

Spatial correlation is reduced by a perceptually derived wavelet transform, described in section 2. In section 3 a multiresolution vector quantizer (VQ) with pyramidal vector tailoring is then proposed to reduce linear and non linear correlations of the transform domain. To properly quantize the coefficients, an input weighted distortion measure is introduced to generate the codebooks, as described in section 4. The complete coding scheme is presented in section 5, together with simulation results. Finally, conclusions are given in section 6.

Résumé

Dans cet article, une technique de compression des images couleurs, combinant la transformée en ondelette et la quantification vectorielle (VQ) est présentée. Les coefficients de la transformée en ondelette sont codés en utilisant un dictionnaire basé sur la fréquence et possédant une structure multirésolution. De plus, les corrélations intra et inter bandes sont exploitées grâce à la vectorisation. Une mesure de distorsion pondérée pour chaque sous-bande est utilisée dans le dictionnaire de manière à prendre en compte l'importance variable des bruits de quantification. Ainsi, les mots-codes obtenus reflètent l'importance relative des coefficients transformés, ce qui permet d'améliorer les performances du codeur.

2 Perceptually derived wavelet transform

A tree structure and rectangular separable biorthogonal Gabor-like wavelet transform [1] has been chosen to perform octave band partitioning, which is motivated by typical image statistics and by the spatial frequency sensitivity of the human visual system. Furthermore, the resulting multiresolution structure is suitable for a generic coding scheme by discarding or taking into account a certain number of levels in the multiresolution data structure. The use of the Gabor-like wavelet transform is motivated by the optimal joint localization in the spatial and spatial-frequency domain of the Gabor functions, together with their perceptual characteristics [2]. The pyramidal decomposition results in a series of subbands corresponding to different spatial frequency resolutions and orientations.

3 Multiresolution VQ with pyramidal vector forming

Data compression is achieved by transmitting only a subset of the suitably quantized coefficients generated by the wavelet transform. In [3], Berger shows that, under the condition of laplacian source, the optimal scalar quantizer, in term of mean square error for an ideal entropy coding of coefficients, is the uniform quantizer. However, this distortion measure is not relevant from a perceptual point of view, due to the different relevance of the quantization error in the subbands. Furthermore, an independent scalar quantization scheme does not exploit the residual statistical correlations inside and between each subband (intra/inter band correlations) [4]. In particular, correlations between subbands which belong to the same frequency orientation (horizontal, diagonal and vertical) still remain. These considerations motivate the replacement of scalar quantizer by a multiresolution vector quantizer with pyramidal vector forming (PVQ) [5].

The lowest resolution level in the pyramidal data (DC component) is the most important one and leads to the highest distortion if badly transmitted, therefore a PCM technique is applied to transmit this part of the data. The remaining levels in the multiresolution data are encoded using a VQ with a vector forming capable to exploit the cross-level correlations inside the pyramid.

The luminance coefficients of the same frequency orientation (horizontal, diagonal and vertical) are combined together with the chrominance ones, assuming all components (Y,U,V) are sampled at the same spatial frequency. Three different vectors are built, according to the different frequency orientations, each vector containing coefficients from low and high frequency subbands. Coefficients are taken from square blocks at the same spatial location in the pyramid, the size of the block being defined by its position inside the pyramid. In Fig. 1 vector forming applied to the subbands characterized by vertical spatial frequency orientation is described, the same strategy being applied to the horizontal and vertical ones.

Thanks to the energy compaction property of the transform coding [6], the vector size can be reduced without loosing performance. Meanwhile, larger block size can be applied in the transform domain, avoiding the annoying blocking artifact due to the independent quantization of each block in the image domain.

4 Weighted distortion measure

The exploitation of the intra/inter band correlations is not the only criterion that needs to be taken into account in the design of a quantizer applied to the transform domain. The structure described in section 3 combines into a unique vector luminance and chrominance components from different subbands, thus taking advantage of the correlation between the coefficients to achieve a more efficient coding. Nevertheless,

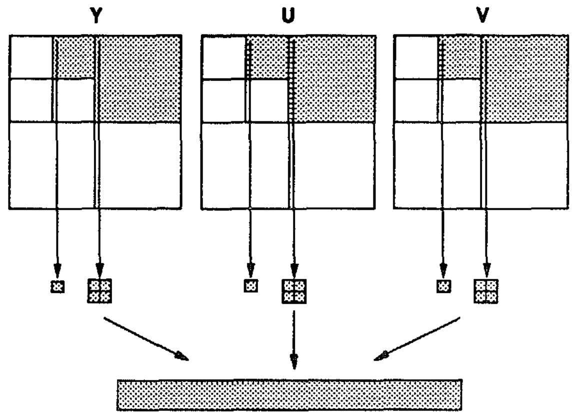

Figure 1: Pyramidal vector forming.

the coefficients combined into the same vector have different importance in terms of the visual quality of the decoded image. In other words, the information carried on by the vector is not equally distributed between its components. As a consequence, the distortion introduced by the quantization stage has to be unequally assigned to the vector components. In this way, higher distortion will be assigned to those vector components having less influence on the quality of the compressed image. In order to apply this concept to the coding scheme previously described, a weighted distortion measure has been introduced in the vector quantizer design.

Given $\mathbf{x}$ a k-dimensional source vector

$$\mathbf{x} = (x_1, x_2, \ldots, x_k)^T \quad \text{with } \mathbf{x} \in \mathbf{R}^k, \tag{1}$$

a vector quantizer $\mathbf{Q}$ of dimension k and size N is a function that maps $\mathbf{x}$ into one of the N output points, $\mathbf{y}_i$, contained in the finite set $\mathcal{C}$, where

$$\mathcal{C} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N) \quad \text{and} \quad \mathbf{y}_i \in \mathbf{R}^k, \\ \forall i \in \mathcal{I} = \{1, 2, \ldots N\}. \tag{2}$$

The set $\mathcal{C}$ is called the *codebook* and the output or reproduction points $\mathbf{y}_i$ are called *codewords*. The resolution, or rate, of the vector quantizer is $r = (log_2 N)/k$. Associated with the codebook there is a partition of $\mathbf{R}^k$ into N regions or cells, $\mathbf{R}_i$, defined as follows:

$$\mathbf{R}_i = \{\mathbf{x} \in \mathbf{R}^k : \mathbf{Q}(\mathbf{x}) = \mathbf{y}_i\}, i \in \mathcal{I} \tag{3}$$

In this way we associate a vector $\mathbf{y}_i$ with each cell $\mathbf{R}_i$ and the quantizer assigns the code vector $\mathbf{y}_i$ to $\mathbf{x}_i$ if $\mathbf{x}_i \in \mathbf{R}_i$.

When $\mathbf{x}$ is quantized as $\mathbf{y}_i$, a quantization error occurs, and a distortion measure $d(\mathbf{x}, \mathbf{y}_i)$ can be defined in order to evaluate the performance of the quantizer. In the proposed approach, we have replaced the classical quadratic distortion measure by means of a weighted squared error of the form

$$d_w(\mathbf{x}, \mathbf{y}_i) = (\mathbf{x} - \mathbf{y}_i)^t \mathbf{W}_\mathbf{x} (\mathbf{x} - \mathbf{y}_i) \quad (4)$$

where $\mathbf{W}_x$ is a symmetric and positive defined weighting matrix that depends on the input $\mathbf{x}$. This type of distortion measure has been discussed by [7], while applications to the image domain have been p-resented in [8]. It is worthwhile to notice that in case $\mathbf{W}_x$ is a diagonal matrix with diagonal values $w_{jj} > 0$, then from Eq.(4) we obtain

$$d_w(\mathbf{x}, \mathbf{y}_i) = \sum_{j=1}^{k} w_{jj}(x_j - y_{ji})^2 \quad (5)$$

where the influence of the weighting elements w_{jj} on the different vector components is explicit. Given the distortion as in Eq.(4), the centroid of the the partition $\mathbf{R}_i$ results to be:

$$cent(\mathbf{R}_i) = \left\{ \sum_{\mathbf{x} \in \mathbf{R}_i} \mathbf{W}_\mathbf{x} \right\}^{-1} \left\{ \sum_{\mathbf{x} \in \mathbf{R}_i} \mathbf{W}_\mathbf{x} \mathbf{x} \right\} \quad (6)$$

as derived in [7].

The advantage of weighted centroids appears when we considered the characteristics of the data involved in this coding scheme and the strategy adopted to built the codebooks. A sample distribution, defined by the training sequence, could be used in the codebook design, as is done in the LBG algorithm [9]. The resulting codebook will minimize the given distortion by taking advantage of the shape of the training set's statistical distribution. However, the population of the training set reflects the statistical characteristics of the encoding data and not their visual importance. In other words, the centroid position will be decided without taking into account the necessity of a better coding of some vector components with respect to the others. On the opposite, the introduction of a weighting matrix will shift the codeword position to better represent the most important vector components.

5 Coding strategy and simulation results

This section describes the complete coding scheme combining perceptual wavelet decomposition and VQ for color image compression. Firstly spatial correlation is reduced by the Gabor-like wavelet transform. Afterwards, two strategies previously proposed are applied. On one side a multiresolution and frequency oriented VQ exploits the residual inter/intra band correlations. On the other, a weighted distortion measure in the codebook design allows to properly control the quantization errors in the different subbands.

Considering the vector forming strategy described in section 3, weights have been assigned to allow higher distortion on the chrominance components with respect to the luminance one. Furthermore, the luminance coefficients from the lower subbands have been weighted more compared to the corresponding coefficients in the higher frequency subbands. Various technique have been proposed in literature to evaluate weighting coefficients [10], while these weighting factors have been determined in an empirical way.

After the wavelet transform and before the quantization, a dynamic threshold have been applied on the coefficients. The goal is to discard those coefficients whose value is near to zero. As a consequence, the statistic of the data is changed without introducing perceptual distortion. This turn out in allowing more codewords to be assigned to less-frequent but more important vectors.

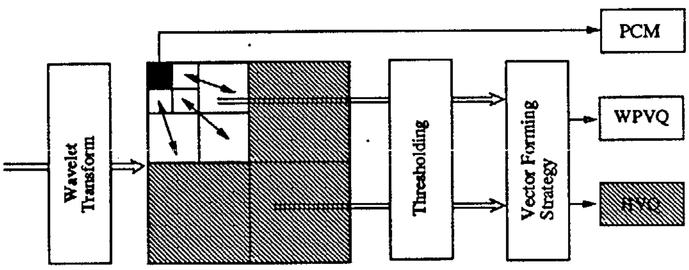

Figure 2: Coding scheme.

The complete compression scheme is described in Fig. 2, the same structure being adopted for generating the training set as well as for coding of new data. Simulations have been performed on several CIF format images (size 288×360, Y, U and V components, 24 bpp) as follows. A three levels pyramidal decomposition has been applied and the first two levels have been combined by means of the pyramidal vector forming, taking one pixel from the first layer and a block of 2×2 from the second one. The resulting 15 elements vectors are then coded by the weighted distortion based codebooks. In the third level the U and V components have been discarded, due to their low visual importance, while blocks of 4 × 4 pixel have been built with the Y coefficients. This scheme results in six different codebooks, where three are assigned to the first and second layer (WPVQ), and the last three to the third one (HVQ). For each level, a different codebook is applied to each frequency orientation.

A training database of six images has been used to built the codebooks, each image representing 1620 training vectors for each of the six codebooks. Table 1 and 2 compare the performance of the coding scheme with (WPVQ) and without (PVQ) weighting distortion measure in the pyramidal vector quantizer design and coding. In case no weighting distortion is used, $\mathbf{W}_x$ is equal to the identity matrix. The tables are organized as follows: in the first and second columns codebook sizes are given for the specified quantizers, followed by quality evaluation of the coded image and achieved compression ratio (CR). A threshold keeping 90% of the coefficients has been adopted in combination with the weighting distortion measure.

The performance has been evaluated in terms of Peak SNR for the luminance (Y) and chrominance (U,V) components of the test image "Lena". Codebook sizes of 128 and 256, for the WPVQ/PVQ and HVQ, allow to code the color image with 0.97 bpp and 1.04 bpp respectively. In Fig. 3 the luminance component of the test image coded with WPVQ is compared to the original one. It can be noticed that the use of an appropriate weighting distortion measure allows to reduce the distortion of the luminance component while reducing the quality of the chrominance ones. Nevertheless, this higher distortion in the color component is not noticeable in the decoded image.

WPVQ	HVQ	Y[dB]	U[dB]	V[dB]	CR
128	128	28.30	33.12	31.36	24.75
256	128	28.86	32.70	31.22	22.96

Table 1: Simulation results: test image "Lena" coded with weighted distortion measure.

PVQ	HVQ	Y[dB]	U[dB]	V[dB]	CR
128	128	27.73	33.69	32.13	24.75
256	128	28.24	33.81	32.21	22.96

Table 2: Simulation results: test image "Lena" coded with weighting matrix $\mathbf{W}_x = I$.

6 Conclusion

In this paper a coding scheme for color still images has been presented. Spatial redundancy is reduced by means of a perceptually derived wavelet transform. To improve the performance of the compression scheme, a VQ with a multiresolution and frequency oriented codebook structure is proposed, allowing to exploit the residual correlations between and across the subbands. Furthermore, a weighted distortion measure is introduced in the codebook design to appropriately weight quantization noise in luminance and chrominance coefficients combined into the same vector. Performance comparison has shown an improvement in term of both objective and subjective quality for a given compression ratio when the weighted distortion is applied.

Figure 3: Original (right) and compressed (left) luminance component of the color test image "Lena" coded at 1.04 bpp by means of WPVQ.

References

[1] T. Ebrahimi. *Perception Based Localized Linear Operators -Application to Sequence Compression-*. PhD thesis, Swiss Federal Institute of Technology, Lausanne, 1992.

[2] J. G. Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169–1179, July 1988.

[3] T. Berger. Optimum quantizers and permutation codes. *IEEE Trans. on Information Theory*, 18(6), November 1972.

[4] G. Galand, E. Lacson, G. Furland, and J. Menez. Subband Coding of Images using Adaptive VQ and Entropy Coding. In *Image 'Com 90*, pages 106–110, November 19-21 1990.

[5] I. Moccagatta and M. Kunt. A Pyramidal Vector Quantization Approach to Transform Domain. In *European Signal Processing Conference*, Brussels, Belgium, 25-28 August 1992.

[6] J. Katto and Y. Yasuda. Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients. In *Visual Comm. and Image Processing '91*, pages 95–106, 1991.

[7] R. M. Gray and E. Karnin. Multiple Local Optimal in Vector Quantizer. *IEEE Info. Theory*, IT-28(2):256–261, March 1982.

[8] P. C. Cosman, K. L. Oehler, A. A. Heaton, and Gray. R. M. Three-structured Vector Quantization with Input-Weighted Distortion Measures. In *Visual Comm. and Image Processing '91*, 1991.

[9] Y. Linde, A. Buzo, and R.M. Gray. An Algorithm for Vector Quantizer Design. *IEEE Trans. on Comm.*, 1:84–95, January 1980.

[10] B. Macq. *Perceptual Transforms and Universal Entropy Coding for an Integrated Approach to Picture Coding*. PhD thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1989.