

# AN IMAGE CODING SCHEME BASED ON PERCEPTUALLY CLASSIFIED VQ FOR HIGH COMPRESSION RATIOS

*Iole Moccagatta and Murat Kunt*

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

## ABSTRACT

In this paper we present a coding scheme for color images aiming at high compression ratios. It is based on perceptually classified Vector Quantization (VQ), where the different classes are chosen to achieve a better quality of the decoded image. Spatial correlation is reduced by a tree structured wavelet decomposition, then prediction of the insignificant coefficients is performed across subbands. Afterwards, the reduced set of data is organized in vectors in such a way that residual intra and inter band correlation are exploited. Finally, such vectors are coded by a classified VQ. Reconstructed images with good quality and a compression ratio higher than 100:1 have been produced by the proposed scheme, which has been shown to outperform the JPEG standard.

## 1. INTRODUCTION

The goal of image coding is to reduce data rate for transmission or storage, while maintaining an acceptable image quality. The use of subband/wavelet decomposition is well known for compressing still images and video, thanks to its inherent energy compaction property and to its perceptual tolerability to error coding. After performing spatial decorrelation of the visual data by means of the above cited approach, both lossless and lossy compression can be performed. The latter is achieved by applying a proper quantization on a subset of the coefficients generated by the spatial decomposition.

Even though a reduction of the spatial correlation is performed by subband/wavelet decomposition, the coefficients generated by such filtering are non independent. For example, similarities are evident among the values of the coefficients located in corresponding spatial/frequency locations of the different subbands. In the following, we will refer to such similarities as inter bands correlation.

Recent researches have been developed to exploit the inter band similarities within the framework of pyramidal decomposition. In [1, 2] the information of the lower frequency subbands is used to predict the upper ones of the same spatial frequency orientation. Other approaches [3, 4] define the significance or insignificance of the higher subband coefficients using the low pass subband information. Nevertheless, none of these schemes takes into account the non stationary characteristics of the visual data to perform an appropriate coding of the perceptually relevant information.

In this paper, we present a coding scheme for color images based on perceptually classified Vector Quantizer (VQ) of the wavelet coefficients, where the different classes are chosen to achieve a better quality of the decoded image.

Some of the coefficients generated by the wavelet decomposition are discarded prior to quantization by predicting their insignificance across subbands. Then a multiresolution VQ with Pyramidal vector tailoring and Weighted distortion measure (WPVQ) is proposed to reduce linear and non linear correlation of the transform domain [5]. Furthermore, the vectors built to perform the VQ are classified according to the visual information they carry.

The paper is organized as follows. In Sec. 2, the applied wavelet transform is briefly described and different prediction schemes to detect the insignificant coefficients are presented. More details concerning the WPVQ are given in Sec. 3. In Sec. 4, the concept of perceptually classified WPVQ is presented, together with different classification criteria. Finally, in Sec. 5 simulation results of a classified WPVQ applied to color image coding are presented, together with the final conclusions.

## 2. WAVELET TRANSFORM AND CROSS-SUBBAND PREDICTION

The tree structure and rectangular separable biorthogonal Gabor-like wavelet transform [6] is chosen to perform octave band partitioning. The pyramidal decomposition results in a series of subbands corresponding to different spatial frequency resolutions and orientations.

The decision strategy concerning which coefficients have to be discarded prior to quantization is similar to the one proposed by [4]. The coefficients are organized in tree structures built according to the pyramidal subband decomposition. That is, each coefficient in the upper level is the root of a tree, whose leaves correspond to a block of coefficients of the same spatial location and frequency orientation in the next (i.e. lower) subband. The tree grows level by level, as each leave becomes itself the root of a block of coefficients in the following subband.

The so built trees are then classified in significant or insignificant ones, the latter named zero-trees according to [4]. The luminance and chrominance coefficients are treated in the same way, the only difference being that the lower subbands of the chrominance components are *a priori* discarded, due to their visual irrelevance.

To predict insignificance across subbands, two strategies are proposed [7]. In the first case, referred as  $\mathcal{A}$ , the prediction is performed on the last (i.e. lower) level of the pyramid, knowing the previous ones. If we suppose a three level decomposition, a coefficient is defined as the root of a zero-tree if and only if itself and all its sons, except the ones of the last level, are insignificant with respect to the same threshold. In the second case, referred as  $\mathcal{B}$ , a full prediction of the insignificant coefficients across the subbands is acquired. Therefore if the root of the tree is considered insignificant with respect to a given threshold, all the luminance coefficients contained in the same tree are considered insignificant too and the tree is classified as a zero-tree. The chrominance coefficients follow the classification of the corresponding luminance ones.

The choice of predicting the insignificance rather than the significance is based on the observation that the first one is easier compared to the second one. Furthermore, additional prediction of the coefficients across the subbands does not allow an important reduction of the bit rate [1]. On the opposite, by using the zero-trees less relevant coefficients can be discarded, therefore more bits are available for the coding of the few important data.

## 3. MULTIREOLUTION WPVQ

The set of coefficients reduced by the above described technique is then organized in vectors according to the Pyramidal VQ (PVQ) scheme proposed in [5]. The luminance coefficients of the same frequency orientation are combined together with the chrominance ones, assuming all components (Y,U,V) are sampled at the same spatial frequency. 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. These vectors are therefore coherent with the zero-trees described in Sec. 2

Furthermore, the classical quadratic distortion used to construct the codebook is replaced by the weighted distortion measure [8], thus resulting in the Weighted PVQ (WPVQ) [5]. This way we take into account the perceptual influence of the quantization noise in the subbands.

## 4. PERCEPTUALLY CLASSIFIED WPVQ

The vectors built according to the WPVQ scheme described in the previous section contain information related to a specific part of the image. Due to the high non-stationary characteristics of natural images, the use of a single codebook to quantize all the vectors will result in sub-optimal performance. To justify such statement, we have to refer to the principles of the clustering algorithm, i.e the LBG [9], used to derive the codebook. The coding performance of a VQ is related to the statistical characteristics of the training set used to derive the codebook, as well as to the adopted distortion measure.

If the training set is populated with a small fraction of a specific class of data, such class will be poorly represented by the elements of the codebook, i.e. by the codevectors. Furthermore, the distortion measure does not always assure that a data belonging to a specific class will be coded with a codevector representative of the same class. A typical consequence of such behavior is the poor VQ coding of edges in the image domain when the MSE distortion measure is adopted, as pointed out in [10]. In order to preserve the features associated to each class, vector belonging to one class have to be coded with a codevector belonging to the same class. This solution, known as Classified VQ (CVQ) [10], has proved its efficiency in preserving edges when applied to the image domain.

The goal of this paper is to apply the CVQ to the vectors built in the spatial/frequency domain according to the previously described strategy. A crucial choice of the CVQ scheme is related to the classification cri-

terion, as it has to satisfy two requirements. First, it has to assure a separation of the data in classes which have a different influence on the quality of the decoded image. Second, the side information which addresses the codebook chosen for the coding/decoding of each vector has to be keep very low.

Different classification schemes will be proposed in the following. The first one is strictly related to the pyramidal decomposition described in Sec. 2. It consists in separating the vectors in three classes, each class containing vectors whose coefficients belong to subbands with the same frequency orientation. In this way we assure that a vector of a specific frequency orientation class is coded by a codevector representing the same class. Furthermore, a wise scanning technique of the decomposed image can avoid the use of side information. However, this scheme does not separate classes which are not equally represented in the training set. To achieve this goal, alternative criteria can be defined.

In [2], the value of the variance on the current subband is extracted from the upper band information and used to address a finite state scalar quantizer. To avoid the side information generated by this scheme, we propose to classify the pixels of the first level of the pyramid (DC) in relevant classes with respect to the human visual system. For example, an edge detector could separate the edge points from the non edge ones. Furthermore, a segmentation can be performed to separate edge, texture and homogeneous regions [11]. The classification of the DC coefficients is then down-projected into the lower subbands, according to the relation root-leaves described in Sec. 2. Furthermore, no classification information needs to be sent, as the classification criterion could operate at both the receiver and transmitter-end of the coding scheme.

Codebooks corresponding to the various classes are then designed. This is done by first classifying the training set into subsets, each containing vectors belonging to one class. Then, we design each sub-codebook separately using the LBG algorithm. This approach is then combined with the previously described WPVQ, thus resulting in a classified WPVQ. Such a scheme allows a better coding of the visually important details of the image, despite their limited statistical relevance.

## 5. SIMULATION RESULTS AND FINAL CONCLUSIONS

In the following, preliminary results of the proposed coding scheme are presented, where the classification of the vectors has been performed according to the frequency orientation of the subbands. The complete coding scheme is described in Fig. 1.

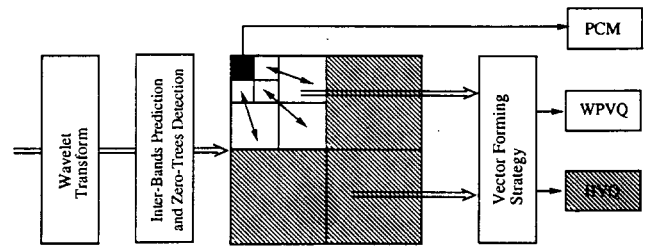


Figure 1: Block diagram of the coding scheme.

The first two levels of the decomposition are treated by the WPVQ described earlier. In the third level the U and V components are discarded, due to their low visual importance, while the Y coefficients is encoded by the classical VQ, referred as HVQ in the following. A different codebook is applied to each frequency orientation, thus resulting in six different codebooks, where three are assigned to the WPVQ, and the last three to the HVQ.

The results presented here are obtained by coding the test image "Lena" (size  $288 \times 360$ , Y, U and V components, 24 bits/pixel), which is outside the training set.

Table 1 shows the performance of the proposed coding scheme for very low bit rate. The high compression ratio is achieved by choosing a high threshold to detect the roots of the zero-trees. The results are given in terms of Peak Signal Noise Ratio (PSNR) versus rate (bits/pixel) of the compressed image. High compression ratio (larger then 100:1) can be reached by using the scheme  $\mathcal{B}$ , which allows to code the test image at a rate of 0.15 bits/pixel keeping an acceptable visual quality, as can be noticed by the luminance component described in Fig. 2. Furthermore, in the same figure you can see as the proposed scheme outperforms the standard JPEG with Arithmetic Coding mode (JPEG-AC) in terms of both visual quality and PSNR.

To conclude, in this paper we have presented a scheme for color image coding based on wavelet decomposition and CVQ. Simulation results show that high compression ratios can be achieved by classifying the vectors according to the frequency orientation of the subbands. Other classification criteria, presented in Sec. 4 and based on the segmentation of the DC information, are under developments.

## 6. REFERENCES

- [1] A. Pentland and B. Horowitz. "A Practical Approach to Fractal-Based Image Compression". In *Data Compression Conference*, pp. 176-185, 1991.

Pred. Scheme	WPVQ	HVQ	Y[dB]	bits/pixel
$\mathcal{A}$	128	128	28.14	0.24
$\mathcal{A}$	256	128	28.62	0.23
$\mathcal{B}$	128	128	27.78	0.17
$\mathcal{B}$	256	128	28.19	0.15

Table 1: Coding performances obtained on the color test image “Lena” using the two prediction schemes and varying the codebook size of WPVQ and HVQ.



Figure 2: Luminance component of the color test image “Lena”: top) original, 24 bits/pixel, middle) coded at 0.15 bits/pixel with the scheme  $\mathcal{B}$  (PSNR = 28.19 dB), bottom) coded at 0.16 bits/pixel by JPEG-AC (PSNR = 23.48 dB).

- [2] T. Naveen and J. Woods. “Subband Finite State Scalar Quantization”. In *IEEE Conf. on Acoustic Speech and Signal Processing*, Vol. V, pp. 613–616, 1993.
- [3] N. Mohsenian and N.M. Nasrabadi. “Subband Coding of Video using an Edge-based Vector Quantization Technique for Compression of the Upper Bands”. In *IEEE Conf. on Acoustic Speech and Signal Processing*, pp. 233–236, 1992.
- [4] J.M. Shapiro. “Embedded Image Coding Using Zerotrees of Wavelet Coefficients”. *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, pp. 3445–3462, December 1993.
- [5] I. Moccagatta and M. Kunt. “A Weighted Distortion Measure for Vector Quantization in the Transform Domain”. In *14th Grets Symposium on Signal and Image Processing*, Vol. 1, pp. 459–462, Juan les Pins, France, September 1993.
- [6] T. Ebrahimi. *Perception Based Localized Linear Operators -Application to Sequence Compression-*. PhD thesis, Swiss Federal Institute of Technology, Lausanne, 1992.
- [7] I. Moccagatta and M. Kunt. “VQ and Cross-Band Prediction for Color Image Coding”. In *Picture Coding Symposium-PCS '94*, Sacramento, California, USA, September 1994.
- [8] R. M. Gray and E. Karnin. “Multiple Local Optimal in Vector Quantizer”. *IEEE Info. Theory*, Vol. IT-28, No. 2, pp. 256–261, March 1982.
- [9] Y. Linde, A. Buzo, and R.M. Gray. “An Algorithm for Vector Quantizer Design”. *IEEE Trans. on Comm.*, Vol. 1, pp. 84–95, January 1980.
- [10] B. Ramamurthi and A. Gersho. “Classified Vector Quantization of Images”. *IEEE Trans. of Comm.*, Vol. COM-34, No. 11, pp. 1105–1115, November 1986.
- [11] O. Egger, W. Li, and M. Kunt. “High Compression Image Coding Using an Adaptive Morphological Subband Decomposition”. *Proceedings of the IEEE*, Vol. 83, No. 2, February 1995.