TREIZIÈME COLLOQUE GRETSI - JUAN-LES-PINS DU 16 AU 20 SEPTEMBRE 1991

VECTOR QUANTIZATION OF COLORED IMAGES BY MEANS OF ADAPTIVE VISUAL MASKING FUNCTIONS

G. Benelli (2), L. Alparone (1), F. Argenti (1)
A. Mecocci (1), L. Vecchietti (1)

(1) Dept. Electronic Engineering, University of Florence, via S. Marta n.3, 50139 Florence, (ITALY). (2) Dept. Electronic Engineering, University of Pavia, via Abbiategrasso 209, 27100 Pavia, (ITALY)

Résumé Récemment beaucoup d'attention a été attribuée au problème de la réduction de la largeur de bande pour les images en couleurs et différentes techniques ont été développées, en particulier les techniques DPCM e HDPCM. En général ces techniques au lieu d'utiliser trois codeurs scalaires individuels, chacun pour chaque composante du signal, emploient la Quantification Vectorielle (VQ) qui garantit des rapports de compression plus élevés. Dans cet article un système pour codage vectoriel psychovisuel est présenté et qui évalue les activités de l'image et utilise les fonctions adaptatives "visual-masking" pour générer le correspondant codebook. En utilisant une telle méthode plusieurs améliorations ont été acquises. Cela est du au fait que le codebook n'a pas été désépaissi de façon adaptative en introduisant une laison correcte entre l'activité de l'image et les paramètres des fonctions "visual-masking" utilisées pendant la génération du codebook.

Abstract. Recently many attention has been devoted to the problem of color image bandwidth reduction and several different techniques have been implemented, which use DPCM (Differential Pulse Code Modulation) or HDPCM (Hybrid Differential Pulse Code Modulation). In general these techniques instead of using three individual scalar coder, one for each signal component, employ Vector Quantization (VQ) that grants higher compression ratios. In this paper a system for psychovisual vector coding is presented which evaluates the image activity and uses adaptive visual-masking-functions to generate the corresponding codebook. By using such a method further improvement in the compression ratio has been achieved. This is due to the fact - that the codebook is not pruned in an arbitrary way, but is generated in an adaptive way by introducing a proper link between image activity and the parameters of the visual-masking-functions used during the codebook computation.

1. INTRODUCTION

Today many image transmission and processing systems use color information as a fundamental integral part. The general preference for color pictures, the greater information contents of color images, and the need of higher spatial resolution that is typical in HDTV (High Definition TV), leads to a huge amount of data. This makes efficient transmission very important and imposes the development of suitable bandwidth reduction techniques.

Many different schemes for color image compression have been designed and tested. PCM (Pulse Code Modulation) based on mean square error minimization [1] or on just noticeable color differences and color signal preprocessing [2], [3]. DPCM methods, based on intra/interframe prediction [4], [5] or based on noise shaping that has been suggested to be an efficient means of reducing the noise visibility [6]. Other interesting techniques employ unitary transforms as: the DCT (Discrete Cosine Transform), the HDT (Hadamard Discrete Transform) or the DWT (Discrete Walsh Transform) [7], [8]. Other methodologies use segmentation of the image into constant chrominance regions (Plateau Coding) [9] or the vectorial representation of colored pixel blocks as in Vector Quantization (VQ) [10].

One important feature in all these schemes is the importance of the quantization law used to encode the color signal in a digital form. The quantizer must achieve a low bit-rate and, at the same time, a good picture quality. In this paper reference to a system based on DPCMVQ (Differential Pulse Code Modulation with Vector Quantization) is made. The critical point in the VQ phase, is the placement of the output vectors in the plane of prediction-errors. From psychovisual studies it has been proved that the human vision system can not perceive small output errors in those parts of the picture that are characterized by fine details i.e. where there is high local image activity. This is known as "Visual-Masking Effect". Recently several methods based on VQ have been proposed to take advantage of such "defects" in the visual system and to obtain compressed images virtually identical to the original ones [11], [12]. In order to achieve this advantage, the quantization vectors are chosen so that the possible quantization errors are smaller than a visual masking threshold computed by means of suitable "Visual-Masking-Functions" (VMFs).

Many different techniques employing the previous approach have been proposed to generate appropriate codebooks. Minor attention has been devoted to the problem of further increasing the compression ratio by means of a reduction of the number of codebook vectors based on the analysis of image properties. In this paper this last topic is covered and a procedure is proposed that evaluates the image activity and uses adaptive parametric VMFs to generate codebooks comprising a lower number of quantization vectors. By using this method further improvement in the compression ratio has been achieved, while the reconstructed image fidelity still remains high.

In Sect. 2 some known results about the masking effect and the methodology used to measure the luminance as well as the chrominance maskings will be reviewed. In Sect. 3 the algorithm for vector placement will be described, while in Sect. 4 a new method used to reduce the number of codebook vectors will be proposed. In Sect. 5 the experimental results are reported and discussed, while in Sect. 6 some conclusions and future developments are summarized.

2. VISUAL MASKING EFFECT

The human visual system when dealing with color images exhibits many interesting behaviors [13]. In particular it has been observed in many experiments that by increasing the details contents of an image, the visual system sensibility to errors is decreased [14]. This phenomenon has been termed "Masking Effect". In other words the human visual system becomes less and less sensitive to errors in those regions where image "activity" is larger. The image "activity" takes into account the variability of the image components: for example, a low-activity region is generally characterized by an high degree of homogeneity, while an high-activity region comprises zones of high gradient-values in one or more color components.

The visual-masking-effect gives the possibility of coding in a coarser way the information that is present in those regions characterized by higher activity, because errors introduced by a coarse representation cannot be noticed by a human observer.

In early measurements of the "Visual Masking Effect" [15] it turned out that this effect is a function of both the luminance and the chrominance activity. The prediction errors eY and eC on the luminance and chrominance signals made by a DPCM stage, can be used as a good measure of such activities. In fact, the prediction errors are more relevant where the activity is higher. When these errors are used as activity estimators, the resulting masking thresholds (MY and MC for Y and C respectively) are approximately given by the following two formulas

$$M_C(e_Y, e_C, Y, C) = b_C(C)\cdot\sqrt{|e_C|}\cdot\left[1+\frac{|e_Y|}{128}\right]\cdot\left[1.1-\frac{255-Y}{1000}\right] \qquad (1)$$

$$M_Y = b_Y\cdot\sqrt{|e_Y|}\cdot\left[1-\frac{Min\{|e_C|,\,|C|\}}{256}\right] \qquad (2)$$

where bY and bC are two constants of value 1.7 and 1.6 respectively [11], [15]. It should be noted that: a) the higher the luminance activity the lower the chrominance error visibility; b) the higher the absolute luminance level, the lower the chrominance error visibility.

3. CODEBOOK CONSTRUCTION

The VMFs discussed in the preceding section, give information about the sensibility (MY and MC) of the visual system to errors done under different viewing conditions. The quantizer fundamental task is that to assign the input vector (eY,eC) to a new quantization vector (eqY,eqC) belonging to the codebook. The placement of (eqy,eqc) with respect to (eY,eC) is chosen in such a way that the visibility thresholds defined by means of MY and MC are not exceeded. By taking into account the definitions of MY(eY,eC,C) and MC(eY,eC,Y,C) the visibility constraints can be satisfied by imposing

$$|e_Y-e_{qY}| < M_Y(e_Y, e_C, C)$$
$$|e_C-e_{qC}| < M_C(e_Y, e_C, Y, C)\,. \qquad (3)$$

In De Lameillieure-Bruyland experiments the previous formulas have been used for the placement of the quantization vector.

An example of the codebooks generated by means of the functions (1) and (2) is reported in Figure 1. It should be noted

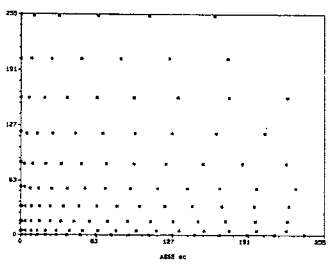

Figure 1 Examples of codebook 200.200 (Luminance.Chrominance)

that the codebook depend on the luminance, chrominance, and prediction errors; however from the experiments it turned out that the codebook structure mainly depends on the value of C.

We have performed many different experiments by applying different codebooks (i.e. codebooks evaluated for randomly chosen values of Y and C) both on real images and on synthetic images. The interesting fact is that even if the actual values used during codebooks generation are very different, the reconstructed images are visually identical to the original ones. Some differences can be found only in the S/N ratio, but this parameter is known to be of little value as an index of visual quality.

A possible conclusion is that the original method proposed in [12] produces codebooks with a redundant number of vectors so that the actual values of Y and C do not influence in a strong way the final image reconstruction quality.

4. CODEBOOK REDUCTION

By taking into account the preceding observations, further improvements in the compression ratio by suitably discarding the redundant codebook vectors can be achieved.

A very simple methodology for codebook reduction has been proposed by De Lameillieure and Bruyland themselves and is based on the fact that the errors produced by the DPCM stages are generally small, i.e. they tend to concentrate around the origin in the error prediction-plane. So, if the vectors produced by the codebook generation procedure are decimated by eliminating those vectors which are far from the origin, the additional error introduced will not be high.

Nevertheless, this method is not efficient because the degradation produced by the decimation is concentrated only in certain regions of the error prediction-plane. This fact can be of dramatic impact if in the original image some luminance or chrominance abrupt transitions are present.

Another possible approach is that of randomly decimating the codebook with a probability density more concentrated far way from the origin. Although this approach may produce good results, it is not adequate in many practical situations because it does not take into account the original distribution of the quantization vectors.

The new method proposed in this paper follows a different approach which has two main properties:

- the user can control the final number of codebook vectors by choosing a suitable multiplicative parameter;

- the vector distribution produced respects the original distribution of the codebook vectors in the error-prediction-plane.

The fundamental idea derives from the analysis of the dependencies among the (eqy,eqc) vector placement in the error-prediction-plane and the visibility thresholds MY and MC. It is evident that if these thresholds are increased the mean distance among the codebook vectors become larger. As a consequence, keeping constant all the other parameters, the final number of vectors in the codebook decreases. In order to obtain higher threshold values leaving unaltered the visual system response "shape", the values given by MY(eY,eC,C) and MC(eY,eC,Y,C) (equations (2) and (1)) have been multiplied by a positive scaling factor $\alpha(N)$ greater than one.

By using this strategy, besides to reducing N, the reconstruction errors introduced by the decimation process are less critical because the vectors are eliminated in a way that respects the visual system behavior.

Another interesting point addressed during the experiments was the link between the final required bit-rate and the number of vectors which are needed to obtain such a bit-rate. The link among the final bit-rate, the image activity, and the number of quantization vectors has been established on a statistical basis. The activity can be measured in many different ways, in our experiments the prediction errors from the two DPCM stages have been employed for activity estimation. The formula used to obtain such an estimate is

$$A = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\sqrt{[e_C^2(i,j) + e_Y^2(i,j)]} \qquad (4)$$

where A is the image activity estimate, r and c represent the number of rows and columns respectively, while eC and eY are the prediction errors in the chrominance and luminance.

To obtain an activity estimate by using (4), the image must be entirely scanned in advance. This action can be computationally expensive: in the actual implementation only a subset of say 10% points may be randomly sampled and used to evaluate the image activity. The estimate based on this random subsampling has been proven suitable for all practical purposes and the resulting coding procedure becomes faster. The final formula that has been obtained is the following one

$$B_r = \log_{10}\left[\frac{5}{6} N\sqrt{A}\right] \qquad (5)$$

where Br is the desired final bit-rate, A is the image activity, while N is the number of vectors which are needed in the codebook. The complete procedure can be summarized as follows:

1) evaluate the mean image luminance $\underline{Y}$ and the mean image chrominance $\underline{C}$;

2) evaluate the image activity A by means of the proposed random subsampling method;

3) by using equation (5) evaluate the number N of desired codebook vectors stated the desired bit-rate;

4) select the scaling factor $\alpha(N)$ on the basis of the value N obtained in the previous step 3);

5) by using the value $\underline{Y}$, $\underline{C}$, evaluate MY(eY,eC,$\underline{C}$) and MC(eY,eC,$\underline{Y}$,$\underline{C}$) then scale by $\alpha(N)$ and generate the compressed codebook to be used in the DPCMVQ.

The previous procedure has also been modified to take into account the eventual variations of luminance and chrominance in different regions of the image. In order to take into account these variations the whole image has been divided into squared subimages and the previous procedure is applied separately to each subimage. By using this last adaptive procedure an increase in performance has been obtained.

5. EXPERIMENTAL RESULTS

Many different experiments have been performed on real images, taken from an HDTV sequence, in order to test the goodness of the proposed quantization methodology.

A first set of results are relative to the methodology applied to the original images considered as a whole and are reported in the following

Activity = 12.8   Codebook = 110Y, 16C

- expected bit-rate = max
  actual = 5.13 bit/pix. SNR = 35.93 Db

- expected bit-rate = 4.90 bit/pix.
  actual = 4.64 bit/pix. SNR = 33.54 Db

- expected bit-rate = 4.50 bit/pix.
  actual = 4.14 bit/pix. SNR = 31.31 Db

- expected bit-rate = 4.20 bit/pix.
  actual = 3.87 bit/pix. SNR = 31.10 Db

It should be noted that the formula (5) is estimated on a statistical basis and so the actual values obtained in the experiments are generally different from the predicted values.

A second set of results obtained by partitioning the original images into subimages and is reported in the following

- expected bit-rate = max
  actual = 6.24 bit/pix. SNR = 37.24 Db

- expected bit-rate =4.50 bit/pix.
  actual = 5.29 bit/pix. SNR = 29.20 Db

- expected bit-rate = 4.00 bit/pix.
  actual = 4.39 bit/pix. SNR = 27.41 Db

- expected bit-rate = 3.80 bit/pix.
  actual = 4.15 bit/pix. SNR = 27.11 Db

As can be easily seen the proposed compression methodology gives good results. As can be expected the performance obtained by subdividing the original image into subimage is higher.

6. CONCLUSIONS AND FUTURE TRENDS

In the paper it has been presented a new technique for adaptive psychovisual compression in DPCMVQ. The method is based on the reduction of the number of codebook vectors; this reduction is performed in a way that respects the Visual-Masking-Effect exhibited by the human visual system. A method to obtain adaptive behavior by subdividing the original image into subimages has been also proposed. The suggested method has been tested by using synthetic as well as real images and the performance obtained gives good results.

In the future it will be explored the possibility to obtain a higher degree of adaptivity by means of more sophisticated techniques whose aim is to subdivide the original image into arbitrarily shaped subregions of different activity and of different chrominance/luminance characteristics (the subregions must be square-shaped in the actual implementation) and then to code each subregion by using the optimum codebook stated the subregion characteristics. Another important point that will be considered is the extension of the technique to the compression of dynamic image sequences.

REFERENCES

[1] J. Max, "Quantizing for minimum distortion", IEEE Trans. Inform. Theory, vol. IT-6, pp. 7-12, Mar. 1960.

[2] F.J. In der Smitten, "Data-reducing source encoding of color picture signals based on optical chromaticity classes", Nachrictentech. Z., vol. 27, pp. 176-181, 1974.

[3] A. Netravali, C.B. Rubinstein, "Luminance Adaptive Coding of Chrominance Signals", IEEE Trans. on Communications, vol. COM-27, pp. 703-710, April 1979.

[4] H. Buley, L. Stenger, "Inter/Intraframe coding of color TV signals for transmission at the third level of the digital hierarchy", Proc. IEEE, vol. 73, pp. 765-772, Apr. 1985.

[5] B. Girod, "Design of switched predictors for interframe DPCM coding of television signals", in Proc. Int. Conf. Digital Signal Processing, pp. 570-576, 1984.

[6] B. Girod, "Reconstruction noise shaping in the context of predictive TV signal coding", in Proc. Int. Conf. Commun., Amsterdam, The Netherlands, pp. 711-717, 1984.

[7] R.J. Clarke, "Transform Coding of Images", Academic Press, London, 1985.

[8] A.K. Jain, "Image Data Compression: A Review", Proc. IEEE, vol. 69, n. 3, pp. 349-389, Mar. 1981.

[9] J.O. Limb, C.B. Rubinstein, "Plateau Coding of the Chrominance Component od Color Picture Signals", IEEE Trans. Commun., vol. COM-22, n. 6, pp. 812-820, Jun. 1974.

[10] H.M. Hang, B.G. Haskell, "Interpolative Vector Quantization of Color Images", IEEE Trans. Commun., vol. COM-36, n. 4, pp.465-470, Apr. 1988.

[11] R. Schäfer, "DPCM coding of the chrominance signals for the transmission of colour TV signals at 34 Mbits/s", Signal Processing, vol. 6, n. 3, pp. 187-199, Jun. 1984.

[12] J. De Lameillieure, I. Bruyland, "Parallel HDPCM with a Vector Quantizer Based on Masking Function", 3rd Int. Workshop on HDTV, Torino, Aug. 1989.

[13] Y. Le Grand, "Optique physiologique", Editions de la Revue d'Optique, Parigi, 1956.

[14] L. Stenger "Quantization of TV chrominance signals considering the visibility of small color differences", IEEE Trans. Commun., vol. COM-25, n. 11, pp. 1393-1406, Nov. 1977.

[15] B. Girod, H. Almer, L. Bengtsson, B. Christensson, "A subjective evaluation of noise-shaping quantization for adaptive intra/interframe DPCM coding of color television signals", IEEE Trans. Commun., vol. COM-36, n. 3, pp. 332-346, Mar. 1988.