

Vector Representation Of Chrominance For Very Low Bit Rate Coding Of Video

Maciej Bartkowiak*, Marek Domański* and Peter Gerken[◇]

*Politechnika Poznańska, Instytut Elektroniki i Telekomunikacji, ul. Piotrowo 3a, 60-965 Poznań, Poland
Phone: +48 61 782 762, Fax: +48 61 782 572, E-mail: mbartkow@et.put.poznan.pl, domanski@et.put.poznan.pl

[◇]Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover,
Appelstraße 9A, 30167 Hannover, Germany
Phone: +49 511 762 5304, Fax: +49 511 762 5333, E-mail: gerken@tnt.uni-hannover.de

Abstract

A chrominance vector quantization technique is proposed as a preprocessing step prior to any kind (e.g. DCT-based or OBASC) of video coding. The operation converts the stream of two-component vectors into a scalar stream of chrominance labels. Therefore the coder processes two channels only: one luminance and one chrominance. After decoding the two chrominance channels are reconstructed from the stream of labels of chrominance codebook entries. Experimental results with still images show recognizable improvement of the subjective quality by a constant compression ratio.

1. INTRODUCTION

Palette representation of color images is well-known from computer graphics where it has been used in order to reduce the amount of data needed to refresh the monitor screen. The principle is that the pixel values are addresses of elements of a predefined table called a palette. These elements consist each of three coordinates of the colors defined by the palette. The process of designing and applying color palettes is in fact a vector quantization technique. Some respective solutions have been already described in the references [3-6,13].

A color video sequence is in fact a vector signal where each pixel of an individual frame is represented by a three-element vector. This representation is highly redundant because usually only a small part of all possible combinations is present in an individual frame and even in a whole sequence [7,8]. The commonly used techniques process the both chrominance channels almost separately taking no advantage of their mutual dependencies.

Our idea is to use vector quantization of chrominance for very low bitrate channels where extremely high compression ratios are required. Before input to the source coder (for example DCT-based coder [11,12] or object-based analysis-synthesis coder [2,10]) an image or a video sequence shall be preprocessed in that way that the two chrominance components of the sequence are converted to a scalar signal being a stream of chrominance labels. The technique includes automatic codebook design and update

according to changes of the frame content.

The assumption for the work is that there is as little as possible interaction with the following coder so that any type of source coder can be used. Similarly at the decoder side only some postprocessing shall be performed for recalculation of the actual color coordinates.

2. CODING STRATEGY

The approach is based on two-dimensional vector quantization on the chrominance plane. In order to reduce the perceivable distortions introduced by the vector quantization, it is recommended to transform the chrominance input signals U and V into the a^* and b^* coordinates from the CIE $L^*a^*b^*$ color system [1]. Application of perceptually more homogenous coordinates yields in codebooks improved in terms of subjective quality.

Therefore vector quantization with the nearest neighbor search uses the Euclidean distance on the a^*b^* plane which approximates the perceptual differences between colors. Nevertheless the color transformation does not play a crucial role in the proposal and can be omitted.

The vector quantizer transforms the two input chrominance components into the scalar signal being a stream of labels of (a^*,b^*) or (U,V) pairs. At first (cf. Fig.1), a basic codebook is designed which is later enlarged by inserting interpolated entries. The size of the basic codebook has been fixed to 32 because experimental results show that for natural images and sequences in QCIF/CIF picture format [12] this is a reasonable value that does not lead to visible degradation of the picture quality. Note that the codebook entries represent chrominance values only and that even with this small number of entries, there is still the possibility to generate lots of colors in combination with the individual luminance values. The basic codebook is computed for each frame and in the cases of video sequences it has to be transmitted at least for the first frame. For the next frames, all the entries are compared with those from the previous frame. The new codebook is sent if a dramatic change of the scene is detected. Otherwise the same codebook is used for consecutive frames. The codebook entries are losslessly encoded and

transmitted as side information.

The 32 basic codebook entries are ordered and mapped onto the range from 0 to 255 (or from 16 to 240 in order to be compliant with standard input data formats). The order of the entries is very important because it deeply affects the performance of the system. The differences between the values representing the consecutive codebook entries are proportional to the distance on the chrominance plane used. The intermediate values represent chrominances which can be calculated by linear interpolation between two neighboring codebook entries. These intermediate values form together with the basic codebook entries the enlarged codebook.

The pixel chrominances are mapped using the nearest neighbor search to a chrominance labelled by an integer which range from 0 to 255 (or 16 to 240).

Due to quantization in the video coder, the values of chrominance labels can be changed. The decoder assigns to the decoded chrominance labels those pairs of interpolated chrominance values which are represented by them. For this purposes the basic codebook is interpolated at the decoder side in the same way as at the coder side.

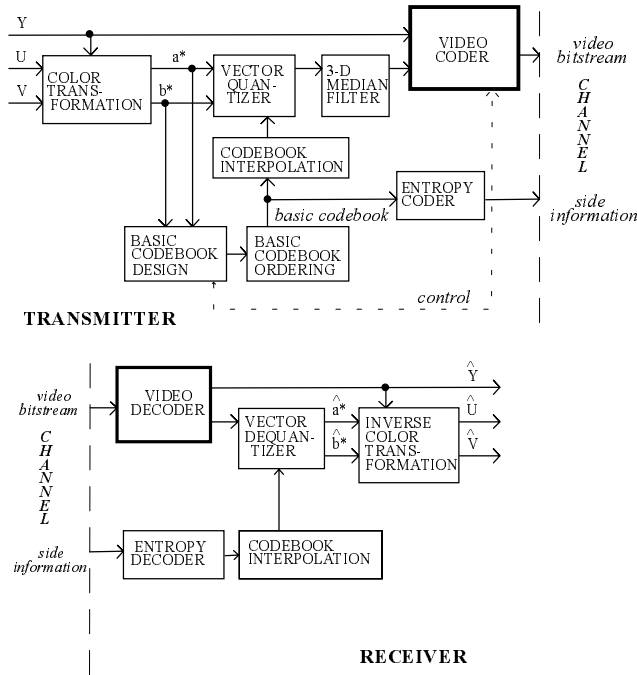


Fig. 1. General structure of the transmitter and receiver.

3. DESIGN OF BASIC CODEBOOK

There exist several techniques to design codebooks in the color spaces [4-6,13]. A common approach is to start with a relatively poor (in the sense of total square error) codebook and improve it using the Linde-Buzo-Gray (LBG) algorithm [9]. On the contrary to those techniques some algorithms based on binary splitting without iterative optimization have been developed [4]. The method proposed here is also a binary-split technique.

As a measure of a vector distance, the Euclidean

norm on the chrominance plane is used. At the beginning, the codebook has only one entry being a vector with its components calculated as the mean values of all chrominances in a frame. In the first step, the set of all chrominances is divided into two subsets according to the following rule:

1. The origin of the coordinate diagram is set to the end point of the initial vector.
2. The direction that maximizes the function

$$M(\varphi) = \sum_n \|X_n\|^2 \cos^2 \alpha_n \quad (1)$$

is determined. Here, $\|X_n\| \cos \alpha_n$ is a distance between the chrominance and the representing vector projected onto the searched direction and n is the index of a pixel in the frame

We rearrange

$$\begin{aligned} M(\varphi) &= \sum_n |X_n|^2 \cos^2 [\arg\{X_n\} - \varphi] = \\ &= \frac{1}{2} \sum_n |X_n|^2 + \frac{1}{2} \sum_n |X_n|^2 \cos 2[\arg\{X_n\} - \varphi] \end{aligned} \quad (2)$$

We notice that $\frac{1}{2} \sum_n |X_n|^2$ is constant.

Moreover,

$$\begin{aligned} \sum_n \cos [2 \arg\{X_n\} - 2\varphi] &= \\ \sum_n \operatorname{Re}\{e^{-j2\varphi} e^{j2 \arg\{X_n\}}\} &= \operatorname{Re}\{e^{-j2\varphi} \sum_n e^{j2 \arg\{X_n\}}\} \end{aligned} \quad (3)$$

so $M(\varphi)$ is maximized for

$$\varphi = -\frac{1}{2} \arctan 2 \frac{\sum_n \operatorname{Re}\{X_n\} \operatorname{Im}\{X_n\}}{\sum_n \operatorname{Re}^2\{X_n\} - \sum_n \operatorname{Im}^2\{X_n\}} \quad (4)$$

3. The set of colors is divided into two subsets (Fig. 2). The border between these two subsets is the line through the origin which is perpendicular to the above determined direction (i.e. the direction defined by the angle φ).

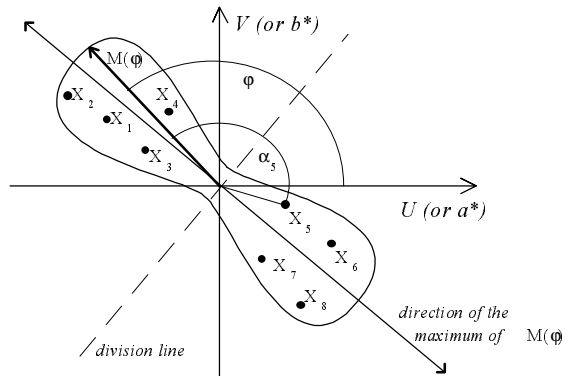


Fig. 2. The splitting rule.

4. The two new vectors representing the two subsets are calculated (their coordinates are mean values of the coordinates of all chrominances of the respective subsets).

The above described procedure is repeated iteratively.

At each step, only the set of chrominances with the highest total square error of its vector representation is divided into two subsets. In general, n steps of this algorithm result in an n -element basic codebook (Fig.3).

Consecutive frames of video sequences usually show similar chrominance histograms and therefore produce codebooks with similar tree structure. This property makes the frame-to-frame comparison of the codebooks easy. This comparison is necessary for the decision whether the codebook (or optionally a part of it) must be retransmitted.

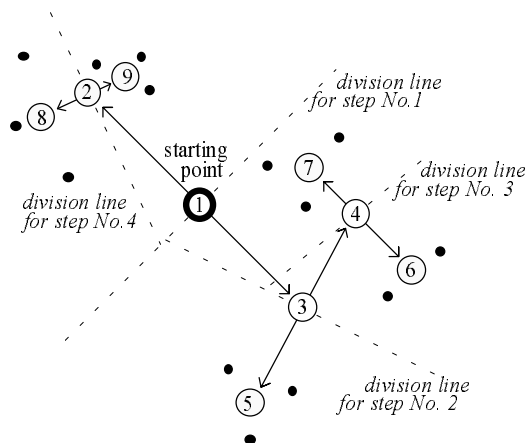


Fig.3. Binary splitting example.

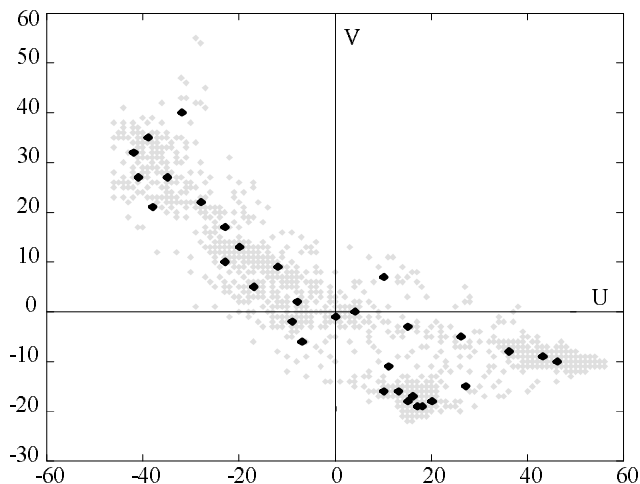


Fig.4. Basic codebook entries (black dots) of the first frame of the test sequence "CLAIRE" shown on the set of all chrominances present in this frame (grey points).

4. BASIC CODEBOOK ORDERING

Random ordering of the basic codebook would lead to very broad power spectrum of the signal fed into the coder. This signal is an image with pixel values being the labels of the codebook entries. The goal of proper ordering should be to make this image as low-frequency as possible. The authors have examined some strategies of ordering and they choose a strategy of simultaneous basic codebook generation and ordering. At each step one codebook entry

is substituted by two new entries as described above. The set of codebook entries is already ordered and the two new entries are put onto the place of the removed one. The order that minimizes the distances to the neighboring entries is chosen out of the two possibilities of ordering the two new entries. In fact, this strategy leads to relatively good results (Fig.5).

5. CODEBOOK INTERPOLATION

In order to get finer quantization, the basic codebook is augmented with interpolated codebook entries.

The ordered set of the basic codebook entries is mapped onto the set of integers in the range of 0 to 255 (or 16 to 240 to be compliant with video formats). The integers assigned to the codebook entries are henceforth called labels. A label difference of two consecutive basic codebook entries is set proportional to the respective distance on the chrominance plane. All other integers from the above mentioned range are assigned to the interpolated codebook entries. Therefore the longer the distance between two consecutive codebook entries the more interpolated entries are inserted in between. The chrominances corresponding to the interpolated codebook entries are equally distributed on the straight lines between consecutive basic codebook entries (cf. Fig.5.)

The above rule is known on the decoder side, therefore there is no need to transmit the information about interpolated codebook entries. Only frame-to-frame updates of the basic codebook entries have to be transmitted as side information if needed.

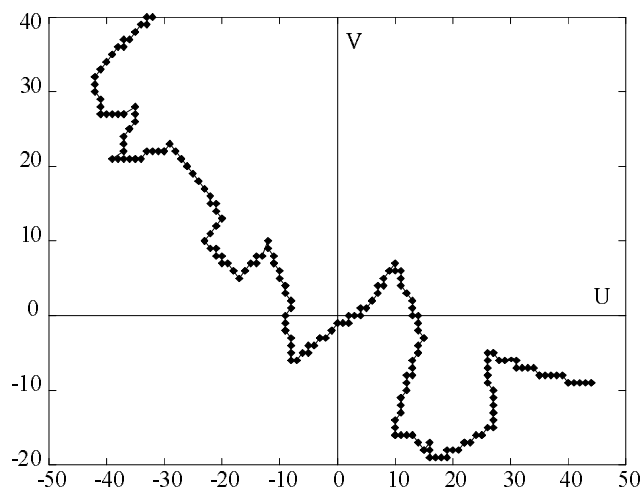


Fig. 5. Ordered entries of the whole codebook of the first frame of the sequence "CLAIRE".

6. MEDIAN FILTERING

After quantization some impulses are visible in the stream of chrominance labels. These impulses correspond to high frequencies. This would lead to lower efficiency of the overall coding system. The impulses can be removed

using 3-D median filters. The authors get good results with 7-point 3-D median with cross mask in all the three dimensions. Such a filter smooths the quantized chrominance sequence but it does not introduce visible degradation.

7. RESULTS AND CONCLUSION

The technique has been examined using standard videophone test sequences "CLAIRE" and "SALESMAN" in QCIF format.

The first series of tests used the technique proposed together with intraframe JPEG coding [11]. The results proved that application of the above described chrominance vector quantization substantially improve subjective quality at the same chrominance compression ratio. In particular, the annoying effects of the false colors are strongly reduced at high compression ratios of order of 1:40 and greater (Fig. 6). Similar results at higher compression ratios have been obtained for popular still test images like "LENA" and "BOATS".

The second series of experiments used the technique together with the H.263 coder [12]. Now, the authors are dealing with the problem of adopting the control procedures of the coder to the task of encoding a scalar stream labels of the codebook entries instead of two chrominance channels.

Future investigations are aimed at combining the proposed technique with the object based analysis-synthesis coder (OBASC) [2,10].

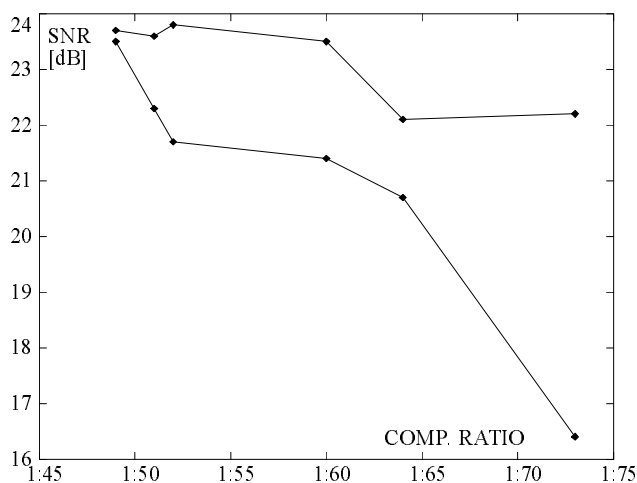


Fig. 6. SNR versus compression ratio for the first frame of the sequence "CLAIRE". Color differences have been measured according to the formulas from the paper [14].

ACKNOWLEDGEMENT

The work has been supported under the NATO Linkage Grant HTECH.LG 941338.

REFERENCES

- [1] G.Wyszecki, W. Stiles, Color science, Wiley 1982 .
- [2] H.G.Musmann, Object based analysis synthesis coding, IEEE Int. Symposium on Circuits and Systems, Tutorials, eds.: C.Tommazou et al., London 1994.
- [3] J.Barilleaux, R.Hinkle, S. Wells, Efficient Vector Quantization for Color Image Encoding, Proc. ICASSP 1987, vol.2, pp.740-743.
- [4] R.S.Gentile, E.Walowit, J.P.Allebach, Quantization and multilevel halftoning of color images for near original image quality, J.Opt.Soc.Amer.-A, Vol.7, No.6, 1990
- [5] M.T.Orchard, Ch.A.Bouman, Color Quantization of Images, IEEE Transactions on Signal Processing, December 1991
- [6] M.Domański, M.Bartkowiak, Color image archi- vization for medical purposes, Journal on Communica- tions, vol. XLV, pp.66-68, July-August 1994.
- [7] M.Bartkowiak, M.Domański, Palette Representation for Data Compression of Color Video, Proc. XVIII Nat. Conf. Circuit Theory Elec. Syst, pp. 473-478, Polana Zgorzelisko 1995.
- [8] M.Bartkowiak, M.Domański, Color statistics in image sequences and their implementations for VLBC, 2nd Int. Workshop on Image Processing, pp. 68-72, Budapest 1995.
- [9] Y.Linde, A.Buzo, R.Gray, An algorithm for vector quantization design, IEEE Trans. on Commun., vol. COM-28, pp. 84-95, 1980.
- [10] P.Gerken, Object-based analysis-synthesis coding of image sequences at very low bit rates, IEEE Trans. Circuits Syst. Video Techn., vol 4, pp. 228-235, 1994.
- [11] ISO International Standard 10918, Digital compression and coding of continuous-tone still images, Geneva 1994.
- [12] ITU, Draft recommendation H.263, Video coding for narrow telecommunication channels at < 64kbit/s, April 1995.
- [13] E. Roytman, C.Gotsman, Dynamic color quantiza- tion of video sequences, IEEE Trans. Vizualization and Computer Graphics, vol.1, pp. 274-286, 1995.
- [14] I.H.Godlove, Improved Color-Difference Formula, with Applications to Perceptibility and Acceptability of Fadings, J. Opt. Soc. Am., 41, 11, pp.760-772 (Nov., 1951)