A PERFORMANCE MODEL FOR THE MPEG CODER

G. Calvagno, G.A. Mian, A. Moro, R. Rinaldo
Dipartimento di Elettronica e Informatica
Via Gradenigo 6/a, 35131 Padova, Italy
Tel: +39-49-827 7731, Fax: +39-49-827 7699,
E-mail: calvagno@dei.unipd.it

ABSTRACT

The MPEG video coding standard provides the syntax and semantics of bit streams representing compressed video. The underlying algorithm uses block matching motion compensation and block based DCT, with run-length coding of the quantized coefficients. It is important to derive models that allow to predict, for a given input sequence, the algorithm performance in terms of quality versus bit rate. In this work, we show that a simple model can be used to this purpose, despite the complexity of the overall MPEG algorithm. The model can be conveniently used to determine the quantizer parameters that give a desired quality or bit rate. For instance, in buffer control, it is necessary to precisely adapt the input rate to the buffer content in order to prevent overflow and underflow.

1 INTRODUCTION

The video coding technique described in the MPEG standard [1] comprises a DPCM loop that uses block matching motion compensation to reduce the temporal redundancy between adjacent frames of the input sequence. To spatially decorrelate the difference images, a two dimensional discrete cosine transformation (DCT) is applied to blocks of samples of dimension 8×8. From time to time, the DPCM loop is reinitialized and coding of the motion compensated difference images (inter-frame coding) is replaced by coding of the original frames (intra-frame coding). The resulting DCT coefficients are then quantized using an adaptive uniform threshold quantizer with conveniently chosen reconstruction values.

It is well known that each of the DCT coefficients can be modeled as an uncorrelated random variable with generalized Gaussian probability density function (pdf) given by

$$p_\nu(a) = \frac{\nu}{2\sigma_i\sqrt{\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}}\Gamma(1/\nu)} \exp\left[-\left(\frac{|a|}{\sigma_i}\sqrt{\frac{\Gamma(3/\nu)}{\Gamma(1/\nu)}}\right)^\nu\right], \tag{1}$$

where (excluding the case of the DC coefficient in intra-frame mode) the parameter ν is approximately equal to 0.7 [2], and the variance σ_i^2 can be estimated from the sample variance of the i-th DCT coefficient.

From the quantizer characteristic $\mu(x)$ and the pdf in (1), one can compute the entropy of the i-th quantized coefficient as

$$H_i = -\sum_{k=-L}^{L} P_{i,k}\log_2 P_{i,k} \tag{2}$$

where $P_{i,k}$ denotes the probability of obtaining the k^{th} symbol at the quantizer output. Similarly, one can compute the quantization error variance (i.e., the distortion) that is given by

$$\sigma_{e,i}^2 = \mathrm{E}\left[(x_i - \mu(x_i))^2\right]. \tag{3}$$

The probabilities $P_{i,k}$ in (2) and the variance in (3) can be numerically computed by means of the absolute moment function

$$\begin{aligned} F_\gamma^\nu(u) &= \int_0^u |a|^\gamma p_\nu(a)da \\ &= \mathrm{sign}(u)\tfrac{\alpha^\gamma\Gamma((\gamma+1)/\nu)}{2\Gamma(1/\nu)} P\left((\gamma+1)/\nu, \left(\tfrac{|u|}{\sigma}\tfrac{\Gamma(3/\nu)}{\Gamma(1/\nu)}\right)^\nu\right) \end{aligned}$$

where

$$P(b,y) = \frac{1}{\Gamma(b)}\int_0^y a^{b-1}e^{-a}da$$

is the incomplete Gamma function [3]. Specifically, the absolute moment function of order zero, F_0^ν, is used to obtain the distribution of the generalized Gaussian random variable, $P_\nu(a) = \frac{1}{2} + F_0^\nu(a)$ for $a \geq 0$ and $P_\nu(a) = \frac{1}{2} - F_0^\nu(-a)$ for $a < 0$, from which the probabilities $P_{i,k}$ and the entropy H_i can be computed. Similarly, the distortion $\sigma_{e,i}^2$ can be computed from the absolute moment function of order two, $F_2^\nu(u)$.

Starting from the above considerations, in the following sections we will derive a simple model that allow to predict, for a given input video sequence, the performance of the MPEG coder in terms of quality versus bit rate. The model can be conveniently used to determine the quantizer parameters that give a desired quality or bit rate. This is particularly useful for buffer control in constant rate coding, where it is necessary to precisely

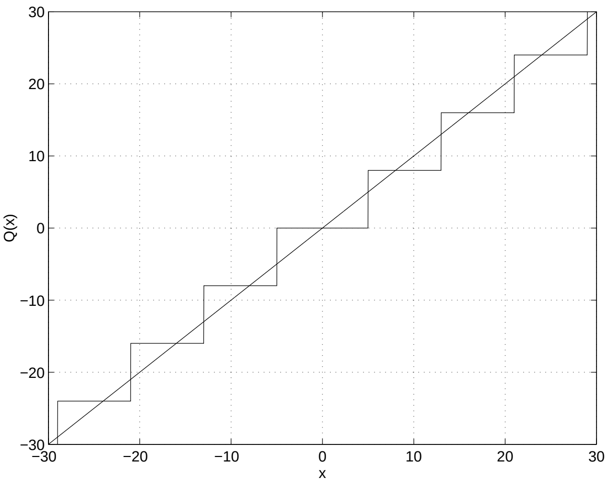

Figure 1: The MPEG quantizer characteristic for intraframe coding and `mquant`=4.

adapt the input rate to the buffer content in order to prevent overflow and underflow. Similarly, the performance prediction given by the model can be used to adapt the coder parameters to have a constant quality reconstructed video sequence.

2 MPEG QUANTIZER PERFORMANCE

In the MPEG coder, intraframe images and motion compensated difference images are divided into 8×8 blocks which are DCT transformed before quantization. The quantizer proposed by the standard computes the quantization levels for a macroblock on the basis of the available bit rate and of the local activity [1]. The quantizer used in MPEG is a modified uniform quantizer specified by a single parameter `mquant`. DCT coefficients are rounded to the nearest integer before quantization.

For intraframe coding, the quantizer entails a dead-zone around the origin, whose amplitude $(-t, t)$ depends on the parameter `mquant`. The reconstruction levels are $k\Delta$, $k \in Z$, where $\Delta = 2\,\texttt{mquant}$ is the step size of the uniform quantizer. The reconstruction levels do not coincide with the midpoint of each quantization interval, as it would be in a uniform quantizer. Fig. 1 shows the quantizer characteristic for `mquant`=4.

For interframe coding, the quantization step is $\Delta = 2\,\texttt{mquant}$, and the quantizer has a dead-zone twice as large as the quantization step, i.e., $t = \Delta$. The reconstruction levels are the midpoints of each quantization interval.

In this section we compare, in the case of intraframe coding, the performance of the MPEG quantizer with that of a uniform quantizer, of a dead-zone quantizer with a modified threshold $t = \Delta/2 + \Delta/16$ and reconstruction levels $k\Delta$, and of an optimum uniform quantizer where the reconstruction levels are computed as the center of mass of the input coefficients pdf in each quantization interval.

Transform coding in the MPEG standard can indeed be viewed as a sophisticated quantization scheme. Let us denote with σ_d^2 the variance of the input image to the transform coder. Input coefficients are transformed into 8×8 vectors of DCT coefficients, whose variances $\sigma_i^2, i = 1, ..., M = 64$ are related to σ_d^2 by

$$\sigma_d^2 = \frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 \tag{4}$$

due to the orthogonality of the DCT transformation. If we assume independence of the quantized DCT coefficients, their entropy, in bit/pixel, results to be

$$H = \frac{1}{M} \sum_{i=1}^{M} H_i, \tag{5}$$

where H_i is the entropy of the i-th quantized DCT coefficient. Moreover, the distortion in the reconstructed image is $\sigma_e^2 = \sum_{i=1}^{M} \sigma_{e,i}^2 / M$, where $\sigma_{e,i}^2$ is the quantization error variance for coefficient i.

The overall performance of the transform coding system can therefore be expressed by the shape factor $\epsilon^2(H)$ of an *equivalent* quantizer, namely

$$\sigma_e^2 = \epsilon^2(H) \sigma_d^2 2^{-2H}. \tag{6}$$

Using (2) and (3) in Section 1, we can compute the distortions $\sigma_{e,i}^2$ and entropies H_i for a given quantizer, once the input pdf is assigned. Fig. 2 shows the shape factor $\epsilon^2(H)$ calculated from (4), (5) and (6) in the case of intracoding for the quantizers under investigation. In this case, σ_d^2 is equal to the input image variance σ_x^2. The variance of each DCT coefficient is estimated by averaging the sample variances of several frames of the image sequence *Calendar*. According to the standard, the values for the MPEG quantizer are calculated including rounding of the DCT coefficients before quantization. In the simulations, all the MPEG visibility mask weights are set to the constant value 16, although similar conclusions can be drawn in the case of variable weights. The MPEG quantizer and the modified-threshold quantizer have a performance very similar to that of the optimal uniform quantizer, despite the fact that the reconstruction levels are immediately determined from Δ and do not need to be computed from the input pdf. Notice that the MPEG and modified-threshold quantizer have reconstruction levels toward the end point of the quantization interval closest to the origin, and that the same happens for the optimum uniform quantizer for a generalized Gaussian distribution. Moreover, the modified-threshold quantizer has a good performance also at high bit rates, while the MPEG quantizer should not be used at bit rates greater than 4 bit/pixel. The values `mquant`=2 (H = 4.4 bit/pixel) and `mquant`=1 (H = 5.3 bit/pixel) for the MPEG quantizer correspond to $\epsilon^2(H) \simeq -2$ dB and $\epsilon^2(H) \simeq -0.25$ dB, and are not

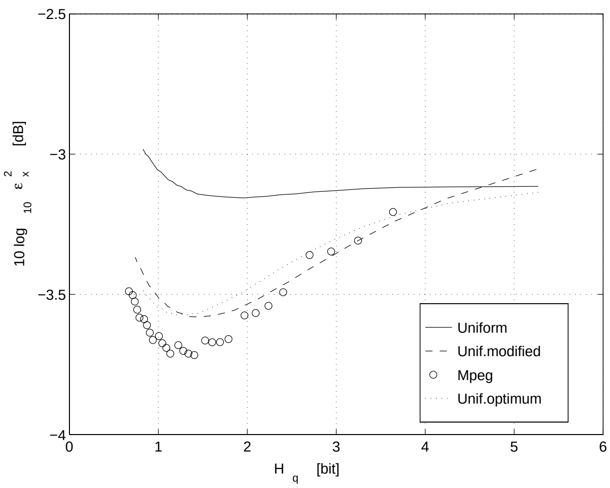

Figure 2: Shape factor for different quantizers.

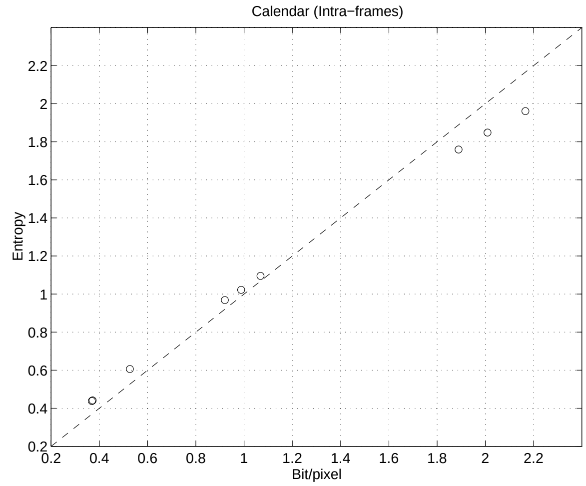

Figure 3: Entropy vs. actual bit rate.

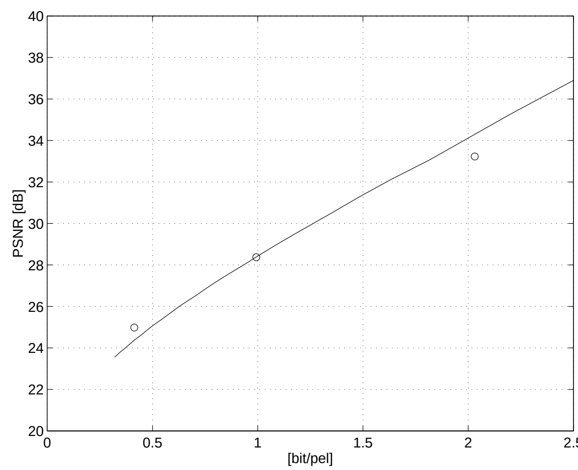

Figure 4: *Calendar* intra frames: PSNR vs. Entropy.

shown in the figure. As a matter of fact, the overall characteristic gives poor performance at high bit rates, because rounding of the DCT coefficients before quantization introduces an additional error of comparable amplitude. In our experiments with real-world video sequences, we found that the use of the modified-threshold quantizer gives a slight overall PSNR improvement over the MPEG quantizer.

For difference images, the variances σ_i^2 depend on the bit rate. For overall bit rates of 2 Mbit/s, 4 Mbit/s and 9 Mbit/s, we find again similar shape factors for the MPEG quantizer, the modified-threshold quantizer and the optimum uniform quantizer.

3 MPEG CODER PERFORMANCE

In this section, we show that the generalized Gaussian model and the coefficient independence assumption can be successfully used to predict the overall rate-distortion performance of the MPEG coder for real-world video sequences. Assuming a generalized Gaussian model with $\nu = 0.7$ for the DCT coefficients, we estimate the variances of the random variables from the sample variances, and compute the distortion and entropy using (4), (5) and the results of Section 1. As a matter of fact, the run-length coding technique used in MPEG can be viewed as an efficient way to code the vectors of DCT coefficients. We expect therefore that the entropy (5) can be used to predict the actual number of bits used to code the vectors.

Figure 3 shows the entropy calculated from (5) versus the actual bit rate obtained from the MPEG coder (running at 2 Mbit/s, 4Mbit/s and 9 Mbit/s) for three intra-coded frames of the luminance of the video sequence *Calendar*. Ideally, for a given quantizer, the entropy should be equal to the actual bit rate (dashed line). Note the very good correspondence between the entropy (5), calculated with the proposed model, and the rate obtained for actual images.

In Figure 4, the continuous curve plots the PSNR predicted by the model for some intra-coded frames of *Calendar* as a function of the entropy. The curve computed with the model is relative to DCT coefficient variances averaged over three frames. In the same plot, the circles correspond to the actual bit rate and distortion obtained with the MPEG coder at three bit rates. Each circle corresponds to averages over three intra-coded frames of the luminance of the video sequence *Calendar*. Again, there is a very good correspondence between the values predicted by the model and the ones obtained with the actual MPEG coder. Figure 5 shows a similar plot for the image sequence *Flower Garden*.

For interframe coding, the input difference image variance depends on the bit-rate. Figure 6 shows the average PSNR for seven difference images of type P of the image sequence *Calendar* at three bit rates. The symbols correspond to the actual performance of the MPEG

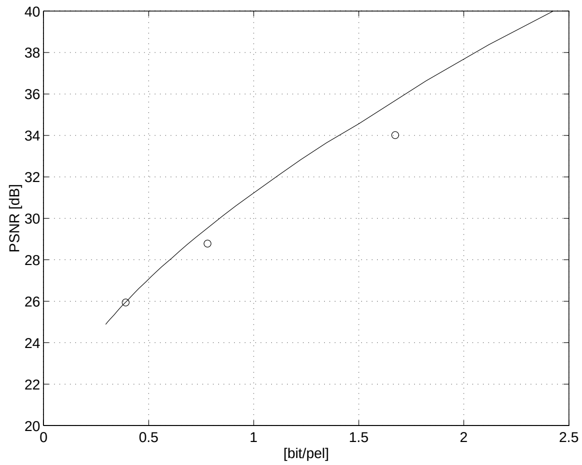

Figure 5: *Flower* intra frames: PSNR vs. Entropy.

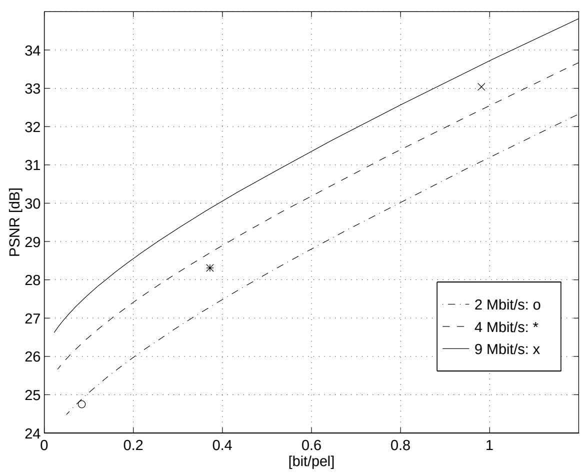

Figure 6: *Calendar* P frames: PSNR vs. Entropy.

coder, while the curves are calculated using the model on the basis of averaged variances. Again, there is a very good correspondence between the values predicted by the model and the actual performance of the coder. Figure 7 shows a similar plot for the image sequence *Flower Garden*.

Figure 8 shows the average PSNR for 18 difference images of type B of the image sequence *Calendar* at various bit rates. It can be seen from the figure that the proposed model accurately predicts the actual bit rate and distortion also in this case.

References

[1] "Coded Representation of Picture and Audio Information," ISO/IEC JCT1/ SC29 WG 11 MPEG, Test Model 5, Draft Revision 2, April 1993.

[2] P.H. Westerink, *Subband Coding of Images,* Ph.D. Thesis, Delft University, The Netherlands, 1989.

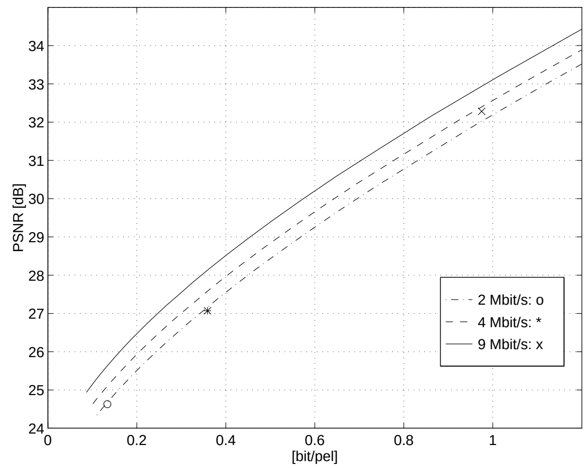

Figure 7: *Flower* P frames: PSNR vs. Entropy.

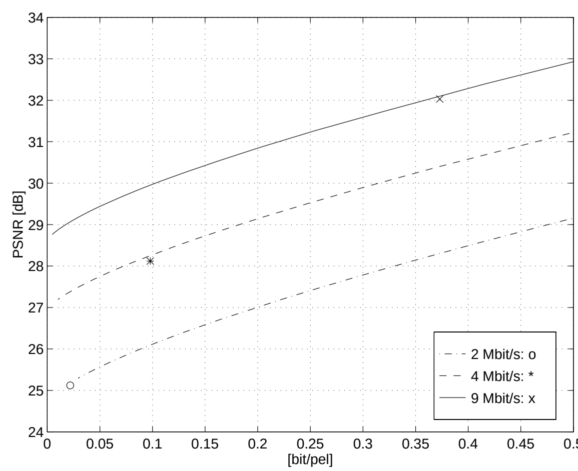

Figure 8: *Calendar* B frames: PSNR vs. Entropy.

[3] M. Abramowitz and A. Stegun, *Handbook of Mathematical Functions*, New York: Dover Publications, 1964.

[4] K.A. Birney and T.R. Fischer, "On the Modeling of DCT and Subband Image Data for Compression," *IEEE Trans. on Image Processing,* vol. 4, no.2, pp. 186-193, Feb. 1995.

[5] H. Gharavi, A. Tabatabai, "Subband Coding of Monochrome and Color Images," *IEEE Trans. on Circuits and Systems,* vol. 35, no.2, pp. 207-214, Feb. 1988.

[6] G. Calvagno, C. Ghirardi, G.A. Mian and R. Rinaldo, "A buffer control technique for video coding", *Proc. 1995 European Symposium on Advanced Networks and Services: Compression Technologies and Standards for Image and Video Communications*, Amsterdam, March 1995, pp. 31–40.