1—4

Efficient Target Bit Allocation Scheme in a Rate-Distortion Sense

W. Y. Lee and J. B. Ra

Dept. of Electrical Engineering
Korea Advanced Institute of Science and Technology
373-1 Kusongdong, Yusonggu, Taejon, Republic of Korea

Abstract : Bit allocation is a critical problem in video encoding such as MPEG. To improve the quality of the reconstructed sequence for a given bit rate, the assigned target bits for a group of pictures (GOP) must be allocated to each picture efficiently. In this paper, we derive a target bit allocation algorithm for more efficient rate control, by assuming that the average rate-distortion curve for an input source is logarithmic. This target bit allocation is based on Shannon's rate-distortion theory, which deals with the minimization of source distortion subject to a channel rate constraint. Simulation results show that the proposed target bit allocation algorithm provides better performance than the one in MPEG-2 Test Model 5 (TM5).

1. Introduction

The rate control algorithm plays a critical role in the hybrid DPCM/DCT video coders like the MPEG [1]. Data compression in these coders is achieved by the quantization operation which maps the samples of block DCT coefficients into a finite set. To exploit the nonuniform distribution of the block energy over DCT coefficients, the quantized DCT coefficients are encoded using a variable length code, typically a Huffman/run-length code, which generates a variable rate bit stream. In order to transmit the variable rate bit stream over a fixed rate channel, a channel buffer regulating the output bit stream, is required. To prevent the buffer from overflowing or underflowing, the fluctuation of the output bit rate should be limited within a reasonable range. And this control can be achieved by adjusting the quantization level [1]. Typical technical issues associated with the channel buffering technique include the size of the channel buffer, and the frequency and degree of adjustment of quantization levels.

To determine the proper quantization step size for a given buffer fullness, an analytical model of the relationship between buffer fullness and quantization is used in the MPEG-2 [1]. A probabilistic model in the coding rate prediction was also tried [2]. To improve the quality of the reconstructed image sequence further, buffer control methods based on the distortion-rate sense were considered [3, 4]. To apply these methods to various pictures in the GOP, however, the appropriate estimation of bits needed for each type of picture is important. The existing target bit allocation schemes where the complexity for pictures is assumed to be constant [5, 6], are not efficient in a rate-distortion sense.

In this paper, we derive a new formula for target bit allocation. This target bit allocation scheme is based on Shannon's rate-distortion theory, and is induced by assuming the exponential rate-distortion curve for an input image. The algorithm minimizes the average distortion in a GOP for a given target bit rate. Its coding performance is examined and compared with the standard MPEG-2 coding scheme.

2. Rate-Distortion Framework for Target Bits

Let us consider the relationship between the distortion and rate for a GOP; an average rate-distortion curve $R(D)$ or average distortion-rate curve $D(R)$. It is noteworthy that $R(D)$ depends on a set of quantizers, entropy coder, and statistics of the input image [7]. It is well known that a high resolution uniform quantizer yields an exponential distortion-rate curve when the distortion is a mean square error (MSE). It is also known that the average distortion-rate curve of many input images can be assumed to be exponential [3].

Let us assume that the average coding bit rate for I, P, and B-pictures are denoted by R_I, R_P, and R_B, respectively; and the numbers of I, P, and B-pictures in a GOP, are denoted by N_I, N_P, and N_B. Then we have the following relation.

$$R_I + N_P R_P + N_B R_B = R_{GOP}. \qquad (1)$$

Here R_{GOP} is the total target number of bits for all the pictures in a GOP. We now consider a target bit allocation problem in a rate-distortion sense. The constrained problem which minimizes the average distortion $D(R)$ for a target average bit rate R can be converted to an equivalent unconstrained problem by using a Lagrange multiplier λ [8]. Thus the unconstrained problem becomes the minimization of the Lagrangian cost function which is defined as

$$J(\lambda) = D(R) + \lambda R. \qquad (2)$$

It can be shown that at R-D optimality, all pictures operate at points with a constant optimal slope on their rate-distortion curves [9]. Thus, by assuming that a rate-distortion curve for an input image is exponential, we will drive the algorithm which minimizes the average distortion of a GOP for a given R.

3. Proposed Target Bit Allocation Scheme

Let us assume that the average distortion-rate curve for each picture type is exponential[3, 7], i.e.,

$$D_I(R_I) = \sigma_I^2 \alpha_I^{-R_I},$$
$$D_P(R_P) = \sigma_P^2 \alpha_P^{-R_P},$$
$$D_B(R_B) = \sigma_B^2 \alpha_B^{-R_B}, \qquad (3)$$

where σ_I^2, σ_P^2, and σ_B^2 are the variances and $\alpha_I(>1)$, $\alpha_P(>1)$, and $\alpha_B(>1)$ are the parameters. Since $\lambda_I = \lambda_P = \lambda_B$ at optimal points, we have

$$\alpha_I^{-R_I} \sigma_I^2 \ln \alpha_I$$

$$= \alpha_P^{-R_P} \sigma_P^2 \ln \alpha_P$$
$$= \alpha_B^{-R_B} \sigma_B^2 \ln \alpha_B. \quad (4)$$

Using Eqs. (1), (3), and (4), and defining the parameters

$$C_I = \sigma_I^2 \ln \alpha_I,$$
$$C_P = \sigma_P^2 \ln \alpha_P,$$
$$C_B = \sigma_B^2 \ln \alpha_B, \quad (5)$$

we can derive the following relations.

$$R_I = \frac{R_{GOP} - N_P \log_{\alpha_P} \frac{C_P}{C_I} - N_B \log_{\alpha_B} \frac{C_B}{C_I}}{1 + N_P \log_{\alpha_P} \alpha_I + N_B \log_{\alpha_B} \alpha_I}, \quad (6)$$

$$R_P = \frac{R^*_{GOP} - N_B \log_{\alpha_B} \frac{C_B}{C_P}}{N_P + N_B \log_{\alpha_B} \alpha_P}, \quad (7)$$

$$R_B = \frac{R^*_{GOP} - N_P \log_{\alpha_P} \frac{C_P}{C_B}}{N_B + N_P \log_{\alpha_P} \alpha_B}, \quad (8)$$

where R^*_{GOP} denotes $N_P R_P + N_B R_P$. Eqs. (6-8) represent the average bit rate for each picture type in steady state. To apply these expressions to the motion compensated hybrid DPCM/DCT video coders like the MPEG [1], by using Eqs. (3-5), we can modify them as follows.

$$T_I = \max\{S_I + \frac{R_r - (S_I + N'_P S_P + N'_B S_B)}{1 + N'_P \frac{D'_P}{D'_I K_P} + N'_B \frac{D'_B}{D'_I K_B}}, \frac{bit\_rate}{8 \times picture\_rate}\}, \quad (9)$$

$$T_P = \max\{S_P + \frac{R_r - (N'_P S_P + N'_B S_B)}{N'_P + N'_B \frac{D'_B K_P}{D'_P K_B}}, \frac{bit\_rate}{8 \times picture\_rate}\}, \quad (10)$$

$$T_B = \max\{S_B + \frac{R_r - (N'_B S_B + N'_P S_P)}{N'_B + N'_P \frac{D'_P K_B}{D'_B K_P}}, \frac{bit\_rate}{8 \times picture\_rate}\}, \quad (11)$$

where T_I, T_P, and T_B denote the target numbers of bits for the next I, P, and B-picture in the GOP; S_I, S_P, and S_B are the numbers of bits generated from a last encoded picture; and D'_I, D'_P, and D'_B are the resulting distortion (MSE), respectively. K_P and K_B denote universal constants which are dependent on the quantization matrices; N'_P and N'_B are the numbers of remaining P-pictures and B-pictures in the GOP; and R_r represents the number of remaining bits for encoding in the GOP. *bit_rate* denotes the transmission channel rate, and $bit\_rate/(8 \times picture\_rate)$ is the minimum number of bits which should be assigned to each picture in the worst case.

Rate control in video coding follows three steps: target bit allocation, rate control, and adaptive quantization, which is the same procedure as in MPEG-2 TM5. At the beginning, target bit allocation step estimates the number of bits which is available for picture coding by using Eqs. (6-8). In general, this step assigns more bits to I-picture than P-picture, and P-picture than B-picture. When a picture is encoded, a value of distortion D'_I, D'_P, or D'_B is updated, and a new target number T_I, T_P, or T_B is obtained for the next picture coding. Then, the quantization parameter for a macroblock is obtained by using the rate control and adaptive quantization step as in TM5.

4. Experimental Results and Discussions

To show the feasibility of the proposed target bit allocation algorithm, we apply this algorithm to the TM5 by modifying its bit allocation scheme [1]. The simulation is carried out on five video sequences; 30 frames of Flower garden, Mobile & calendar, Cyclegirl, Table tennis, and Susie sequences, respectively. Each sequence has a resolution of 352 pixels × 240 lines and 30 frames per second. The sequences are encoded with a rate of 1.5 Mbps. A GOP consists of one intra frame, four predicted frames, and ten bidirectionally predicted frames. The motion vector has a half pixel resolution, and covers a search range of -8 to 7.5 pixels in both the horizontal and vertical directions. The buffer size of 25 × 16,385 bits is prepared and the initial buffer state is set to the value of 20/31 × 1,500,000/30 corresponding to the initial virtual buffer fullness of the intra frame in the TM5.

Coding performance of the proposed target bit allocation scheme is compared with the previous target bit allocation schemes [1, 5, 6] (see Table 1). The MPEG-2 TM5 is also tried for the reference. As shown in the table, the proposed scheme outperforms the existing methods in PSNR comparison. Fig. 1 shows the simulation graphs for the Flower garden and Table tennis sequences. The figure illustrates the PSNR and buffer occupancy of various schemes for these two video sequences. As expected, the proposed scheme is superior to the MPEG-2 TM5 scheme, and controls the buffer fullness well to prevent buffer overflow and underflow.

5. Conclusions

This paper presents an efficient and suboptimal rate control algorithm which improves the performance of hybrid DPCM/DCT video coders for a given bit rate constraint. In this algorithm, an input image is assumed to have a exponential rate-distortion characteristic. The target bit allocation scheme minimizes the average distortion in a GOP for a given bit rate. This algorithm is easily applicable to a real-time MPEG-2 encoding system with a reasonable extra computational burden. The experimental results show that the proposed scheme improves PSNR performance compared with the ordinary MPEG-2 rate control scheme.

6. Acknowledgment

This work was supported by MunWha Broadcasting Corporation (MBC), Korea.

References

[1] ISO / IEC JTC1 / SC29 / WG11 and ITV-TS SG 15 EG for ATM video coding, MPEG-2 Video Test Model 5, April 1993.

[2] J. Zdepski, D. Raychaudhuri, and K. Joseph, "Statistically based buffer control policies for constant rate transmission of compressed digital video," *IEEE Trans. Commun.* vol. 39, pp. 947-957, June 1991.

[3] J. Choi and D. Park, "A stable feedback control of the buffer state using the controlled Lagrange multiplier method," *IEEE Trans. Image Process.* vol. 3, pp. 546-557, Sept. 1994.

[4] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. Image Process.*, vol. 3, pp. 26-40, Jan. 1994.

[5] L. Wang, "Bit rate control for hybrid DPCM/DCT video codec," *IEEE Trans. Circuit Systems Video Technol.*, vol. 4, pp. 509-517, Oct. 1994.

[6] G. Keesman, I. Shah, and R. K.-Gunnewiek, "Bit-rate control for MPEG encoders," *Signal Process.: Image Commun.* vol. 6, pp. 545-560, Nov. 1995.

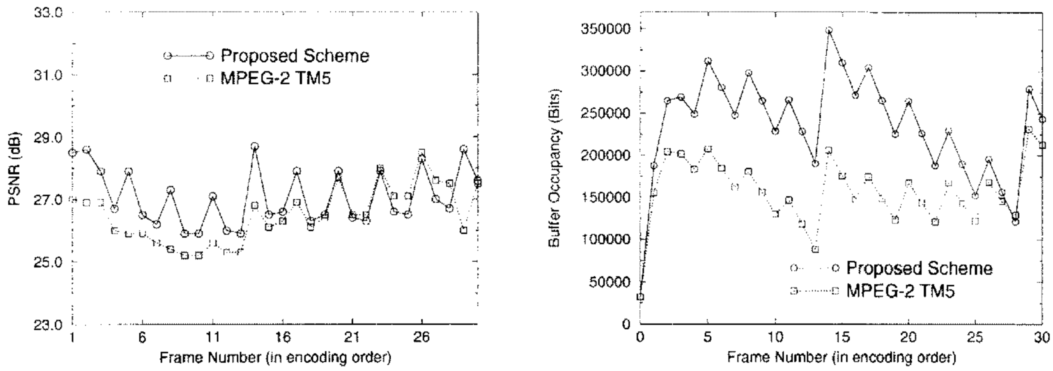
[7] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression,* Kluwer Academic Publishers, 1992.

[8] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operation Rec.*, vol. 11, pp. 399-417, 1963.

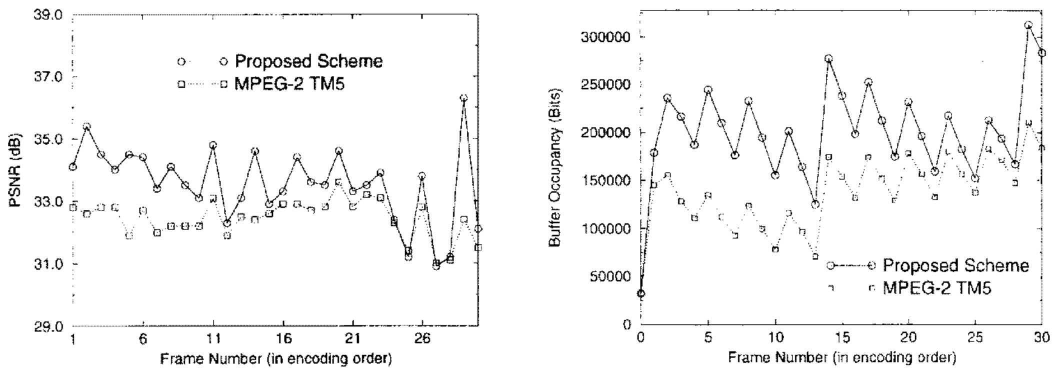
[9] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Process.*, vol. 2, pp. 160-175, Apr. 1993.

Algorithm	Flower	Mobile	Table	Cyclegirl	Susie
TM5 $K_P = K_B = 1.0$	27.23 dB	25.78 dB	31.46 dB	33.66 dB	40.00 dB
Wang's [5]	28.09 dB	26.57 dB	32.41 dB	34.37 dB	40.18 dB
Keesman's [6] $K_P = K_B = 1.0$	27.02 dB	25.55 dB	31.62 dB	33.60 dB	39.87 dB
Proposed $K_P = K_B = 1.0$	28.70 dB	27.09 dB	33.56 dB	34.71 dB	41.41 dB
TM5 $K_P = 1.0$ $K_B = 1.4$	28.01 dB	26.49 dB	32.44 dB	34.17 dB	40.59 dB

Table 1: Average PSNRs obtained from MPEG-2 coding by using various picture target bit allocation schemes.



(a)



(b)

Figure 1: Performance and buffer occupancy of MPEG coding by using two picture target bit allocation schemes for (a) *Mobile and calendar* and (b) *Table tennis* sequences.