Optimal Bit Allocation for MPEG-4 Video Coding

Seung-Hwan Kim and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
1 Oryong-Dong, Buk-Gu, Gwangju, 500-712, Korea
{kshkim, hoyo}@gist.ac.kr

ABSTRACT

In this paper, we propose an optimal bit allocation scheme for MPEG-4 video coding. In the proposed scheme, we assign more bits to I-frames and some selected P-frames. In order to decide optimal target bits for each intra frame, we need to calculate spatial complexity of the intra frame. We also define an absolute temporal correlation factor to measure temporal correlation between successive frames in the video sequence. Based on the temporal correlation factor, we can determine the position to insert a P-frame of high quality. Experimental results demonstrate that the proposed rate control scheme has significantly reduced the frame skipping, while increasing the average PSNR value by up to 0.9dB for various test sequences.

Keywords: MPEG, Bit Allocation, Temporal Correlation

1. INTRODUCTION

Bit allocation is a classic source coding problem, where a given number of bits are distributed among a finite set of quantizers to minimize the average distortion measure. It has been studied extensively and has recently found wide applications in MPEG video encoding [1], where bit allocation is more commonly referred to as rate control.

Rate control is one of the most important parts in video coding. We control the video bit rate to meet the target rate while pursing maximum video quality. The bit rate varies depending on its prediction mode, motion vector choices, etc. Once we decide the prediction mode and motion vectors, the quantization parameter (QP) controls the residual data, which is obtained by subtracting the motion compensated frame from the current frame [1].

In MPEG encoding, there are three different frame types: I-frame, P-frame, and B-frame. While the I-frame is the intra-coded frame that is coded by itself without any motion compensation, the P-frame is a predictive-coded frame that is coded with motion compensation from the previous anchor frame. The B-frame is an interpolative-coded frame that is coded with motion compensation from both the previous and the future anchor frames. In MPEG standards, the quantization step sizes are typically from zero to 31. Our goal is to find optimal target bits and quantizer choices for each frame.

In MPEG-4, the Q2 rate control algorithm is used as a reference algorithm for rate control [2]. The target bit rate is computed based on the bits available and the previous encoded frame bits while maintaining latency using a video buffering verifier. If the previous frame is complex and consumes excessive bits, more bits should be assigned to the current frame; however, if there are fewer bits left for encoding, only a fewer bits will be assigned to this frame because of the required bit rate. Once the target bit rate for the current frame is determined, we need to find the optimal QP value to meet the target bit rate.

The MPEG-4 Q2 rate control algorithm provides good buffer management because it decides target bits for each frame using the information, such as the previous encoded frame bits, buffer occupancy, and safety margin factor. However, it does not consider the prediction effect between frames. The most compression is obtained by the motion estimation (ME) and motion compensation (MC) processes in video coding. It means that picture quality of the I-frame has a significant effect on quality of the following frames. However, the I-frame of high quality may cause the buffer to overflow. Experimental results indicate that the average coding bits for the I-frame require up to 10 times of the bits for the P-frame [3, 4]. Therefore, we should assign more bits for the intra frame to enhance the coding efficiency. In order to avoid buffer overflow, we also decrease the target bits for non-intra frames by assigned extra bits for the I-frame of high quality.

Although the MPEG-4 Q2 rate control scheme inspects stability of the buffer management during the encoding operation, it does not consider the prediction effect among successive frames; hence, it may cause severe performance degradation in video coding. In this paper, we try to provide an optimum bit allocation scheme by considering the temporal correlation among successive frames. After we explain a picture-level rate control mechanism, we describe an optimum bit allocation strategy for the I-frame, reflecting the complexity of the I-frame, and we discuss a stable buffer control scheme to prevent buffer overflow and underflow.

2. MPEG-4 Q2 RATE CONTROL SCHEME

In this section, we describe the MPEG-4 Q2 frame-based rate control algorithm. There are four major stages in the MPEG-4 Q2 frame rate control algorithm [4].

1) Initialization
2) Estimation of the target bit rate before encoding
3) Determination of QP
4) Updating modeling parameter

At the first stage, we initialize several parameters, such as two coefficients of the quadratic model, quantization values for I- and P-frame, buffer size, buffer level, etc. Using these initialized parameters, the MPEG-4 Q2 rate control algorithm obtains the initial target bits T_i by

$$T_i = \begin{cases} \alpha \cdot R/(\alpha \cdot N_I + \beta \cdot N_B \cdot N_P), & \text{for I frame} \\ \beta \cdot R/(\alpha \cdot N_I + \beta \cdot N_B \cdot N_P), & \text{for B frame} \\ R/(\alpha \cdot N_I + \beta \cdot N_B \cdot N_P), & \text{for P frame} \end{cases} \quad (1)$$

where α and β are weighting factors for I-frame and B-frame. R represents the total remaining number of bits in the group of picture (GOP). N_I, N_B, and N_P represent the number of remaining I-frames, B-frames, and P-frames in the GOP, respectively.

In the second stage, we estimate the target bit rate before encoding. We determine the target bits for the current frame by three steps. At first, the initial target bit is estimated by

$$T_c = 0.95 \cdot T_i + 0.05 \cdot T_P \quad (2)$$

where T_P is the total bits used in the previous frame and T_c is the updated target bits. T_c is adjusted further based on the buffer occupancy a and buffer vacancy b. In addition, the target bit is guaranteed by the minimum bit rate $B/30$ from the following equation.

$$MAX\left(\frac{B}{30}, \frac{(a+2b)\cdot T_c}{2a+b}\right) \quad (3)$$

where B is the bit rate of the sequence.

When we determine QP, the final target bits obtained from the previous two stages are used to calculate the QP by

$$T_c = X_1 \cdot \frac{M}{\text{QP}} + X_2 \cdot \frac{M}{\text{QP}^2} \quad (4)$$

where M is the mean absolute difference (MAD) of the current frame after motion compensation, and X_1 and X_2 are quadratic model parameters.

In the final stage, quadratic model parameters are updated continuously based on encoding results of the current frame as well as the specified number of previous frames. After encoding each frame, the buffer level is updated by adding the target bits for the current frame.

3. OPTIMAL BIT ALLOCATION

3.1 Bit Allocation for I-frame

In video coding, we employ motion estimation (ME) and motion compensation (MC) operations to reduce temporal redundancy between frames. Since the major data reduction is occurred by ME/MC, the performance of ME /MC has a large impact on video coding efficiency.

In MPEG video coding standards, we encode the video sequence by the unit of GOP. We have at least one intra frame in each GOP and other frames are motion compensated from the intra frame directly or indirectly, i.e., the intra frame affects coding efficiency of all other frames in the same GOP through the ME/MC processes. Therefore, if the I-frame is reproduced at high quality, succeeding frames in the same GOP would have enhanced picture quality. Quality or distortion propagation is bounded to the same GOP.

The high-quality I-frame provides a good reference for better coding efficiency of succeeding frames. However, it may cause buffer-level surge and frame skipping during the encoding process. In order to solve this problem, we need to find an optimal target bits for the I-frame. In the MPEG-4 standard, the weighting factor α for the I-frame is assumed to be three; however, it is not always true [5]. From extensive experiments on various test sequences, we find that the weighting factor for the I-frame depends on the complexity of the intra frame. We also obtain the optimal target bits, T_{I_opt}, for the I-frame as

$$T_{I\_opt} = (1+\delta)\cdot T_{I\_Q2} \quad (5)$$

where T_{I_Q2} represents the target bits for the I-frame in the MPEG-4 Q2 rate control algorithm. The value of δ is a positive constant, which is represented as a function of the complexity of the intra frame and the temporal correlation between successive frames. However, since it is difficult to estimate the temporal correlation, we use only the spatial complexity of the intra frame. Fortunately, we find that the spatial complexity of the intra frame is inversely proportional to the temporal correlation. In general, the range of δ is from 0.15 to 0.35. For example, "Akiyo" and "News" sequences have the value of δ=0.35, and "Foreman" sequence has the value of δ=0.15.

Since MPEG video coding employs the block DCT transform, a good yet simple spatial complexity measure is the mean absolute value (MAV) of DCT coefficients, defined by [1]

$$MAV_{DCT} = \frac{1}{N \times M}\sum_{u=0}^{N-1}\sum_{v=0}^{M-1} ABS[F(u,v)] \quad (6)$$

where $F(u, v)$ is 8x8 block DCT coefficients of the entire I-frame. The relationship among Q_I, MAD_{DCT}, and target bits B can be described by [1]

$$Q_I = \frac{16.34}{B^{2.05}} \times MAV_{DCT}^{\ 1.0+0.29\times\ln(B)} \quad (7)$$

Since Eq. (7) provides a mechanism to control the number of data bits before actually encoding, it is very useful to determine the optimal quantization step size.

3.2 Adaptive Target Bit Control

In the proposed rate control scheme, we allocate more bits for the I-frame than the MPEG-4 Q2 rate control scheme, which may cause buffer overflow. In order to solve this problem, we design a stable buffer management method.

In the proposed rate control scheme, the target bits T_{GOP} in the GOP is

$$T_{GOP} = T_I + \sum_{n=1}^{N-1} T_p(n) = \alpha \cdot T_{GOP} + \sum_{n=1}^{N-1} T_p(n) \quad (8)$$

where T_I and $T_P(n)$ represent the target bits for I-frame, and P-frame, respectively. In Eq. (8), we define the target bits for the I-frame as a function of the target bits for the given GOP. In the proposed scheme, the average target bits T_{P_avg} for the P-frame is represented by

$$T_{P\_avg} = \frac{1}{N-1}(T_{GOP} - \alpha \cdot T_{GOP}) = \frac{(1-\alpha)}{N-1} T_{GOP} \quad (9)$$

The average target bits $T_{P\_avg\_Q2}$ for the P-frame in the MPEG-4 Q2 rate control algorithm is

$$T_{P\_avg\_Q2} = \frac{1}{N-1}(T_{GOP} - T_I) = \frac{1-k}{N-1} T_{GOP} \quad (10)$$

where k represents the weighting factor for the I-frame. In order to prevent buffer overflow, we need to decrease the amount of the target bits for P-frames. Therefore, the final target bits for P-frames are determined by

$$T_{target} = T_{target} \cdot \left(\frac{T_{P\_avg}}{T_{p\_avg\_Q2}}\right) = T_{target} \cdot \left(\frac{1-\alpha}{1-k}\right) \quad (11)$$

Using this adaptive target bit allocation, we can avoid buffer overflow. Figure 1 compares buffer occupancies of two rate control schemes for the “Akiyo” sequence. From Fig. 1, we observe that the proposed rate control scheme maintains a more stable buffer state than the MPEG-4 Q2 rate control algorithm.

3.3 Absolute Temporal Correlation

In the proposed rate control scheme, we select some P-frame and assign more bits to it to obtain high picture quality. We call it as a high-quality P-frame (HQP). The optimal bit allocation scheme for the P-frame consists of two parts. One is to select the optimal position to insert HQP and the other is to determine the amount of bits to encode HQP.

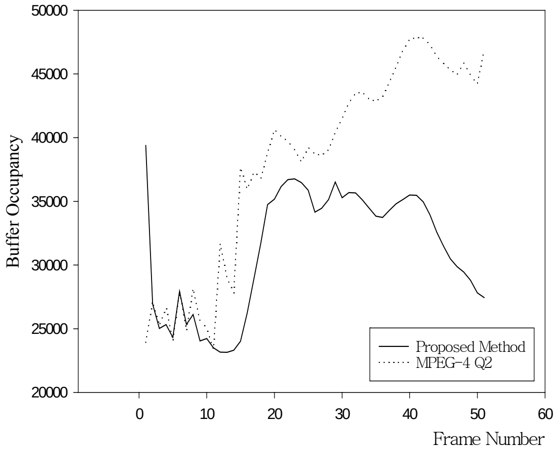

Fig. 1. Buffer Occupancies

In order to find the optimal position to insert HQP, we need to estimate temporal correlation between successive frames. If the estimated temporal correlation is high, HQP can provide a significant impact on picture quality of the following frames. Figure 2 represents the temporal correlation of a video sequence.

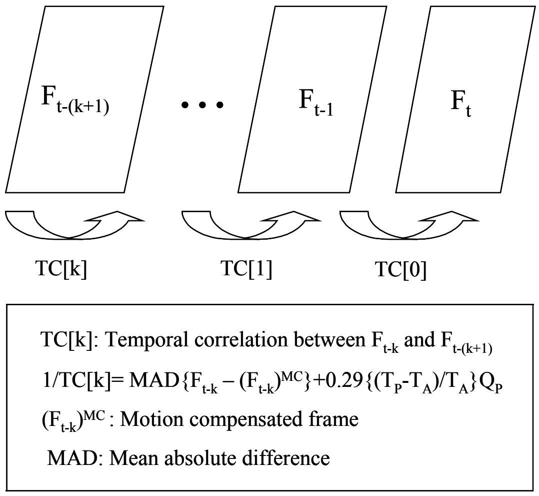

Fig. 2. Temporal Correlation

In order to apply the proposed rate control scheme, we need to find the temporal correlation between successive frames in the original video sequence. There are several measures to estimate the temporal correlation. The mean absolute difference (MAD) is a good example of them; however, MAD obtained during the encoding process only represents the relative temporal correlation because each frame is encoded with a different amount of bits.

In order to estimate the absolute temporal correlation between frames, we need to define a new measure considering the different amount of bits assigned to each frame. In this paper, we adopt the R-D cost function, used in the H.264/AVC [6]:

$$J_{MODE} = D + \lambda R \tag{12}$$

where D represents the distortion, λ = 0.85xQ_P^2, and R is the bit rate. From this point, we can replace D with MAD and R with used bits in the previous frame. When we use MAD, the value of λ is replaced by the square root of λ. Therefore, the absolute temporal correlation (*TC [k]*) between successive frames is represented by

$$1/TC[k] = MAD + 0.29 \cdot \{(T_P - T_A)/T_A\} \cdot Q_P \tag{13}$$

where T_P is the total number of coding bits used in the previous frame, T_A is the expected average bits in the frame, Q_P is the quantization step size in the previous frame. In Figure 3, we can observe the distribution of the inverse of the temporal correlation in “Foreman” sequence.

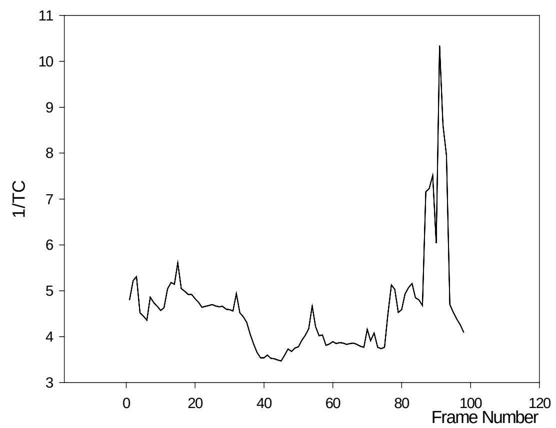

Fig. 3. Temporal Correlation for “Foreman”

3.4 Bit Allocation for P-frame

Using the distribution of the absolute temporal correlation, we can select the HQP position by

$$(TC[0] < T_{avg}) \text{ and } (TC[0] < TC[1] - \alpha \cdot T_{avg}) \tag{14}$$

where

$$T_{avg} = \frac{1}{N}\sum_{k=1}^{N} TC[k], \quad \alpha : \text{constant} \tag{15}$$

After finding the HQP insertion position, we calculate the weighting factor of the selected P-frame. The target bits for HQP are proportional to the absolute correlation factor because the correlation factor is the inverse of the slightly changed mean absolute difference (MAD). Hence, we can directly obtain their relationship between the correlation factor and the target bits by

$$T_{HQP} = \frac{TC[1]}{TC[0]} \cdot T_C \tag{16}$$

Figure 4 represents the framework of the proposed rate control scheme, which consists of two parts: one is to select the optimal target bits for the intra frame, and the other is to determine the position and target bits for HQP.

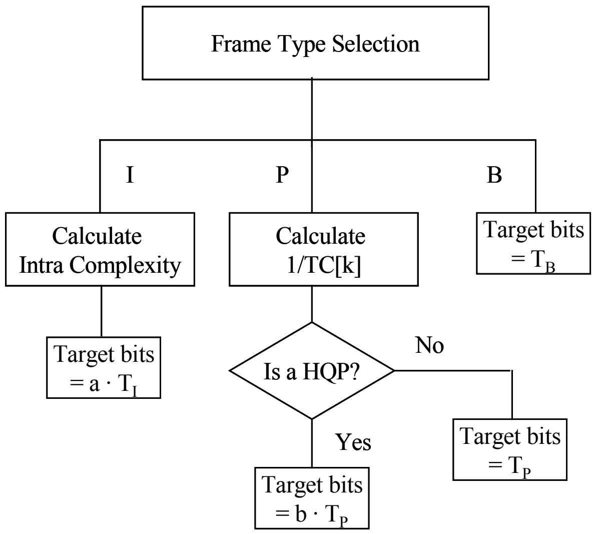

Fig. 4. Framework of The Proposed Scheme

In order to decide the optimal position to insert HQP, we measure the absolute temporal correlation between successive frames. If the absolute temporal correlation is high, HQP can efficiently affect the following frames. In Fig. 4, the weighting factors a and b are already obtained in the previous section.

4. EXPERIMENTAL RESULTS

In order to evaluate performance of the proposed scheme, we compare its results with those of the MPEG-4 Q2 rate control algorithm. We employ various MPEG-4 test sequences of the CIF format: “Foreman,” “Akiyo,” “Silent,” “News,” and “Mobile.”

Table 1 shows average PSNR values and the number of coding bits for “Akiyo” sequence. In Table 1, we observe that more bits are assigned to the I-frame in the proposed scheme; hence, we can obtain an I-frame of better quality. Succeeding P-frames are then motion compensated from the high-quality I-frame, and this process is repeated until the end of the current GOP. From Table 1, we recognize the better prediction effect on the succeeding frames in the proposed scheme. When the numbers of overall coding bits are the same in both rate control schemes, the proposed scheme provides higher average PSNR values than the MPEG-4 Q2 rate control algorithm.

Table 1. Average PSNRs and Coding Bits

	MPEG-4 Q2		Proposed Scheme		PSNR Gain
Frame	Used bits	PSNR	Used bits	PSNR	(dB)
1	19,280	32.05	24,616	33.89	+1.84
2	1,176	32.04	904	33.87	+1.83
3	1,128	32.18	472	33.85	+1.67
4	2,792	32.61	952	33.92	+1.31
5	3,584	32.98	5,240	34.66	+1.68
6	1,976	33.14	1,040	34.64	+1.5
7	8,568	34.30	4,192	35.02	+0.72
8	1,176	34.30	5,312	35.44	+1.14
9	2,232	34.44	1,048	35.38	+0.94
10	4,656	34.78	6,840	35.90	+1.12
11	936	34.75	872	35.82	+1.07
12	6,136	35.18	1,256	35.79	+0.61
13	904	35.08	3,784	36.02	+0.94
14	1,392	35.17	1,240	35.95	+0.78

Table 2 compares the average PSNR values with the proposed and the MPEG-4 Q2 rate control algorithms. The proposed scheme improves coding efficiency over the MPEG-4 Q2 scheme. For "News" sequence, we have improved the average PSNR values by up to 0.96 dB. In Table 2, we observe that video sequences, which have high temporal correlation between frames, provide better performance improvement with the proposed scheme.

Table 2. Average PSNR Values for Test Sequences

Sequence	PSNR (dB)		
	MPEG-4 Q2	Proposed Scheme	Gain
Akiyo	37.69	38.11	+0.42
News	32.39	33.35	+0.96
Silent	33.57	34.02	+0.45
Foreman	32.65	32.77	+0.12
Mobile	25.42	25.59	+0.17

Table 3 shows the numbers of frame skipping in both schemes. All the sequences are coded at 30 fps; however, each sequence is encoded at different bit rates. While "Akiyo" and "News" are coded at 120 kbps, "Silent," "Foreman," and "Mobile" are coded at 196 kbps, 210 kbps, and 384 kbps, respectively.

Table 3. Comparison of the Number of Frame Skipping

Sequence	Number of Frame Skipping		
	MPEG-4 Q2	Proposed Scheme	Gain
Akiyo	15	0	+15
News	8	0	+8
Silent	7	0	+7
Foreman	18	16	+2
Mobile	10	11	-1

We compare average PSNR values at various bit rates in Fig. 5 to Fig. 9, where the solid line indicates results by the MPEG-4 Q2 rate control scheme and the dotted line indicates results by the proposed scheme. Since "Akiyo" and "News" have high temporal correlation between successive frames, the proposed rate control scheme provides better performance. However, "Foreman" has small correlation between frames, limiting performance improvement.

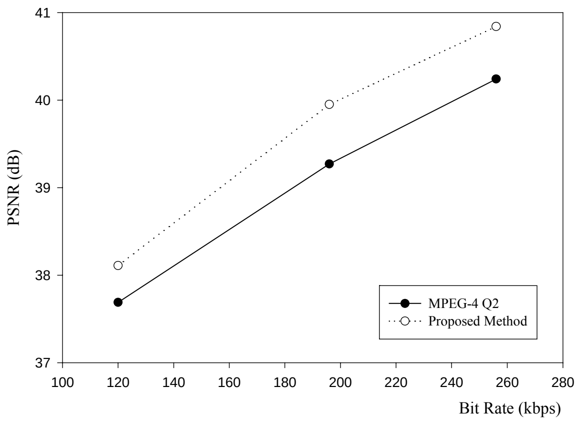

Fig. 5. PSNR Values for "Akiyo"

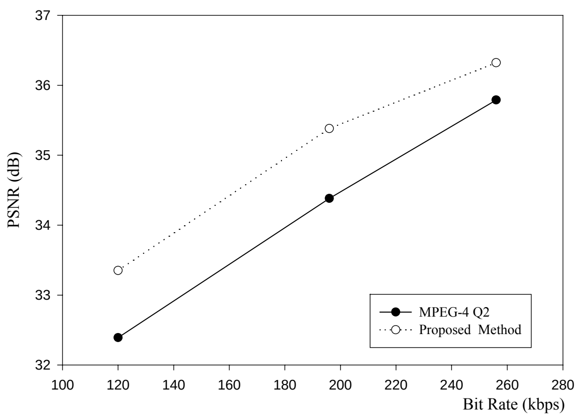

Fig. 6. PSNR Values for "News"

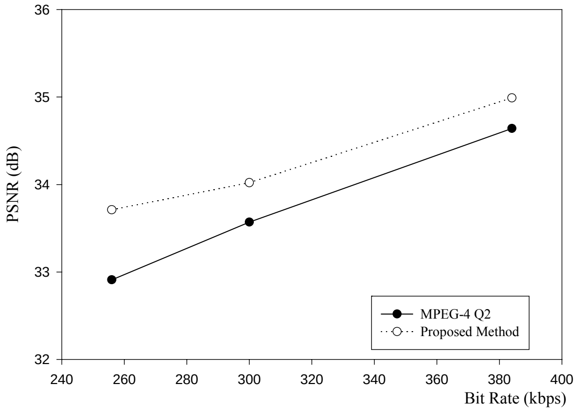

Fig. 7. PSNR Values for "Silent"

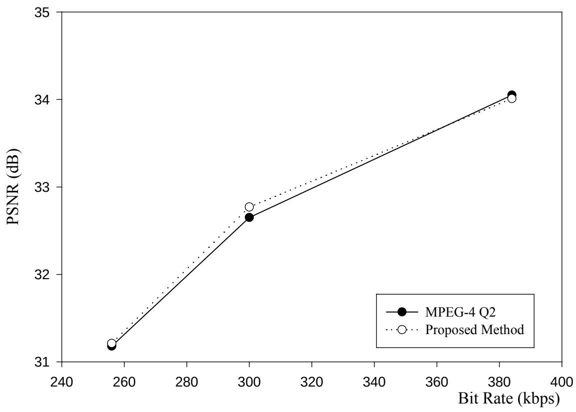

Fig. 8. PSNR Values for "Foreman"

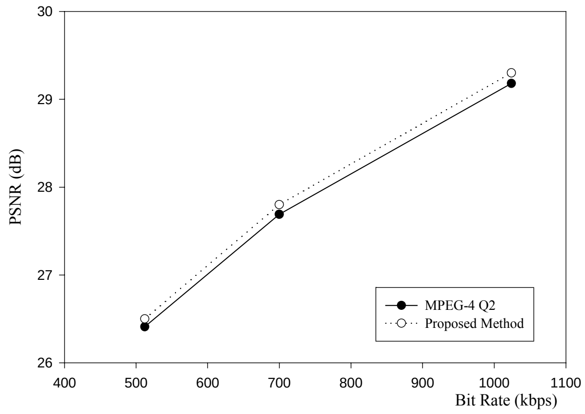

Fig. 9. PSNR Values for "Mobile"

4. CONCLUSIONS

In this paper, we propose a new bit rate control scheme for the MPEG-4 video coding algorithm. In the proposed scheme, we calculate the optimal number of coding bits for the intra frame considering the spatial complexity of the I-frame. We estimate the absolute temporal correlation between successive frames to find the best position and the optimal number of coding bits for P-frames. Although we have assigned more coding bits to intra frames, we can avoid buffer overflow by an adaptive buffer control strategy. Experimental results demonstrate that the proposed rate control scheme increases the average PSNR values by up to 1 dB.

ACKNOWLEDGMENTS

This work was supported in part by Gwangju Institute of Science and Technology (GIST), in part by the Ministry of Information and Communication (MIC) through the Realistic Broadcasting Research Center (RBRC) at GIST, and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

REFERENCES

[1] F. Pan, Z. Li, K. Lim, and G. Feng, "A Study of MPEG-4 Rate Control Scheme and Its Implementation," IEEE Trans. Circuit and System for Video Technology, vol. 13, pp. 440-446, May 2003.

[2] H.M. Kim, "Adaptive Rate Control Using Nonlinear Regression," IEEE Trans. Circuit and System for Video Technology, vol. 13, pp. 432-439, May 2003.

[3] B. Tao, B.W. Dickinson, and H.A. Peterson, "Adaptive Model Driven Bit Allocation for MPEG Video Coding," IEEE Trans. Circuit and System for Video Technology, vol. 10, pp. 147-157, Feb. 2000.

[4] J. Lee and B.W. Dickinson, "Rate-Distortion Optimized Frame Type Selection for MPEG Encoding," IEEE Trans. Circuit and System for Video Technology, vol. 7, pp. 501-510, June 1997.

[5] T. Chiang and Y.Q. Zhang, "A New Rate Control Scheme Using Quadratic Rate Distortion Model," IEEE Trans. Circuit and System for Video Technology, vol. 7, pp. 247-250, Feb. 1997.

[6] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G.J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," IEEE Trans. Circuit and System for Video Technology, vol. 13, pp. 688-703, July 2003.