OPTIMAL BIT ALLOCATION STRATEGY BASED ON SPATIAL COMPLEXITY AND TEMPORAL CORRELATION

Seung-Hwan Kim, Geun-Yong Kim, Jin Heo, and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
1 Oryong-dong, Buk-gu, Gwangju, 500-712, Korea
phone: + 82-062-970-2258, FAX: + 82-062-970-2247
Email: {kshkim, hoyo}@gist.ac.kr

ABSTRACT

In this paper, we propose a bit allocation strategy for video coding algorithms, such as MPEG-4 and H.264/AVC. Considering the target bit rate and the spatial complexity of the intra frame, we determine the optimal number of target bits for the intra frame. In order to design an optimal bit allocation strategy for the inter frame, we consider that each inter frame in the group of pictures (GOP) has different importance in terms of motion estimation and motion compensation processes. We analyze the importance of each inter frame according to its position. Experimental results show that the proposed rate control algorithm has reduced the frame skipping rate significantly, while increasing the average PSNR value by up to 1dB.

1. INTRODUCTION

Rate control is one of the most important parts in video coding. We control the video bit rate to meet the given target bit rate while pursing maximum video quality. The bit rate changes according to the frame type, the prediction mode, and the motion vector accuracy. Once we decide the prediction mode and motion vectors, the quantization parameter (QP) controls coding of the residual data, which are obtained by subtracting the motion compensated frame from the original frame [1].

In MPEG, motion estimation and motion compensation (ME/MC) processes play an important role in compressing video data. Hence, the intra frame gives significant effect on the picture quality of succeeding frames in the ME/MC processes [2]. The pre-located inter frames in GOP are also give significant effect on the coding efficiency.

The MPEG-4 VM5 Q2 rate control algorithm has been considered as a good frame-based rate control algorithm and provides a stable buffer management [3]. However, since it does not consider the ME/MC process during bit allocation, it may cause severe performance degradation. There is no reasonable method to determine QP of the initial frame in the MPEG-4 VM5 Q2 algorithm.

In this paper, we discuss an efficient bit allocation strategy for the intra frame, reflecting the spatial complexity of the intra frame. We also describe the optimum bit allocation method for the inter frame and determine the importance of each inter frame according to its poison in GOP.

2. MPEG-4 Q2 RATE CONTROL ALGORITHM

In the MPEG-4 Q2 frame rate control algorithm [4], there are four main operations:

a) Initialization
b) Estimation of the target bit rate before encoding
c) Determination of QP
d) Updating modelling parameters

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

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

where α and β are weighting factors for each frame. R represents the total number of remaining bits in GOP. N_I, N_B, and N_P represent the number of remaining frames in GOP.

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 bits are estimated by

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

where T_P is the total number of coding 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 the buffer vacancy b, which is (1-a). 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. The final target bits T_c obtained from the previous two stages are used to calculate QP by

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

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.

3. OPTIMAL BIT ALLOCATION

3.1 Bit Allocation for the Intra Frame

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 [5] [6]. Hence, the intra frame affects coding efficiency of all other frames in the same GOP.

Table 1 shows the importance of selecting the initial QP value of the intra frame. We have used 10 and 15 as the initial QP value in our experiment. The overall coding bits for both cases are similar, but have different coding efficiencies

Table 1. Initial QP Value and Average PSNR Value

	Initial QP=10		Initial QP=15		PSNR Gain
Frame	Coding bits	PSNR	Coding 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

In order to find the optimum QP value, we need to find the optimal target bits for the intra frame from the given target bits for GOP. From extensive experiments on various test sequences, we have found that the optimal target bits for the intra frame are inversely proportional to the spatial complexity of the intra frame [7]. The purpose of measuring the spatial complexity of the intra frame is to estimate coding bits for the intra frame with the given quantization parameter. By allocating bits to the intra frame according to its spatial complexity, the variation of buffer occupancy can be controlled efficiently [1].

In order to estimate the number of coding bits, we count several low frequency components since most high frequency components are removed in the quantization process. Therefore, we propose a new measure for spatial complexity, called partial mean absolute value of DCT ($PMAV_{DCT}$).

$$PMAV_{DCT} = \frac{1}{N \times M} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} ABS[F(u,v)] \qquad (5)$$

where M and N represent constant values to count several low frequency components. From Table 2, we observe that $PMAV_{DCT}$ estimates the coding bits more accurately than MAV_{DCT} [7].

Table 2. Spatial Complexity and Coding Bits

Sequence	MAV_{DCT}	$PMAV_{DCT}$	QP	Coding Bits
Akiyo	15.16	11.23	15	32256
News	16.15	20.02	15	48424
Silent	22.55	16.95	15	44848
Foreman	26.18	16.43	15	43120
Mobile	34.91	36.70	15	129720

Using $PMAV_{DCT}$, we find the optimum target bits (T) for the intra frame as:

$$T = B / \sqrt{PMAV_{DCT}} \qquad (6)$$

where B is the given bit rate of the sequence. From the optimal target bit rate for the intra frame, we can calculate the quantization value. The relationship among Q_I, $PMAD_{DCT}$, and optimum target bits T can be described by [1]

$$Q_I = \frac{16.34}{T^{2.05}} \times PMAV_{DCT}^{\,1.0+0.29\times\ln(T)} \qquad (7)$$

Since Eq. (7) provides a mechanism to control the number of data bits before actual encoding, it is very useful to determine the optimal quantization step size. Figure 1 shows the resulting PSNR values over various intra QP values. Using Eq. (6) and Eq. (7), we obtained the optimum quantization values 7, 9 and 10 for the sequence “Foreman,” “News,” and “Silent”, respectively.

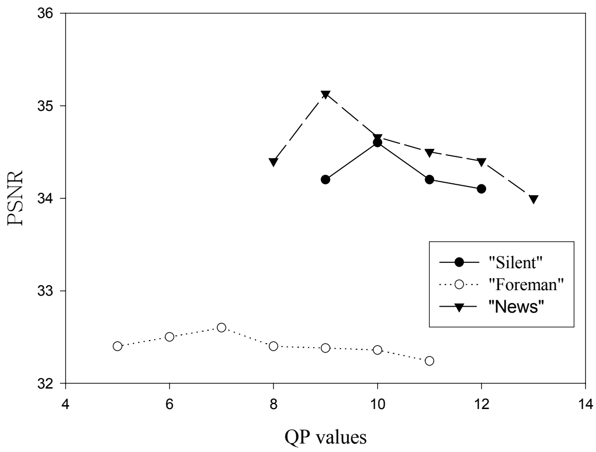

Fig. 1 Adaptive Intra Quantization

3.2 Bit Allocation for the Inter Frames

3.2.1 Various Weighting Functions

An adaptive method for inter frame bit allocation uses a linear weighting factor adjustable to its distance from the reference intra frame [1]. Let N_{SEG} be the GOP length and σ be an adjustable parameter. The adjustable bit $\Delta B_p(n)$ for the P-frame is defined by

$$\Delta B_p(n) = \sigma \frac{N_{SEG} - 2n}{N_{SEG} - 2}, \qquad n = 1, \ldots N_{SEG} - 1 \qquad (8)$$

For the first P-frame, $\Delta B_p(1) = \sigma$, which is the extra number of bits added to the average bits. For the last P-frame,

$\triangle B_p(N_{SEG}-1) = -\sigma$, while for the other P-frames, $\triangle B_p(n)$ changes from $+\sigma$ to $-\sigma$. In order to consider the effect of the ME/MC processes, we designed two weighting functions (sigmoid and cubic), which assign more bits to the pre-located P-frames compared to the linear weighting function. The sigmoid function is defined by

$$\Delta B_p(n) = \sigma\left(\frac{-1}{1+\alpha\exp(-\beta(n-N_{SEG}/2))}+0.5\right) \qquad (9)$$

In the sigmoid function, we control the shape of the function by two parameters α and β. Since the sigmoid function is odd symmetric, the symmetric point is determined by α, and the slope of the function is determined by β. The cubic function is defined by

$$\Delta B_p(n) = -\sigma(n-N_{SEG}/2)^3, \qquad n = 1,\ldots N_{SEG}-1 \qquad (10)$$

In the cubic function, most additional bits are allocated to a few pre-located frames in GOP. From the extensive experiments, we obtained the optimal parameter values listed in Table 3.

Table 3. Optimal Parameter Values

Sequence	Parameters		
	Linear	Sigmoid	Cubic
Akiyo	$\sigma = 0.47$	$\beta = 0.1$	$\sigma = 1.85$
Foreman	$\sigma = 0.18$	$\beta = 0.074$	$\sigma = 1.10$
Carhone	$\sigma = 0.345$	$\beta = 0.065$	$\sigma = 1.05$
Mobile	$\sigma = 0.1$	$\beta = 0.016$	$\sigma = 0.29$

Table 4. PSNR Values in Various Weighting Functions

Sequence	PSNR (dB)			
	Q2	Linear	Sigmoid	Cubic
Akiyo	38.24	38.83	38.82	39.03
Foreman	35.01	35.08	35.07	35.09
Carphone	36.34	36.34	36.33	36.35
Mobile	25.62	25.61	25.60	25.61

3.2.2 Analytical Modelling

Since we do not know the optimum value of σ and β, the above three weighting functions cannot provide the generalized method. Those values also depend on the sequences. In order to design the importance of each inter frame considering its position in GOP, we propose a new analytical model.

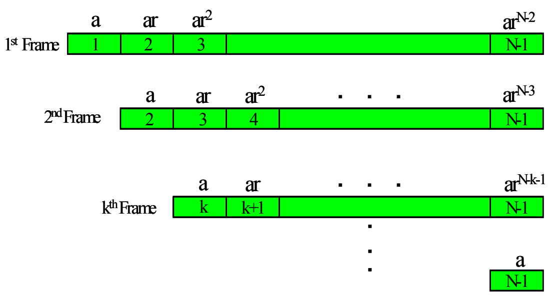

Fig. 2 The Importance of the Inter Frame

In Figure 2, N is the length of GOP, r $(0<r<1)$ is the average temporal correlation factor between frames, and a represents the picture quality of the given frame. If we have the first inter frame with quality of a, the overall video quality from the ME/MC processes is expressed by

$$S_1 = a + ar + ar^2 + \cdots + ar^{N-2} = \frac{a(1-r^{N-1})}{1-r} \qquad (11)$$

Hence, the overall video quality of the k-th inter frame is represented by

$$S_k = a + ar + ar^2 + \cdots + ar^{N-k-1} = \frac{a(1-r^{N-k})}{1-r} \qquad (12)$$

The overall video qualities are then represented by

$$T = S_1 + S_2 + S_3 + \cdots + S_{N-1} = \frac{a}{1-r}\sum_{n=1}^{N-1}(1-r^n) \qquad (13)$$

As a result, we can determine the importance of each inter frames to optimise the ME/MC processes. The equation (14) represents the normalized importance of each inter frame according to its position in GOP.

$$W_k = \frac{S_k}{T} = \frac{1-r^{N-k}}{\sum_{n=1}^{N-1}(1-r^n)} \qquad (14)$$

In "Foreman" and "Mobile" sequences, about 80 percents of data are coded in the ME/MC processes. Hence the value of r is 0.8 in those cases. In "Akiyo" and "News" sequences, about 90 percents of data are coded in the ME/MC processes. For simplicity we fix the value of r by 0.85.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed algorithm, we compare our results with those of the MPEG-4 Q2 algorithm. We employed various test sequences of the CIF format (352 x 288 pixels). The frame rate is fixed at 30 fps.

Table 5 compares the average PSNR values with the proposed and the MPEG-4 Q2 algorithms. For "News" sequence, we have improved the average PSNR values by up to 0.96 dB.

In Table 6, we compare the number of frame skipping. We have also implemented the proposed algorithm with the H.264/AVC reference software version 8.2 and PSNR values are compared in Table 7. Since H.264/AVC determine the QP value before transform coding, we cannot measure the spatial complexity of the intra QP. Hence, we implemented only the bit allocation algorithm for the inter frame.

Table 5. Average PSNR Values for Test Sequences

Sequence	PSNR (dB)		
	MPEG-4 Q2	Proposed	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 6. Comparison of the Number of Frame Skipping

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

Table 7. Average PSNR Values for Test Sequences

Sequence	PSNR (dB)		
	H.264/AVC	Proposed	Gain
Akiyo	38.77	39.20	+0.43
Foreman	36.58	36.78	+0.2

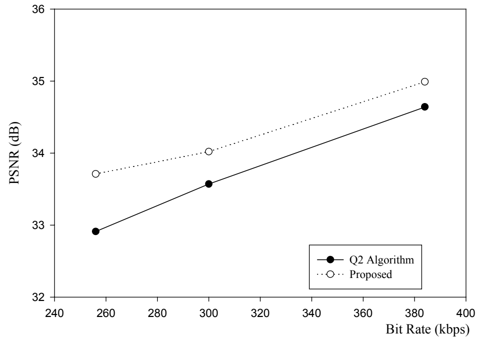

Fig. 3. PSNR Values for "Silent"

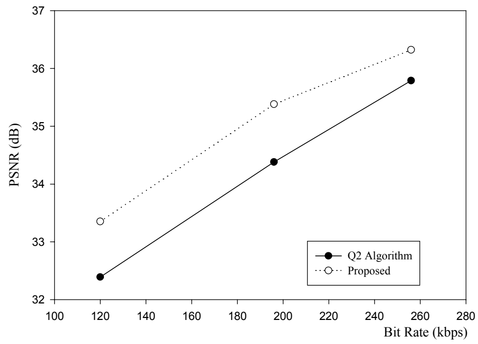

Fig. 4. PSNR Values for "News"

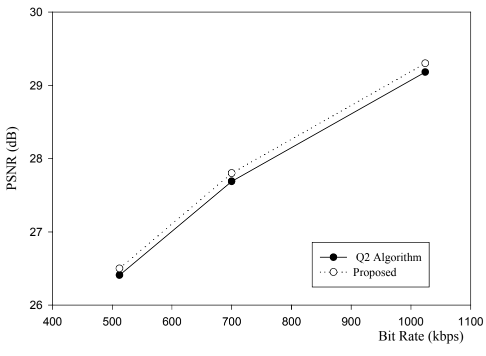

Fig. 5. PSNR Values for "Mobile"

5. CONCLUSIONS

In this paper, we propose an optimum bit allocation algorithm for video coding. In the proposed algorithm, we calculate the optimal number of coding bits for the intra frame considering the spatial complexity of the intra frame and target bit rate for the sequence. In order to find optimal bit allocation strategy for the inter frame, we design a new analytical model, which represents the importance of each inter frame according to its position in GOP. We implemented the proposed method in MPEG-4 part 2 and H.264/AVC. Experimental results demonstrate that the proposed algorithm increases the average PSNR values by up to 1 dB.

ACKNOWLEDGMENTS

This work was supported in part by MIC through RBRC at GIST, in part by MOE through BK21, and in part by DCRC.

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.

[7] S. H. Kim and Y. S. Ho, "Optimal Bit Allocation for MPEG-4 Video Coding," *Proc*. Picture Coding Symposium (PCS), Dec. 2004.