A Simple Adaptive Rate Control for H.264/AVC

Myoung-Jin Kim, Min-Cheol Hong

(School of Electronic Engineering, Soongsil University, Seoul, Korea)

Abstract— **The purpose of this paper is to improve allocation of the number of bits without skipping the frame by accurately estimating the target bits in H.264/AVC rate control. In our scheme, we propose an enhancement method of the target frame rate based H.264/AVC bit allocation. The enhancement uses a frame complexity estimation to improve the existing mean absolute difference (MAD) complexity measurement. Bit allocation to each frame is not just computed by target frame rate but also adjusted by a combined frame complexity measure. Using the statistical characteristic, we obtain change of occurrence bit about QP to apply the bit amount by QP from the video characteristic and applied in the estimated bit amount of the current frame. Simulation results show that the proposed rate control scheme could not only achieves time saving of more than 99% over existing rate control algorithm, but also increase the average PSNR of reconstructed video for around 0.02~0.78 dB in all the sequences.**

***Keywords*— H.264, Computational Complexity, Mean Absolute Difference, Rate Control, Peak Signal-to-Noise Ration**

Manuscript Number: 1674-8042(2010)supp.-0151-04

dio: 10.3969/j.issn1674-8042.2010.supp..41

1 Introduction

H.264/AVC is the latest international video coding standard developed by the Joint Video Team (JVT) of ISO Motion Picture Expert Group (MPEG) and ITU-T Video Coding Expert Group (VCEG). [1–5]. This is mainly intended for video transmission in all areas where bandwidth or storage capacity is limited (e.g. video telephony, video conferencing over mobile channels, and other such services). Many applications using video transmission are affected by time-varying bandwidth channels. Thus, we need to control bit rate algorithms to allow modifying coding parameters according to the channel's variations. Many rate control schemes have been proposed in previous works [6-8]. However, they are difficult to apply directly to H.264 rate control. The other schemes can only supply the needed data after encoding the current frame to determine the appropriate QP. It does not comply with the H.264 RDO procedure. M. Jiang *et al* in [9] have proposed a peak signal to noise ratio (PSNR)-based frame complexity measure to improve the existing MAD-based complexity measure. A normalized MAD as a frame complexity measure is also proposed [10]. These schemes use the quadratic R-D model to compute a QP with an estimated target-bit and an estimated MAD [11]. The estimated MAD is different from the actual computed MAD in the scene transition frame. Thus, an inexact QP is calculated because of the extremely low correlation between the current frame and the previous frames. Although the schemes mentioned above [9-10] improve the quality of video, an inaccurate MAD is still used to obtain the QP for the current frame, and additional computations are required in the pre-analysis. The large computational complexity deters its application in real-time video transmission. Ribas-Corbera and Lei [12] proposed an optimized method to assign target bits to each frame according to frame complexity, which is measured by frame energy. Frames with higher complexity can have more bits, and frames with lower complexity have fewer bits.

To resolve the additional computation problem, we propose a simple and enhanced frame-layer rate control scheme for frame bit allocation by considering both buffer status and frame complexity. We took real-time rate control into consideration to obtain an appropriate QP for the characteristics of inter coding. Then we estimated the frame complexity using the statistical data gathered after encoding each frame to improve the existing MAD-based complexity measure.

Received: 2010-07-20

Project supported: This work was supported by Seoul R&BD Program (10544). In addition, this research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2010-(C1090-1011-0003))

Corresponding author: Myoung-Jin Kim (webzealer@ssu.ac.kr), Min-Cheol Hong (mhong@ssu.ac.kr)

Simulation results show that our proposed method achieves better rate control for inter-coded frames without the degrading the coding performance. The rest of this paper covers the following: Section 2 describes our proposed frame-estimation scheme; Section 3 discusses the results, and Section 4 presents a conclusion.

2 Proposed Rate Control Scheme

2.1 Estimate Target Bits for Current Frame

Similar to earlier standards, H.264/AVC exploits the spatial, temporal and statistical redundancies in the sequence. As the level of redundancy changes from frame to frame, the number of bits generated per frame is also variable. In general, the rate control scheme has been treated in frame layer level and/or in the MB layer level based on fluid traffic model and linear model. To estimate target bits for the current frame, we employ a fluid traffic model based on the linear tracking theory. In this paper, we assume one GOP consisting of first I-frame and subsequent P-frames. Let N denote the total number of frames in GOP**,** n denote the n^{th} frame in the sequence, and $B_c(n)$ denote the occupancy of the virtual buffer after coding the n^{th} frame. The buffer occupancy is updated after coding each frame as:

$$B_c(n) = A(n-1) - \frac{u}{F_r} + 0.5 \quad (1)$$

where $A(n\text{-}1)$ is the number of bits generated by the $(n\text{-}1)^{th}$ frame. We first define a target buffer level, $TBL(n)$, for each P frame as in equation (2), where N_p is the total number of P-frame remaining for encoding, N_{cp} is the number of P-frame coded in the GOP. T_{buf} denote the target bits computed based on the target buffer level, the frame rate, the available channel bandwidth and actual buffer occupancy, which is computed using equation (3). where B_{LB} and B_{UB} are denote the limit of the buffer, as a buffer lower bound and buffer upper bound.

$$TBL(n) = \begin{cases} TBL(n-1) - \dfrac{B_C}{N_P - N_{CP}} & , N_{CP} > 1 \\ B_C - \dfrac{B_C}{N_P - N_{CP}} & , N_{CP} = 0 \end{cases} \quad (2)$$

$$T_{buf} = \frac{u}{F_r} + \gamma \times (TBL(n) - B_C(n))$$
$$T_{buf} = \min\left(B_{UB}, \max(B_{LB}, T_{buf})\right) \quad (3)$$

In this mathematical statement, γ is considered a constant and its typical value is 0.75 but we set the default value at 0.8 to achieve a tight buffer regulation. Meanwhile, the number of remaining bits should also be considered when the target bit is computed as follows

$$T_r = \left(\beta \times \frac{B_C}{N_P - N_{CP}}\right) + \left((1-\beta) \times T_{buf} + 0.5\right) \quad (4)$$
$$T(n) = (\alpha \times W_P) + \left((1-\alpha) \times T_r + 0.5\right)$$

where W_p is the average complexity weight of P-frame, T_r is the number of buffer for each P-frame to encode a frame. β is a meaning of dependence on buffer occupancy and target buffer level as a weighting factor with typical value 0.5. α is a weighting factor with value 0.5.

2.2 Compute QP and Performing QP Adjustment

For a given frame, rate control determines a QP to achieve the frame target bits. To determine the frame QP based on statistical information, we introduce a reference table derived from extensive experiments using various test sequences. The computed average bits of five CIF sequences (slow and smooth sequence “Container”, “News”, normal sequence “Foreman”, fast and detail sequence “Mobile”, “Stefan”) are reported in Table I. The average bits of the P-frame used in the experiment (as shown in Tab. 1), measures the QP and the required bits and it can be derived from equation (6), where $QP_{bits,n}$ shows the estimated number of bits based on QP index n. It was calculated only once but was updated after encoding each frame.

$$QP_{bits,n} = \alpha \times e^{(\beta \times (QP_{n-1}+1))}, \quad (1 \le n \le 51) \quad (5)$$

Using the table, the parameters of the equation (5) can be calculated by approximation. In our work, α and β are derived from Tab. 1, based on statistical data that were considered as a constant values.

Table. 1 Average Bits of P-frames by QP

QP	Cont.	Foreman	Mobile	News	Stefan	QP Range
…						
19	36,361	59,772	155,991	21,704	140,248	77,201
20	28,845	49,106	137,400	18,580	124,006	67,587
21	21,139	42,086	123,580	16,430	111,703	59,867
22	20,027	36,069	110,859	14,542	99,234	52,376
23	16,023	30,299	96,701	12,610	87,396	45,247
24	12,896	25,432	84,335	10,993	75,787	39,617
25	10,634	22,152	75,887	9,831	68,227	34,368
26	8,371	18,285	64,033	8,398	57,865	29,263
...						

In Tab. 1, the QP range (QR) is the range of the number of bits based on the QP index. It can be allocated for encoding the current frame, and is updated by actual bits generated from the previous frames.

$$QP_{bits,n} = QP_{bits,(n+1)} + \frac{\left|QP_{bits,n} - QP_{bits,(n+1)}\right|}{2} \quad (6)$$

Since the bits as a function of QP index take Gaussian distribution, equation (5) can be updated using equation (6). According to the QR, the number of bits in the $QP_{bits,n(0\sim51)}$ is estimated. Using equation (4) and (5), the QP of the current frame (QP_c) can be computed by

$$QP_c = choiceQP(QP_{bits,n(1\sim51)}, T), (1 \le n \le 51) \quad (7)$$

where T is the number of target bits estimated in equation (4), *choiceQP* is a function of finding T from the $QP_{bits,n(1\sim51)}$, it is decided by iterative loop. To maintain the smoothness of visual quality among successive frames, the computed QP_c is limited to a certain range. In our scheme, a limit is set for the QP for encoding the current frame using this equation:

$$QP_c = \min\{QP_p + \Delta QP, \max\{\mathrm{QP}_p - \Delta QP, QP_c\}\} \quad (8)$$

where QP_p is the QP value of previous frame, the increment or decrement of ΔQP is set at ±2.

2.3 Frame Complexity Measure

To get a better target bit estimation and accurate QP, we need to consider the statistical information of the sequence characteristic. The current frame, according to sequence characteristic, is in close correlation to the adjacent frames. Therefore, we use two parameters, which are consisted of weighted combination of two values: 1) the number of bits generated from the previous frame; 2) the number of bits by scaling the average bits from the reference twenty frames. For frame-level rate control, the target bits for each frame are first determined adaptively according to the frame complexity. To estimate the current frame complexity, we use these parameters above. To estimate the number of target bits of P-frame, the complexity weight of P-frame, W_p is computed by

$$W_p = (\lambda \times A(n-1) + (1-\lambda) \times S_{bits}) \times QP_p + 0.5 \quad (9)$$

where W_p is updated after encoding a frame, and is reflected in equation (4). S_{bits} is the average bits computed with the same QP value from the reference frames. λ is a weighting factor and its value is set to 0.67.

3 Experimental Result

The proposed rate control algorithm is tested for various video sequences. All test sequence is encoded with only one I-frame of the first frame followed by P-frames. As a reference for comparisons, the rate control based on PSNR-based frame complexity [9] and the H.264/AVC rate control algorithm were selected [13]. We employed test sequences of the QCIF (176×144 pixels) 4:2:0. The frame rate is fixed at 30 fps, a total of 300 frames were coded without skipping the frames. The H.264 encoder was configured to have one reference frames for inter motion search, (1/4)-pel motion vector resolution, rate-distortion optimized mode decisions, and full search motion estimation with a search range of 16. More results are reported in Table II, this table compares the average PSNR values and average encoding time with the proposed, Ref.9, and the JM.

Tab. 2 indicates that Ref.9 scheme achieved average PSNR gain with similar or lower PSNR deviation as compared to the JM12.1, but it causes a waste of time because of the additional PSNR computation needed in the encoding of the data. Also, JM12.1 and Ref.9 produced an excess of bit quantity in the all sequences, and large computational complexity deterred the application in real-time video transmission.

However, our proposed rate control effectively allocated bit quantity to the target bit rate and achieved time saving of about 99% when compared to the reference [9] and [13]. Furthermore, our scheme has about 0.02~0.78 dB gains in the average PSNR, and we achieved lower PSNR deviation.

In this experiment, CIF (352×144 pixels) also show that the PSNR gain and time saving are pretty much the same. Figure 1 shows the comparison of PSNR against frame number in “Table QCIF” by using JM12.1 scheme, reference [9] scheme, and proposed control scheme. It can be shown that the PSNR fluctuation has been reduced greatly. A good rate control results in higher video quality, lower fluctuation, and a lower mismatch between the target bit rate and the encoded bit rate.

Overall, our scheme shows a much steadier visual quality without wasting encoding time. Better video quality is the result of our QP adjustments and simple frame complexity using encoded statistical information.

4 Conclusions

In this paper, we have presented an efficient real-time rate control scheme without skipping frames. We have effectively allocated the number of bits for H.264/AVC video encoding. Our new and simple frame complexity measurement was developed to enhance the existing MAD-based method and was applied to our bit allocation for real-time rate control. QP accuracy is very important to prevent the overflow or underflow to a target channel that has a low bandwidth. Therefore, we have presented a QP control scheme to adjust the computed QP based mainly on the actual encoding results of previously-coded frames.

As demonstrated in our experiments, in comparison to H.264/AVC rate control [13] and reference [9], our proposed algorithm achieves accurate target bit rates and average PSNR gain and lower PSNR deviation to provide smoother visual quality. The bits produced by each frame are closer to the target bits. These results are very useful in determining the various target bit rates and frame rates in real time application.

References

[1] ISO-IEC/JTC1/SC29/WG11, *Information technology—coding of audio-visual objects—part 10: advanced video coding* Final Draft International Standard, ISO/IEC FDIS 14 496-10, December 2003.

[2] T.Wiegand, “Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264—ISO/IEC 14496-10 AVC),” in *Joint Video Team (JVT) of ISO/ICE MPEG and ITU-T VCEG, VT-G050*, Pattaya, Thailand, March 2003.

[3] T. Sikora, “Trends and perspectives in image and video coding,” *Proceedings of the IEEE*, vol. 93, no. 1, pp. 6–17, 2005.

[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[5] G. J. Sullivan and T. Wiegand, “Video compression—from concepts to the H.264/AVC standard,” *Proceedings of the IEEE*, vol. 93, no. 1, pp. 18–31, 2005.

[6] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172-185, Feb. 1999.

[7] Z. He, Y. K. Kim and S. K. Mitra, "Low delay rate control for DCT video coding via p domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no 8, pp. 928-940, Aug. 2001.

[8] F. Pan, Z. G. Li and K. P. Lim et al, "Adaptive intra-frame quantization for very low bit rate video coding," in *Proc. Int. Symp. Circuits and Systems, ISCAS'*04, pp. 781-784.

[9] M. Jiang and N.Ling, “On Enhancing H.264/AVC Video Rate Control by PSNR-Based Frame Complexity Estimation,” *IEEE Trans. Consumer. Electronics.* vol. 51, no. 1, pp. 281-286, Feb. 2005.

[10] X. Yi and N. Ling, “Improved H.264 rate control by enhanced MAD based frame complexity prediction,” *Journal of Visual Communication and Image Representation*, vol. 17, pp. 407-424, April 2006.

[11] L. Hung-Ju, C. Tihao, and Z. Ya-Qin, “Scalable rate control for MPEG-4 video,” IEEE Trans. Circuits and Syst. Video Technol., vol. 10, pp. 878-894, Sept. 2000.

[12] J. Ribas-Corbera and S.-M. Lei, “A frame-layer bit allocation for H.263+,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 7, pp. 1154-1158, Oct. 2000.

[13] JM12.1, http://iphome.hhi.de/suehring/tml/download

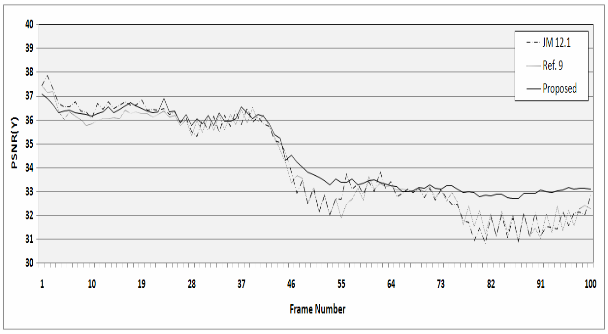

Fig. 1 Performance comparison of the proposed rate control scheme with the PSNR value per frame (Table 128 kbps, @30Hz)

Table. 2 Performance comparison of the proposed rate control scheme with the existing schemes (JM 12.1 and Ref.9)

Seq.	*PSNR(Y)*						*Encoding Time(㎲)*					
	A@15	*B@15*	*B@15*	*A@30*	*B@30*	*C@30*	*A@15*	*B@15*	*C@15*	*A@30*	*B.@30*	*C@30*
Container	*45.79*	*45.81*	*45.93*	*42.27*	*42.27*	*42.45*	*2,607,263*	*2,838,318*	*19*	*2,842,175*	*3,099,443*	*17*
News	*45.55*	*47.59*	*47.84*	*43.24*	*43.24*	*43.44*	*2,543,548*	*2,809,929*	*17*	*2,815,480*	*3,068,442*	*18*
Foreman	*41.85*	*41.85*	*41.86*	*38.00*	*38.01*	*38.02*	*2,689,941*	*2,909,196*	*17*	*2,909,089*	*3,176,842*	*17*
Mobile	*32.52*	*32.43*	*32.60*	*29.08*	*29/08*	*29.23*	*2,704,426*	*2,922.187*	*17*	*2,918,878*	*3,191,028*	*17*
Stefan	*34.19*	*34.11*	*34.16*	*30.16*	*30.14*	*30.23*	*2,612,460*	*2,824,663*	*17*	*2,822,680*	*3,084,532*	*23*
Akiyo	*51.79*	*51.69*	*52.47*	*48.28*	*48.31*	*48.41*	*2,386,903*	*2,727,425*	*17*	*2,732,621*	*2,978,348*	*17*
Coast.	*37.57*	*37.57*	*37.62*	*34.11*	*34.13*	*34.18*	*2,353,558*	*2,881,783*	*18*	*2,885,395*	*3,146,908*	*17*
H.M	*43.79*	*43.79*	*43.84*	*41.93*	*41.93*	*42.03*	*2,604,883*	*2,825,414*	*17*	*2,828,023*	*3,085,352*	*17*
M&D	*47.56*	*47.58*	*47.57*	*44.53*	*44.55*	*44.56*	*2,583,402*	*2,833758*	*17*	*2,838,588*	*3,094,464*	*17*
Pamphlet	*47.94*	*47.98*	*47.99*	*45.28*	*45.32*	*45.34*	*2,555,795*	*2,771,996*	*24*	*2,779,517*	*3,027,020*	*17*
Paris	*41.64*	*41.64*	*41.73*	*36.43*	*36.47*	*36.59*	*2,646,992*	*2,846,193*	*17*	*2,848,318*	*3,108,043*	*17*
Sean	*48.47*	*48.48*	*48.52*	*45.52*	*45.54*	*45.60*	*2,584,869*	*2,791,995*	*17*	*2,804,878*	*3,048,859*	*17*
Sign.	*44.25*	*44.26*	*44.26*	*40.04*	*40.06*	*40.04*	*2,603,462*	*2,799,857*	*17*	*2,803,919*	*3,057,443*	*17*
Silent	*46.39*	*46.37*	*46.48*	*41.97*	*41.95*	*42.08*	*2,652,375*	*2,852,929*	*17*	*2,863,802*	*3,115,398*	*24*
Table	*43.23*	*43.29*	*43.35*	*39.12*	*39.21*	*39.22*	*2,814,098*	*2,871,114*	*17*	*2,871,897*	*3,135,257*	*17*

※ *Computation Complexity (㎲) is only measured time unit for the rate control algorithm, especially at timer, which is the current value of the high-resolution performance counter(※ A@15 → JM12.1 @15, B@15 → Ref.9 @15, C@15 → Proposed @15).*