A New Macroblock-Layer Rate Control for H.264/AVC Using Quadratic R-D Model

Nam-Rye Son, Yoon-Jeong Shin, Jae-Myung Yoo, and Guee-Sang Lee

Dept. Computer Science, Chonnam National University, Youngbong-dong Pug-gu, Gwang-ju, South Korea
{nrson, gslee}@chonnam.ac.kr, syj@gwangju.ac.kr, jmyoo@cad.chonnam.ac.kr

Abstract. In recent years, rate control is an important technique in real time video communication applications using H.264/AVC. Many existing rate control algorithms employ the quadratic rate-distortion (R-D) model, which is determine the target bits for each I, P, B frame. In this paper, we analysis the problems in rate control for JVT video coding using quadratic R-D model. According to the analysis and experimental results, we present a new frame-layer rate control scheme for JVT video coding based on the quadratic R-D model. Also we estimate the target bit rate for macroblock-layer effectively. The experimental results show that with many scene changes between neighboring frames, existing algorithm exceeded transmission bit rate, while proposed algorithm tranquilized bit rate for stable delivery.

Keywords: H.264 video coding, Quadratic R-D Model, Mean Absolute Difference.

1 Introduction

H.264/AVC is the newest international video coding standard. By the time of this publication, it is expected to have been approved by ITU-T as recommendation H.264 and by ISO/IEC as International Standard 14496-10 (MPEG-4 part 10) Advanced Video Coding(AVC)[1]. The H.264/MPEG-4 AVC video compression standard promises a significant improvement over all previous video compression standards. In terms of coding efficiency, the new standard is expected to provide at least 2x compression improvement over the best previous standards and substantial perceptual quality improvements over both MPEG-2[2] and MPEG-4. The standard, being jointly developed by ITU-T and ISO/IEC, will address the full range of video applications including low bit rate wireless applications, SD(Standard Definition) and HD(High Definition) broadcast television, video streaming over the Internet, delivery of HD-DVD content, and the highest quality video for digital cinema applications.

One important component of video codec is rate control. It is a necessary part of an encoder to allocate the suitable number of bits to each video frame and then smooth out the variable bit rate to constant bit rate channel[3]. Rate control algorithms have been widely studied in standards, like MPEG-2, MPEG-4, H.263. For H.264/AVC, there are two different rate control algorithms proposed in the literature. JVT-D030 is

L.-W. Chang, W.-N. Lie, and R. Chiang (Eds.): PSIVT 2006, LNCS 4319, pp. 822 – 831, 2006.

based on MPEG-2 TM5 rate control[4] but is not suitable for low bit rate video applications. Our study bases on another scheme, which is based on fluid traffic model, linear model, and Rate Distortion Optimization (RDO)[5].

In short, since existing bit rate control scheme for JVT video coding codes current macroblock without accurately acknowledge of MAD(Mean Absolute Difference) to be coded. Because this, it cannot deal with drastically change of delivery channel status or video characteristics. Accordingly a new bit rate control scheme for H.264/AVC codec is needed considering characteristics of inter-frames.

The rest of the paper is organized as follows. In Section 2, we address the existing JVT-H014 video coding[5] and present the problem in the quadratic R-D model and in the rate control for JVT video coding. Then in Section 3, we propose a new macroblock-layer rate control using improved MAD. The experiment and results are presented in the section 4. Finally concluding remarks are given in Section 5.

2 Problems in Rate Control for JVT Video Coding

2.1 Estimated Bit Rate for GOP and Allocated the Target Bit Rate

According to [5], frame layer rate control scheme consists of two stages: pre-encoding and post-encoding. The objective of the first stage is to compute Quantization Parameter (QP) for all frames and it composed of two sub steps: (a) determine a target bit rate for each P frame and (b) compute the QP and perform RDO. In the pre-encoding stage, a quadratic R-D model is used to calculate the corresponding QPs, in the post-encoding stages, the model parameters are continually updated and buffer's control is performed. In this section, we summarize and analysis the JVT-H014 video coding used for estimating the target bits.

To estimate target bits for the current frame, a fluid traffic model is employed, which is based on the linear tracking theory[9]. For simplicity, let us assume one Group of Pictures (GOP) is used and the video sequence is encoded first as an I frame, and subsequently P frames. To illustrate rate control modeling, Let N denote the total number of frames in a GOP, $n_j (j=1,2,...N)$ denotes the j^{th} frame, and $B_c(n_j)$ denotes the occupancy of virtual buffer after coding the j^{th} frame. The fluid traffic model is stated as

$$\begin{aligned} B_c(n_1) &= \frac{B_s}{8} \\ B_c(n_{j+1}) &= \min \{\max \{0, B_c(n_j) + A(n_j) - \frac{u(n_j)}{F_r}\}, B_s\} \end{aligned} \tag{1}$$

Where B_s is the buffer size, $A(n_j)$ is the number of bits generated by the j^{th} frame, $u(n_j)$ is the available channel bandwidth, and F_r is the predefined frame rate. The determination of target bits for current P frame is composed of two steps.

Step 1. Determination of target buffer occupancy

Since the QP of the first P frame is given at the GOP layer in this algorithm, the initial value of target buffer level is set as

$$Tbl(n_2) = B_C(n_2) \tag{2}$$

Then the target buffer levels of other P frames in the GOP are predefined by using the following function

$$Tbl(n_{j+1}) = Tbl(n_j) - \frac{Tbl(n_{j-1}) - B_s/8}{N_p - 1} \tag{3}$$

Where N_p is the total number of P frames in the GOP.

Step 2. Computation of target bit rate
By using linear tracking, the target bit allocated for the j^{th} frame is determined based on the target buffer level, the frame rate, the available channel bandwidth, and the actual buffer occupancy as follows:

$$T_{buf} = \frac{u(n_j)}{F_r} + \gamma(Tbl(n_j) - B_c(n_j)) \tag{4}$$

Where γ is a constant and its typical value is 0.75. Meanwhile, the remaining bits are also computed as like function

$$T_r = \frac{R_r}{N_r} \tag{5}$$

Where R_r is the number of bits remaining for encoding this sequence, and N_r is the number of P frames remaining for encoding. The final target bit T is a weighted combination of T_r and T_{buf}.

$$T = \beta \times T_r + (1-\beta) \times T_{buf} \tag{6}$$

Where β is a weighting factor and its typical value is 0.5.

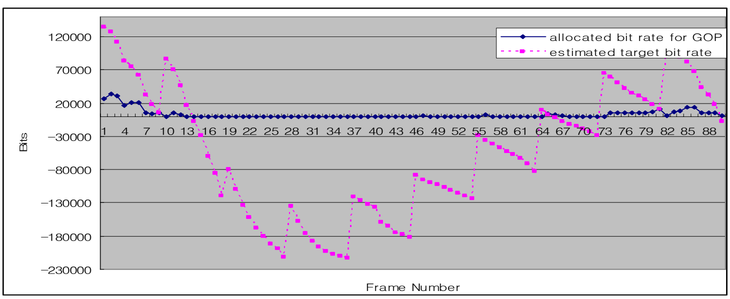


Fig. 1. Estimated target bits and allocated bit rate of GOP for TABLE TENNIS sequence

After the encoding results for TABLE TENNIS sequence, Figure 1 shows allocated bit rate for GOP and estimated target bit rate using JVT-H014 algorithm. In detail analysis, the value of target bit rate is '0' or 'positive' value because of considering both the status of buffer occupation after encoding and remained bit rate in the *eq*.(1). Specially, in the case of scene change, the value of allocated bit rate for GOP is 'negative' because estimated target bit rate is predicted wrongly, namely estimated target bit rate is greater than estimated buffer occupation.

2.2 Rate Control for JVT Video Coding Using Quadratic Rate-Distortion Model

This quadratic R-D model[6] is utilized for QP determination in rate control scheme, where the quadratic R-D model's parameters are estimated using MAD, which is prediction error between previous and current frame. Since quantization parameters are specified in both rate control and RDO, there exists a problem when the rate control is implemented: to perform RDO for MB, a quantization parameter should be first determined for the MB by using the MAD of MB. We take the following specific form of the simple MAD as like figure 2.

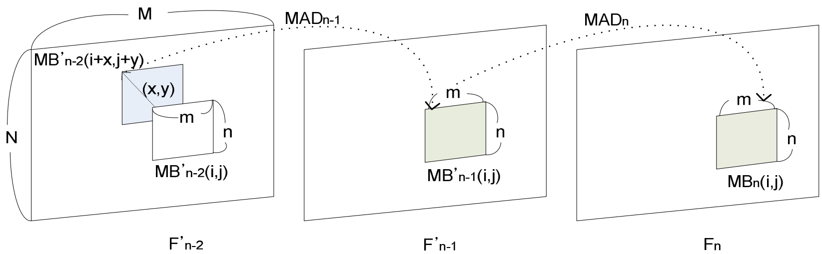


Fig. 2. Prediction of MAD for JVT video coding

$$\begin{aligned} MAD_{n-1}(i,j) &= \frac{1}{m \times n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|MB_{n-1}(i,j) - MB_{n-2}(i+x, j+y)\right| \\ MAD_n &= a_1 \times \frac{1}{X \times Y}\sum_{i=0}^{X-1}\sum_{j=0}^{Y-1} MAD_{n-1}(i,j) + a_2, (X = \frac{M}{m}, Y = \frac{N}{n}) \end{aligned} \tag{7}$$

In *eq*. (7), MAD_n is the mean absolute difference between pre-previous frame(F'_{n-2}) and previous frame(F'_{n-1}) in the JVT video coding, when current frame(F_n) is encoding. And MAD_{n-1} is predicted MAD of previously encoded P frame $(n-1)^{th}$. Where a_1 and a_2 are the two coefficients of the prediction model. The initial values of a_1 and a_2 are to 1 and 0, respectively. They are updated after coding each frame.

Finally a quadratic rate-distortion model is used to calculate the corresponding quantization parameter, which is then used for the RDO for each MB in the current basic unit as follows:

$$T = \frac{x_1 \times MAD_n}{QP} + \frac{x_2 \times MAD_n}{QP^2} \tag{8}$$

Where T is target bit rate, x_1 and x_2 are the model parameters. However, the MAD of current MB is only available after performing the RDO. This is a typical chicken and egg dilemma. Because of this, the rate control for JVT video coding is used to solve to estimate the MAD of current MB. Besides this, we also need to compute a target bit rate for the current MB and to determine the number of contiguous MBs that share the same quantization parameter.

3 New Macroblock-Layer Rate Control

In JVT-H014[5], the target bit is estimated solely based on the buffer fullness, regardless of the inter-frame motion. [7] introduced the MAD ratio as measure of motion complexity to improve the video quality at scene changes. However, MAD ratio is not a good ways for representing the motion contents, as it can only represent the similarity between the current frame and its reference frames. Also scene change, zooming, and fast motion may occur image quality degradation due to little correlation to the previous frames. This section deals with a new bit rate control scheme for H.264/AVC considering characteristics both of previous and current macroblock of frame and predicted buffer status, which controls for buffer overflow and underflow.

3.1 Bit Allocation per Frame for Target Bit Rate

In the section 2.1, we shows that JVT video coding is wrongly estimated the target bit rate for encoding frame. Then we proposed the effective predicted the target bit rate. That is to say, if estimated target bit rate is less than '0', we recognized status of buffer occupancy is overflow. We encoded adding 2 from QP of previous frame(QP_{pre}) in order to both diminish the target bit rate for encoding frame and maintain the smoothness of visual quality among successive frames.

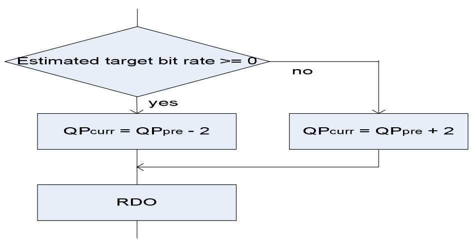


Fig. 3. Flowchart for proposed algorithm

3.2 Improved Quadratic R-D Model and Bit Allocation for Macroblock

There are two considerations to decide the quantization parameter for macroblock.

Case 1: Available bit to encode macro block in a frame is greater than zero
The quantization parameter corresponding to the target bit is then computed by using the modified quadratic mode as in *eq*. (9).

$$T_{MB(i,j)} = \frac{a_1 \times MAD'_{CB(i,j)}}{QP} + \frac{a_2 \times MAD'_{CB(i,j)}}{QP^2} \quad (9)$$

Where, $T_{MB(i,j)}$ is the target bit rate for current macroblock to be encoded. $MAD'_{CB(i,j)}$ is mean absolute difference between previous and current frame considering not MB's pixel value of motion compensation but co-located pixel value of MB for scene change as like figure 4. Also formula of $MAD'_{CB(i,j)}$ is given in *eq*. (10). Also a_1 and a_2 is updated after encoding.

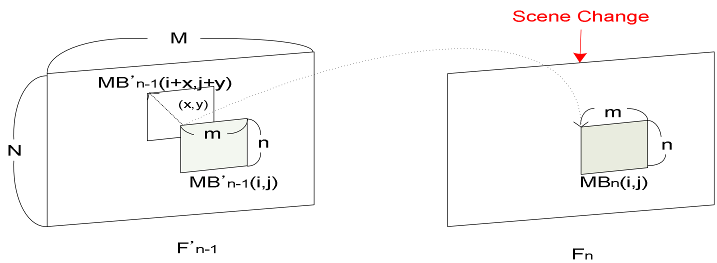


Fig. 4. Prediction of $MAD'_{CB(i,j)}$ for proposed rate control

$$MAD_n(i,j) = \frac{1}{m \times n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|MB_n(i,j) - MB'_{n-1}(i,j)\right|$$
$$MAD'_{CB(i,j)} = \frac{1}{X \times Y}\sum_{i=0}^{X-1}\sum_{j=0}^{Y-1} MAD_n(i,j), (X = \frac{M}{m}, Y = \frac{N}{n}) \quad (10)$$

Case 2: Available bit to encode macro block in a frame is smaller than zero

In this case the bit rate allocated to current macro block would be insufficient. Therefore picture quality of the macroblock and circumferential macroblock of the frame should be smoothness. To this end, as illustrated in figure 5, proposed algorithm adjusts the quantization parameter of the current macroblock in the same location of previous frame according to the macroblock mode of current frame. If current macro block mode is SKIP mode then the bit stream for texture and moving data will be '0'. Therefore use quantization parameter for macro block in previous frame. If current macro block mode is INTER mode bit stream will be less than in INTRA mode. Therefore add 1 to quantization parameter for macro block of previous frame. If current macro block mode is INTRA mode then add 2 to quantization parameter for macroblock of previous frame.

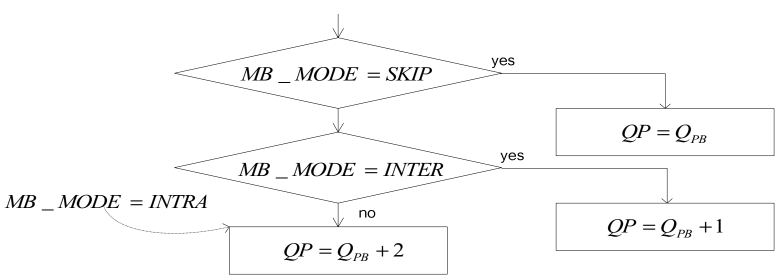


Fig. 5. Decision of quantization parameter for macroblock

3.3 Post-procession

This phase, the algorithm stores MAD which is coding result for macroblock, quantization parameter and coded bit rate(Bit_{MB}). Then it update target bit rate reflecting the coding result for next macroblock and the target bit rate for the frame as *eq*. (11).

$$T = T - Bit_{MB} \tag{11}$$

4 Experimental Results

To demonstrate the performance, we implement the proposed rate control algorithm in H.264 encoder reference software version 9.5[8]. This section presents the experimental results on typical test sequences. As recommended by JVT, the simulation common conditions presented by VCEG-N8111 are used. The test sequences are encoded by JM9.5 at the bit rate 32kbps and 64kbps. The generated bit rates are used as the target bit rates for the encoder with our rate control scheme and the one with JVT-H014 and JVT-D030 rate control scheme. For comparison, the basic unit in JVT-H014 rate control is selected to be a macroblock. In all simulation tests, the encoding parameters have been set next followings. The first picture is INTRA(I-TYPE) coded and the remaining pictures are P-TYPE. For all tests, the CAVLC mode is enabled and ME(Motion Estimation) search windows is set to 16. All other parameters such as de-blocking filter, context initialization and file mode have been carefully selected equivalent. In the test set, we used seven sequences with CCIR-601 parameters (176x144 pixels and a frame rate of 10fps): FOREMAN, CARPHONE, M&D(Mother&Daughter), SILENT, MOBILE, STEFAN, and TABLE TENNIS. In all experiments, the number of encoded frames was 100. We chose a GOP size of 10. The search was selected as 16x16 for P-type frames, and the alternate scan option was turned on for all cases.

Figure 5 depicts the estimated target bit rate and allocated bit rate for TABLE TENNIS sequence in the proposed algorithm. Proposed algorithm is efficiently distributes in the whole frames of both GOP and target bit rate in the scene change situation.

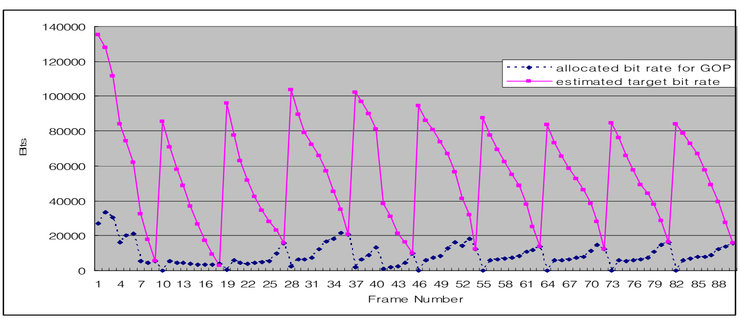


Fig. 5. Estimated target bits and allocated bit rate for TABLE TENNIS in the proposed algorithm

Table 1. The average PSNR values and average encoded bit rate at 32kbps for (I) JVT-D030, (II) JVT-H014, and (III) proposed algorithm

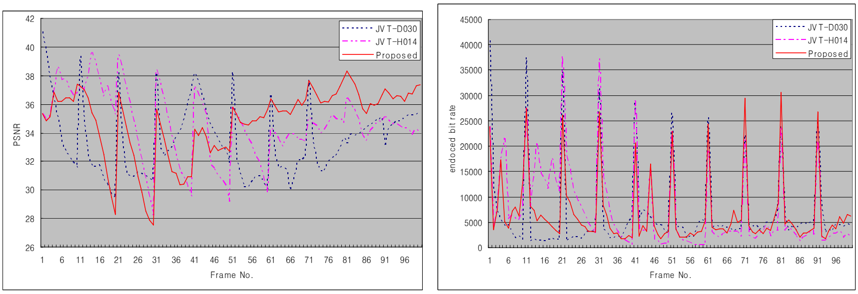
Sequence	JVT-D030		JVT-H014		Proposed	
	PSNR	encoded bit rate	PSNR	encoded bit rate	PSNR	encoded bit rate
FOREMAN	29.59	31.93	31.25	39.09	31.86	32.03
CONTAINER	34.22	31.96	35.91	41.00	35.89	31.84
M&D	36.71	31.98	37.31	39.02	37.72	31.51
SILENT	31.68	31.96	33.09	41.10	32.54	31.46
STEFAN	22.36	32.04	26.30	47.76	25.57	32.05
TABLE TENNIS	29.74	31.98	30.41	42.33	31.08	32.07
MOBILE	21.96	32.01	24.29	52.99	24.50	32.05

Table 1 and Table 2 show the PSNR performance and the encoding the bit rate of performance comparison for 1) JVT-D030, 2) JVT-H014, and 3) Proposed algorithm respectively at the 32kbps and 64kbps. The results indicate that significant PSNR gains can be obtained with the proposed algorithm. It is shown that by the proposed method the generated bit rates are very close to the target bit rates while keeping good video quality. These results justify that the proposed algorithm is superior to JVT-H014 rate control because of correctly predicting the target bit rate of each frame for scene change or high motion.

Table 2. The average PSNR values and average encoded bit rate at 64kbps for (I) JVT-D030, (II) JVT-H014, and (III) proposed algorithm

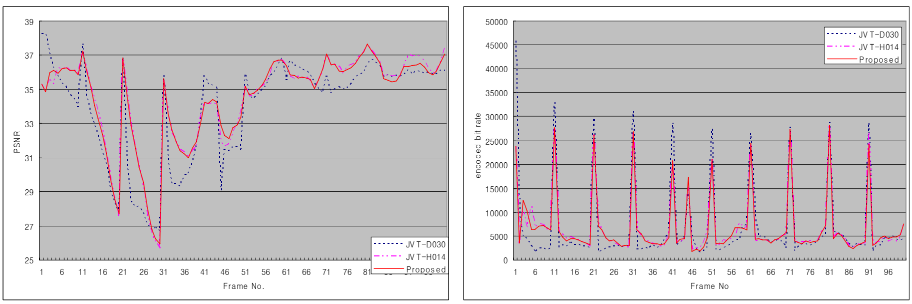
Sequence	JVT-D030		JVT-H014		Proposed	
	PSNR	encoded bit rate	PSNR	encoded bit rate	PSNR	encoded bit rate
FOREMAN	33.42	63.96	34.43	67.74	34.48	63.44
CONTAINER	37.71	63.96	38.89	78.08	38.40	63.17
M&D	40.20	63.94	40.37	67.10	40.80	63.10
SILENT	35.50	63.82	36.62	77.26	35.90	63.99
STEFAN	25.57	64.12	26.30	81.86	27.02	64.11
TABLE TENNIS	33.90	63.98	34.53	73.11	34.90	63.97
MOBILE	25.65	63.99	27.19	77.51	27.23	63.94

Figure 6 and 7 represents value of PSNR and encoded bit stream for FOREMAN and TABLE TENNIS sequences respectively at the 64kbps. In the case of encoded bit rate and PSNR, our proposed algorithm and JVT-D030 does not change abruptly in the high motion, but JVT-H014 is changed according to variance of motion.



(a) (b)

Fig. 6. (a) PSNR value at each frame and (b) encoded bit rate at each frame for FOREMAN sequence for (I) JVT-DO30, JVT-H014, and (III) proposed algorithm



(a) (b)

Fig. 7. (a) PSNR value at each frame and (b) encoded bit rate at each frame for TABLE TENNIS sequence for (I) JVT-DO30, JVT-H014, and (III) proposed algorithm

5 Conclusions

In this paper, we proposed an effective rate control algorithm for abruptly scene change and high motion by more accurately predicting frame complexity using quadratic rate-distortion the statistics of previous and current macroblock of frame. Based on our extensive tests and computational analysis, we show that the coding efficiency achieved by the proposed rate control algorithm is similar to or even better than that of the JVT-D030 and the JVT-H014 rate control algorithm while keeping high motion and scene change.

Acknowledgement

This work was supported by Korea Research Foundation Grant (KRF-2004-041-D00527).

References

1. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-G050r1 (2003)
2. ISO/IEC JTC1/SC29/WG11 MPEG/N0400: MPEG-2 Test Model5 (1993)
3. T. Wiegand, H. Schwarz, A. Joch, F. Lossentini, and G. J. Sullivan: Rate-constrained Coder Control and Comparison of Video Coding Standards. IEEE Transaction on Circuit and System for Video Technology, Vol. 13, No 7(2003) 688–703
4. Siwei Ma, Wen Gao, Peng Gao, Yan Lu: Rate Control for JVT Standard. JVT of ISO/IEC MPEG and ITU-T VCEG, JVT-D030.doc(2002)
5. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Document JVT-H014(2003)
6. Siwei Ma, Wen Gao, Peng Gao, Yan Lu: Rate Control for Advanced Video Coding standard. Proceedings of the 2003 International Symposium on ISCAS '03., Vol. 2(2003) 892-895.
7. M. Q. Jiang, X. Q. Yi, N. Ling: Improved Frame-layer Rate Control for H.264 using MAD Ratio. in Proc. 2004 IEEE Int. Symp. Circuits Syst., Vol 3(2004), 813-816
8. JVT Reference Software Version JM9.5: http://bs.hhi.de/~suehring/tml/
9. F. Pan, Z. Li, K. Lim, G. Feng: A Study of MPEG-4 Rate Control Scheme and Its Improvements. IEEE Trans. Circuit Syst. Video Technol., Vol. 10(2000) 878-894