A Low Complexity Macroblock Layer Rate Control Scheme Base on Weighted-Window for H.264 Encoder

Huibo Zhong, Sha Shen, Yibo Fan, and Xiaoyang Zeng

State Key Lab of ASIC and System, Fudan University
825 Zhangheng Road, Shanghai, 201203, China
fanyibo@fudan.edu.cn

Abstract. Rate control plays a very important role in video coding. A low complexity macroblock (MB) layer rate control scheme for H.264 encoder is presented in this paper. Based on the analysis of the relationship among the quantization parameter (QP), mean absolute distortion (MAD) and the coded bits, a weighted-window model is proposed. A weighted-window based QP decision and MAD prediction model is proposed to reduce the computational complexity of MB-layer rate control. A new rate control scheme based on these models is presented in detail. The experimental results show that the proposed scheme gives a quality improvement of about 0.80dB on the average for all sequences, and about 58% reduction in bit rate mismatch.

Keywords: H.264, rate control, weighted-window, Video coding.

1 Introduction

Rate control plays a very important role in real-time video communication applications. The goal of rate control is to regulate the coded bitstream to meet certain given constraints, such as bit rate, buffer overflow/underflow prevention. As one of the key technologies in regarded to coding performance, rate control has drawn significant research attentions.

1.1 Brief Review of Rate Control Scheme

In the field of video transmission, as the available channel bandwidth for the coding process can be constant or time varying, the rate control schemes can be classified into two major categories: constant-bit-rate (CBR) control for constant channel bandwidth and variable-bit-rate (VBR) control for variable channel bandwidth. The existing rate control schemes focus on the CBR case. There are two main problems in rate control schemes: the first is how to allocate proper bits to each coding unit, and the second is how to select the quantization parameter (QP) to encode each unit with the allocated bits. The problem of optimum bits allocation can be described by the formula as follows:

$$\min\{D\},\ \ \textit{subject to}\ R_c \le R_t \tag{1}$$

Where D denotes the distortion of current encoding unit, R_c and R_t denote the number of bits used to encode the unit and the bits budget respectively. The key point of

K. Schoeffmann et al. (Eds.): MMM 2012, LNCS 7131, pp. 563–573, 2012.

quantization parameter (QP) selection is to find the relation between rate and QP. The relation between rate and QP is usually derived based on a rate-quantization (R-Q) model, and the most common R-Q models are either linear or quadratic. Linear R-Q model has been studied in MPEG-2 TM5 rate control [1]. The quadratic R-Q model, originally proposed for MPEG-4 Q2 rate control [2] [3], was considered better and more accurate than the linear one. This quadratic R-Q model has been adopted in the JM reference software for H.264 rate control [4~6].

1.2 Rate Control in H.264

According to the terms of the unit of rate control operation, rate control schemes for H.264 can be classified into MB-layer, frame-layer, group-of-picture- (GOP) layer rate control [6]. The rate control for H.264 is more difficult than those for other standards such as MPEG-2, MPEG-4, H.263, and so on. This is because the quantization parameters are used in both rate control algorithm and rate distortion optimization (RDO), which results in the following chicken-and-egg-dilemma: to perform RDO for MBs in the current frame, a quantization parameter should be first determined for each MB by using the mean absolute distortion (MAD) of current frame or MB, however, the *MAD* of current frame or MB is only available after the RDO. To solve this dilemma, a linear model using the actual *MAD* of the basic unit in the same position of previous frame is proposed in [6]:

$$MAD_a[k,i] = a_1 \times MAD_a[k,i-1] + a_2 \tag{2}$$

Where a_1 and a_2 are two coefficients of prediction model, MAD_a[k,i] and MAD_a[k,i-1] are the *MAD* of k^{th} MB in current (i) and previous (i-1) frame, respectively. By the value of MAD_a[k,i], the corresponding QP can be computed by following quadratic R-Q model:

$$R = c_1 \frac{MAD_a[k,i]}{Qs} + c_2 \frac{MAD_a[k,i]}{Qs^2} \tag{3}$$

Where the two parameters c_1 and c_2 are model coefficients, and *R* denotes the target bits used for encoding the current basic unit, *Qs* is quantizer step size. QP can be obtained by converting *Qs* as defined in [4].

It is noted that by employing a bigger coding unit, a higher PSNR can be achieved while the bit fluctuation is also bigger. On the other hand, by using a smaller coding unit, the bit fluctuation is less severe. This is very useful in some communication application. By employing a MB-layer rate control, the bit fluctuation can be the smallest; however, this will introduce the highest computational complexity and make it hard for real-time applications. Many previous works have paid attention to the accuracy of (1) and (2). To improve the performance of (2), frame bits allocation and MAD estimation accuracy have been enhanced using a PSNR-based frame complexity measure in [7]. Multi-pass rate control can achieve higher performance than one-pass rate control, [8] has proposed an optimized two-pass rate control with a linear R-Q model, however, it will introduce higher complexity. A weighted model for MAD prediction is proposed using for rate control in embedded systems [9], a simple QP decision method is proposed by judging the MAD of current MAD and previous frame average MAD, only a little better RD performance can be achieved compared with the JM's. Since any model is actually an approximate model and cannot always match ideally with real application, the estimation of model parameters usually has some deviation or errors.

In this paper, base on the analysis among QP, MAD, and coded bit, a Weighted-Window prediction model is proposed. This model removes the complex update of the coefficients in (1) and (2) [10]. A weighted-window model based QP decision and MAD prediction model is proposed to reduce the computational complexity of MB-layer rate control, a new rate control scheme based on these models is presented in detail. The experiment results show that our scheme achieves higher PSNR and smaller overall bit rate mismatch compared with JVT-G012 [6].

The rest of this paper is organized as follow. Section 2 presents the proposed rate control algorithm in detail. The experiment results and the comparison are given in Section 3. Finally, section 4 shows the conclusion.

2 Proposed Weighted-Window Based Rate Control Scheme

Temporal and spatial information are always been used to predict MAD and QP. In this section, a weighted-window model is proposed in this section. The relationship between QP, MAD and coded bit is analyzed to decide the size of the window. Based on this weighted-window model, temporal and spatial information are used. A new QP decision and MAD prediction model are proposed. And finally, the framework of the proposed rate control is presented in detail.

2.1 Weighted-Window Model

Experiments have shown that the average MAD of the current frame becomes bigger if the previous frame is quantized with a larger QP [9]. This is because the motion estimation has to refer to the reconstructed previous frame with more distortion. On the other hand, if the QP in the previous frame is smaller, the average MAD of the current frame should be decreased. Hence, we can jointly considering the average MAD and the average QP of current frame and the previous reference frame. Fig.1 illustrates two windows in current frame and previous frame. The window in current frame can be seen as spatial information and the window in previous frame as the temporal information. We use *Ws* to measure the size of the window, *Ws* equal to 3 denotes that the window is 3x3 macroblock square. Fig.1 is a case with a window size equals to 3. $MB0_c$ means current MB in current frame, and $MB1_c$~ $MB8_c$ are the corresponding MB of Left, Right, Top, Down, Top left, Top right, Down left, Down right, respectively. To find the relationship between the QP and MAD in the two windows, we introduce two variables N_c and N_p whose values are given by:

$$N_c = \sum_{k \in Wc} (\alpha_k \times QP[k,i]) / \sum_{k \in Wc} (\alpha_k \times MAD_a[k,i]) \tag{4}$$

$$N_p = \sum_{k \in Wp} (\beta_k \times QP[k,i-1]) / \sum_{k \in Wp} (\beta_k \times MAD_a[k,i-1]) \tag{5}$$

Where *Wc* and *Wp* denote the windows in current and previous frame respectively, *QP*[k,i], *QP* [k,i-1] are the QP of the MBs in the window of current and previous frame, respectively. MAD_a[k,i], MAD_a[k,i-1] are the MAD of the MBs in the two windows, α_k and β_k are weight factors.

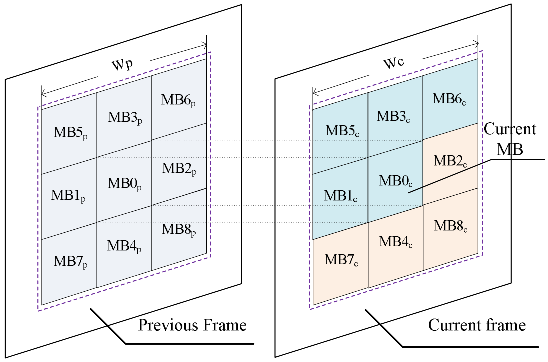
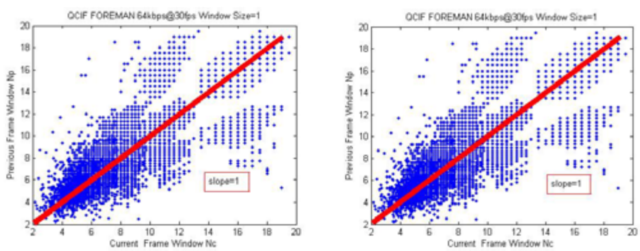
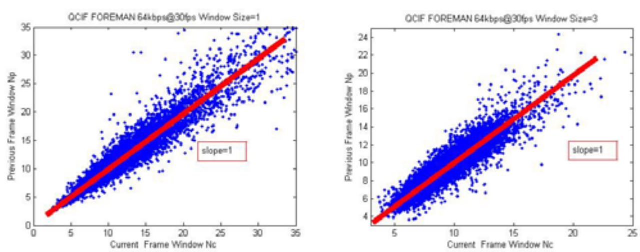


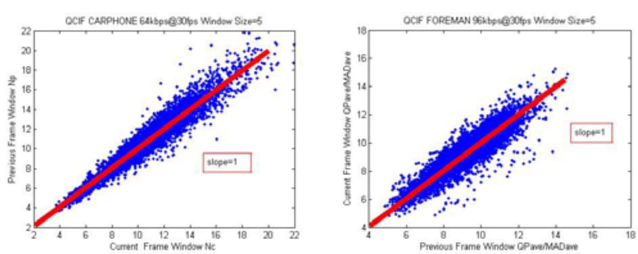
Fig. 1. Proposed weighted-window model with *Ws* is equal to 3



(a) Window size = 1



(b) Window size = 3



(c) Window size = 5

Fig. 2. The relationship N_c and N_p at different window size for test sequences: Carphone (left) and Foreman (right) 96kbps@30fps

To analyze the relationship between N_c and N_p, let $\alpha_0=\ldots=\alpha_k$, $\beta_0=\ldots=\beta_k$, and numerous experiments have been done according (4) and (5) at different window size. The results are shown in Fig.2, the bigger the window size, the closer the correlation between N_c and N_p; when Ws is 3 or larger, N_p is almost linear to N_c as follows:

$$N_c = N_p \quad if \ \ Ws \geq 3 \tag{6}$$

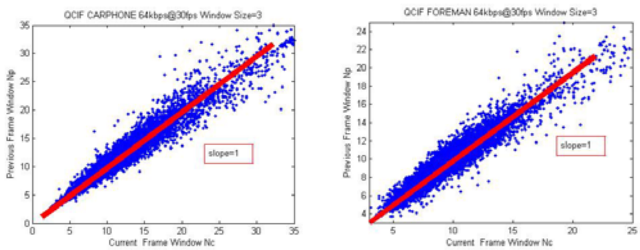


Fig. 3. Relationship between $N_{c,2}$ and N_p:Carphone (left) and Foreman (right) 64kbps @30fps

2.2 QP Computation Model Based on Weighted-Window

As mentioned above, when the window size is 3 or larger, we can have $N_c = N_p$. The more candidates taken into account, the more parameters should be determined, which will introduce more computational complexity. The value of Ws is set as 3 in this paper. As shown in Fig.1, at the current frame, the data of $MB2_c$, $MB7_c$, $MB4_c$, and $MB8_c$ is not available when the encoder is coding $MB0_c$, so (4) has to be adjusted. The data of $MB2_c$, $MB4_c$, $MB7_c$ and $MB8_c$ are substituted by $MB2_p$, $MB4_p$, $MB7_p$ and $MB8_p$ as follows:

$$\begin{aligned} N_{c,2} = \{ & \sum_{k=0,1,3,5,6} \alpha_k \times QP[k,i] + \sum_{k=2,4,7,8} \alpha_k \times QP[k,i-1]\} \\ & / \{ \sum_{k=0,1,3,5,6} \alpha_k \times MAD_a[l,i] + \sum_{k=2,4,7,8} \alpha_k \times MAD_a[k,i-1]\} \end{aligned} \tag{7}$$

Similar relationship between $N_{c,2}$ and N_p can be obtained as Fig.3 shown when $Ws = 3$.

$$N_{c,2} = N_p \qquad when \ Ws = 3 \tag{8}$$

For MB layer, the coded bit of each MB should also be considered. The bit allocated for each MB is by judging the value of the current MAD, the remaining bit, and the complexity of current picture. Combine (5), (7) and the coded bit, we have:

$$QP[0,i] = \{S_c \times N_p \times \phi - \sum_{k=1,3,5,6} \alpha_k \times QP[k,i] - \sum_{k=2,4,7,8} \alpha_k \times QP[k,i-1]\} / \alpha_0 \tag{9}$$

Where S_c is computed as:

$$S_c = \sum_{k=0,1,3,5,6} \alpha_k \times MAD_a[k,i] + \sum_{k=2,4,7,8} \alpha_k \times MAD_a[k,i-1] \tag{10}$$

According to experiment results, the weight factors are set as follows:

$$\alpha_k = \begin{cases} 3, & k=0 \\ 2, & k=1,2,3,4 \\ 1, & k=5,6,7,8 \end{cases} , \quad \beta_k = \begin{cases} 4, & k=0 \\ 2, & k=1,2,3,4 \\ 1, & k=5,6,7,8 \end{cases} \tag{11}$$

Where 0~8 denote the position of MB0~MB8 as shown in Fig.1. ϕ is a regulated factor which value is computed by :

$$\phi = \frac{Bit[0,i]}{\sum_{k\in Wp} \beta_k \times Bit[k,i-1] / \sum_{k\in Wp} \beta_k} \tag{12}$$

Where *Bit*[0,i] is the coded bit for current MB，*Bit*[k,i-1] is the coded bit of the MBs in the window of previous frame. ϕ is used for regulating the obtained QP according to the consumed bits. Before encoding current MB, *Bit*[0,i] is always obtained by bit allocation according to the predicted MAD of current MB. If *Bit*[0,i]> $\sum_{k\in Wp} \beta_k \times Bit[k,i-1] / \sum_{k\in Wp} \beta_k$, it means the bit used for current MB is large than the number of the bits previous window used, ϕ >1 is obtained to achieve a relative large QP, else if *Bit*[0,i]< $\sum_{k\in Wp} \beta_k \times Bit[k,i-1] / \sum_{k\in Wp} \beta_k$, the bit used for current MB is smaller than the previous window used, ϕ <1 is obtained to achieve a relative small QP . The QP obtained by (9) is rounded as follows:

$$QP = \lfloor QP[0,i] + 0.5 \rfloor \tag{13}$$

Where $\lfloor \ \rfloor$ is the floor operation, the further bound of the QP is presented in the following parts.

2.3 MAD Prediction Model Based on Weighted-Window

As we mentioned above, for the current MB, we can obtain the value of QP according to (9), however, $MAD_a[0,i]$ can only be obtained until RDO is done. As the one which is used for QP decision model, we use the same weighted-window to predict the MAD of current MB. According to our experiment, similar relationship exists between the MAD of previous frame and current frame. Same as the QP decision model, a similar MAD prediction model based on weighted-window can also be established. As Fig.1 shows, we use the previous window to predict the MAD of current window in current frame. Two variables M_p and M_c are introduced as follows:

$$M_p = \sum_{k=0}^{8} \beta_k \times MAD_a[k,i-1] / \sum_{k\in Wp} \beta_k \tag{14}$$

$$M_c = \{ \sum_{l=0,1,3,5,6} \gamma_l \times MAD_a[l,i] + \sum_{l=2,4,7,8} \gamma_l \times MAD_a[l,i-1] \} / \sum_{l\in Wp} \gamma_l \tag{15}$$

Where γ_l and β_k are the weight factors, the value of β_k is the same as (11) shown. We analyze the relationship between M_p and M_c in the same way as N_p and N_c. The results are presented in Fig.4. We can have:

$$M_c = M_p \quad when\ Ws = 3 \tag{16}$$

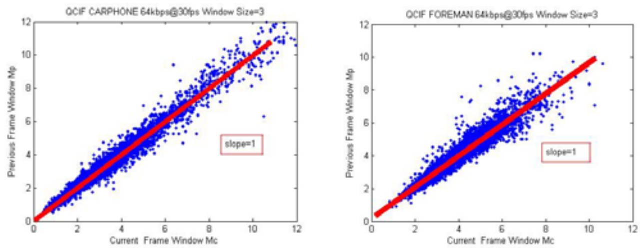


Fig. 4. Relationship between M_c and M_p @Ws=3 for test sequences: Carphone (left) and Foreman (right) 64kbps@30fps

As shown in Fig.4, we can have $M_p = M_c$, $MAD_a[0,i]$ can be computed by:

$$MAD_a[0,i] = \{M_p - (\sum_{l=1,3,5,6} \gamma_l \times MAD_a[l,i] - \sum_{l=2,4,7,8} \gamma_l \times MAD_a[l,i-1]) / \sum_{l \in Wp} \gamma_l\} / \gamma_0 \tag{17}$$

Where the weight factor γ_l is given as following according to the experimental results:

$$\gamma_l = \begin{cases} 3, & l = 0 \\ 1, & l = 1 \sim 8 \end{cases}$$

2.4 Proposed Rate Control Framework

With the proposed QP and MAD model based on the weighted-window, we now present our rate control scheme for H.264. The proposed rate control scheme includes three different coding granularities: the GOP-layer, frame-layer, and MB- layer. At the GOP-level and frame-layer, it is the same way as [6] to allocate target bits and perform the post-encoding regulation. Now we focus on a step-by-step description of the proposed scheme at MB-layer for P-frames as Fig.5 shows.

In Fig.5, for the first I/P frame, the initial QP is computed by the same way as Li's [6], QP_{ave} is the average QP of previous frame, $QP[k-1,i]$ is the QP of the previous MB in current frame. The MB layer bits allocation is according to the predicted MAD obtained from (17) and the average MAD of all coded MBs in current frame as following shows:

$$T_c = T_r \times MAD_a[0,i] / (MAD_{ave,c} \times N) \tag{18}$$

T_c and T_r are the target bits for the current MB and the remaining bits for all MBs which are not encoded yet in the current frame. $MAD_{ave,c}$ denotes the average MAD of all coded MB in current frame and N is the number of the remaining MBs to be encoded. T_r is updated by substracting the total encoded bits of encoded MB from it.

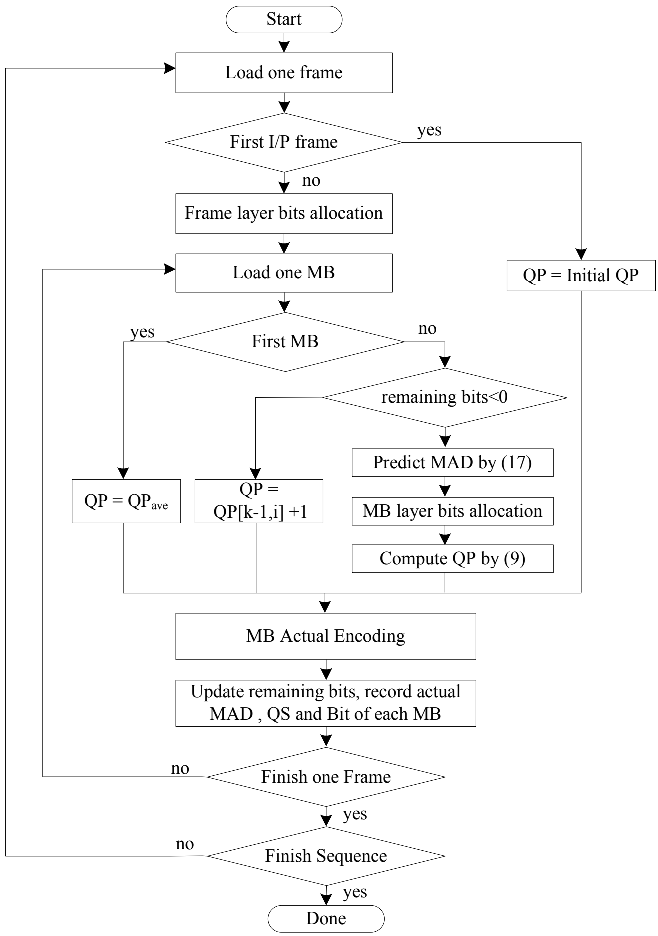


Fig. 5. Proposed rate control scheme framework base on the weighted-window

If current MB is the first MB in current frame, current QP is set to be the average QP of the previous P-frame. For other MBs, QP can be calculated as follows. If the remaining bits are negative, the current QP is set to *QP*[k-1,i]+1 to achieve frame level actual bits that are closer to target bits, otherwise, allocate bits for the current MB by (18), and calculate QP for the MB by (9). The derived QP value should also be restricted by the QP value of the previously encoded MB to reduce blocking artifacts by (19):

$$QP[k,i] = \min\{QP[k-1,i]+1, \max\{QP[k-1,i]-1, QP\}\} \tag{19}$$

Qp[k,i] is the QP value for the k^{th} MB in current frame. To maintain the smoothness of the visual quality within one sequence, the QP value is further adjusted by (18):

$$QP[k,i] = \min\{QP_{ave} + 2, \max\{QP_{ave} - 2, QP\}\} \tag{20}$$

And finally, the QP value should be restricted between 1 and 51, which is provided in H.264/AVC [4]:

$$QP[k,i] = \min\{51, \max\{1, QP\}\} \tag{21}$$

After encoding each MB, the encoder should update the remaining bits and record the data such as the actual QP, coded bit and MAD.

3 Experimental Results

Our proposed rate control scheme is implemented in JM15.1 [5], to evaluate the performance of the proposed rate control scheme, the test parameters for encoding are: 1) CABAC is used; 2) Hadamard transform is used; 3) Max reference frame number is 5, and search range is 16; 4) Fast full search is used; 5) RDO is on, and rate control mode is 0. All other parameters are carefully selected for both algorithms to be equivalent. For each sequence, 300 frames are encoded at 30fps, and the GOP structure is IPPP, the GOP length is 300 if not specified, for each GOP, the first frame is IDR frame, and the following 299 frames are P-frame.

Table 1. Rate control performance comparison between JM15.1 and the proposed

Sequence	Target rate (kbps)	JM 15.1		Proposed	
		PSNR	Rate	PSNR(dB)	Rate
akiyo	48	40.64	48.08(+0.08)	41.70(+1.06)	48.05(+0.05)
	64	42.12	64.10(+0.10)	43.03(+0.91)	64.06(+0.06)
	96	44.51	96.19(+0.19)	45.07(+0.56)	96.07(+0.07)
Carphone	48	32.02	48.09(+0.09)	32.67(+0.65)	48.03(+0.03)
	64	33.26	64.09(+0.09)	33.88(+0.66)	64.03(+0.03)
	96	35.30	96.12(+0.12)	35.83(+0.53)	96.01(+0.01)
Container	48	36.43	48.04(+0.04)	37.17(+0.74)	48.02(+0.02)
	64	37.66	64.04(+0.04)	38.25(+0.59)	64.02(+0.02)
	96	39.34	96.06(+0.06)	39.84(+0.50)	96.04(+0.04)
foreman	48	31.10	48.10(+0.10)	31.50(+0.40)	48.03(+0.03)
	64	32.49	64.08(+0.08)	32.89(+0.40)	64.05(+0.05)
	96	34.50	96.17(+0.17)	34.82(+0.32)	96.06(+0.06)
grandma	48	37.25	48.08(+0.08)	38.18(+0.93)	48.05(+0.05)
	64	38.55	64.06(+0.06)	39.46(+0.91)	64.03(+0.03)
	96	40.92	96.10(+0.10)	41.72(+0.80)	96.09(+0.09)
salesman	48	35.21	48.07(+0.07)	36.25(+1.04)	48.02(+0.02)
	64	37.07	64.04(+0.04)	37.94(+0.87)	63.98(-0.02)
	96	39.90	96.06(+0.04)	40.48(+0.58)	96.04(+0.04)

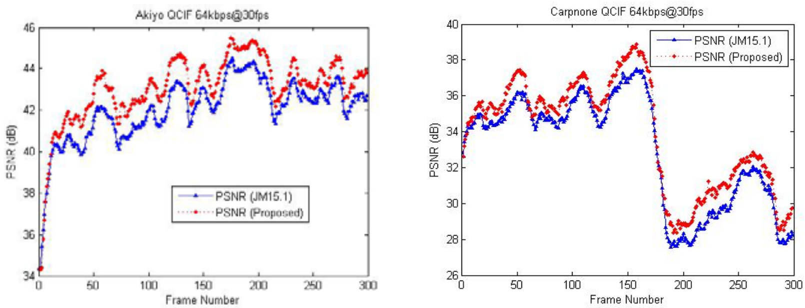


Fig. 6. PSNR comparison frame by frame: Akiyo (left), Carphone (right) 64kbps @30fps

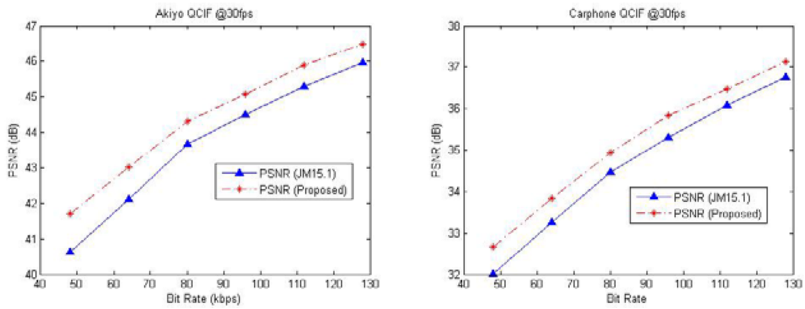


Fig. 7. PSNR comparison at different bit rate: Akiyo (left), Carphone (right) 64kbps @30fps

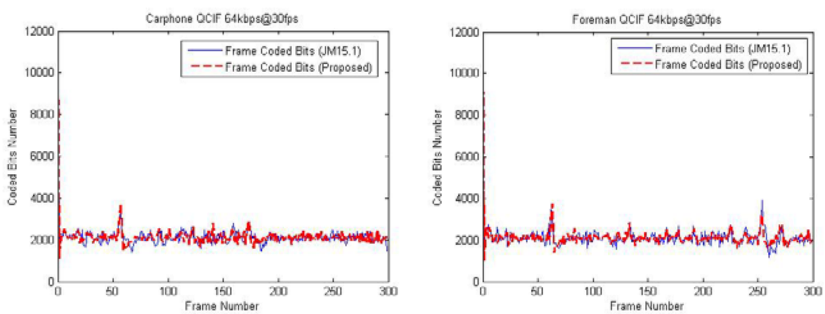


Fig. 8. Coded bits comparison frame by frame: Carphone (left), Foreman (right) 64kbps@30fps

The rate distortion performance comparison is summarized in Table 1. For all test sequences, the target bit rate is set to 48, 64 and 96kbps. It shows that our proposed rate control scheme achieves better results with a largest increase in PSNR about 1.06dB (Akiyo @48kbps) and an average increase for all test sequences of about 0.80dB @48kbps, 0.72dB @64kbps, and 0.55dB @96kbps. Fig.6 illustrates the comparison of the PSNR curse for QCIF sequences Akiyo and Carphone @64kbps frame by frame. Fig.7 gives the PSNR comparison at different bit rate. The results prove that the proposed rate control scheme outperforms the original rate-control scheme proposed in JM15.1 [5].

Table 1 also shows that the bit rate mismatches in our proposed rate control scheme is smaller than that of JM15.1. The average bit rate mismatch in our proposed rate control scheme for all sequences is about 0.069%, 0.055% and 0.054% at different bit rate. The corresponding values are 0.188%, 0.107% and 0.122% respectively in JM 15.1. The average bit rate mismatch for all video sequences is reduced by 58% with our proposed rate control scheme. Fig.8 shows the number of coding bits frame-by-frame of QCIF sequence Carphone and Foreman 64kbps @30fps.

4 Conclusion

Based on the analysis of the relationship among the quantization parameter (QP), mean absolute distortion (MAD) and the coded bits, a weighted-window model is proposed in this paper. A weighted-window model based QP decision and MAD prediction model is proposed to reduce the computational complexity of MB-layer rate control. A new rate control scheme based on these models is presented in detail. The experimental results show that the proposed scheme gives a quality improvement of about 0.80dB on the average for all sequences, and about 58% reduction in bit rate mismatch.

Acknowledgments. This work is supported by the State Key Lab of ASIC & System, Fudan University. Project Number: 11MS004.

References

1. Test Model Editing Committee, MPEG-2, Test Model 5, Doc. ISO/IEC JTC1/SC29 WG11/93-225b (April 1993)
2. Lee, H.J., Chiang, T.H., Zhang, Y.Q.: Scalable Rate Control for MPEG-4 Video. IEEE Trans. CSVT 10, 878–894 (2000)
3. Chiang, T., Zhang, Y.Q.: A New Rate Control Scheme using a New Rate-Distortion Model. IEEE Trans. CSVT, 246–250 (February 1997)
4. ITU-T, H.264, Advanced video coding for generic audiovisual services (March 2005)
5. Joint Video Team (JVT) of ITU-T VCEG and ISO/IEC MPEG, Joint Model (JM) Reference Software Version 15.1, http://iphome.hhi.de/suehring
6. Li, Z.G., Pan, F., Lim, K.P., Feng, G.N.: Adaptive basic unit layer rate control for JVT, JVT-G012. In: 7Th meeting, Pattaya, Thailand (2003)
7. Jiang, M., Ling, N.: Low-delay rate control for real-time H.264/AVC video coding. IEEE Trans. Multimedia, 467–477 (June 2006)
8. Kwon, D.-K., Shen, M.Y., Kuo, C.C.: Rate control for H.264 Video With Enhanced Rate and Distortion Models. IEEE Trans. CSVT, 517–529 (2007)
9. Kuo, C.H., Chang, L.C., Fan, K.W., Liu, B.D.: Hardware/software co-design of low cost rate control scheme for H.264/AVC. IEEE Trans. CSVT 20(2), 250–261 (2010)
10. Chang, L.C., Kuo, C.H., Liu, B.D.: A Two-Stage Rate Control Mechanism for RDO-Based H.264/AVC Encoders. IEEE Trans. CSVT, 660–673 (May 2011)