



ELSEVIER

Contents lists available at ScienceDirect

# Signal Processing: *Image Communication*

journal homepage: [www.elsevier.com/locate/image](http://www.elsevier.com/locate/image)

## A two-pass rate control algorithm for H.264/AVC high definition video coding

Dongdong Zhang<sup>a,b,\*</sup>, King Ngi Ngan<sup>b</sup>, Zhenzhong Chen<sup>b,c</sup><sup>a</sup> Department of Computer Science, Tongji University, Shanghai, China<sup>b</sup> Department of Electronic Engineering, The Chinese University of Hong Kong, China<sup>c</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

### ARTICLE INFO

#### Article history:

Received 25 September 2008

Received in revised form

28 February 2009

Accepted 12 March 2009

#### Keywords:

Two-pass

Rate control

Constant quality

High definition video coding

### ABSTRACT

In this paper, we propose a novel two-pass rate control algorithm to achieve constant quality for H.264/AVC high definition video coding. With the first-pass collected rate and distortion information and the built model of scene complexity, the encoder can determine the expected distortion which could be achieved in the second-pass encoding under the target bit rate. According to the built linear distortion-quantizer (D-Q) model, before encoding one frame, the quantization parameter can be solved to realize constant quality encoding. After encoding one frame, the model parameters will be updated with linear regression method to ensure the prediction accuracy of the quantization parameter of next encoded frame with the same coding type. In order to obtain the expected distortion of each frame under the target bit rate, a GOP-level bit allocation scheme is also designed to adjust the target bit rate of each GOP based on the scene complexity of the GOP in the second-pass encoding. In addition, the effect of scene change on the updating of D-Q model is considered. The model will be re-initialized at the scene change to minimize modeling error. The experimental results show that compared with the latest two-pass rate control algorithm, our proposed algorithm can significantly improve the bit control accuracy at comparable coding performance in terms of constant quality and average PSNR. On average, the improvement of bit control accuracy achieved about 90%.

© 2009 Elsevier B.V. All rights reserved.

### 1. Introduction

The H.264/AVC is the most effective and popular video coding standard, which has been used extensively for various video applications such as high definition television (HDTV) broadcasting, high definition digital video disc (HD-DVD), mobile TV, videoconferencing and internet video streaming. In terms of the requirement of bit rate, these video applications can be classified into two

categories: constant bit rate (CBR) encoding and variable bit rate (VBR) encoding. CBR encoding means that an encoder generates the output bit stream at a constant rate. It is useful for streaming multimedia content on limited capacity channels, but it is not the optimal choice for storage as it cannot allocate enough data for complex sections (resulting in degraded quality) while wasting data on simple sections. As opposed to CBR encoding, VBR encoding vary the amount of output data per time segment so that a better visual quality can be achieved by flexibly allocating the coding bits according to the complexity of different sections. Whatever for CBR or VBR encoding, a well-designed rate control scheme is necessary for an encoder to meet the requirement of the network bandwidth limitation, the transportation delay,

\* Corresponding author at: Department of Computer Science, Telecommunication Building, Tongji University, No. 4800, Chao'an Highway, Jiading, Shanghai 201804, China. Tel.: +86 69589359.

E-mail address: [cessyzh@hotmail.com](mailto:cessyzh@hotmail.com) (D. Zhang).

or the storage size. For CBR encoding, the one-pass rate control is usually designed to improve the matching accuracy between the target bit rate and the actual bit rate and satisfy the low-latency and buffer constraints. The complexity of CBR rate control could not be too high so as to meet the requirement of real-time encoding. The fluctuation of video quality cannot be avoided for CBR rate control due to the varying content in natural scenes. For VBR encoding, the delay or rate constraint is not as strict as real-time video coding. So more complex rate control modules could be used to generate constant quality videos or improve coding performance. It is commonly employed with two-pass method.

Since the first rate control algorithm for H.264/AVC was proposed by Ma et al. [1] in 2002, a lot of improvement work of CBR rate control has been done for H.264/AVC real-time video applications. With the quadratic rate-quantizer (R-Q) model, Li et al. [2] proposed an adaptive basic-unit-layer rate control algorithm to optimize bit allocation according to the relative complexity of each unit and further adjust quantization parameters (QP). This algorithm can provide more accurate target bit matching and better buffer regulation and has been adopted by JVT in the H.264/AVC JM reference model. However, this algorithm has large quality fluctuation among frames since the quality consistency of adjacent frames is not considered. Based on the algorithm in [2], some real-time rate control techniques [3–12] are further proposed to reduce the control complexity, improve the control accuracy and enhance the visual quality for low-delay communications. By taking the advantage of the encoder buffer, smoothing video quality is possible for these algorithms as the buffer can tolerate limited bit rate fluctuation whilst the buffer does not either overflow or underflow. But these single-pass rate control algorithms cannot provide constant quality video services since they do not have sufficient information of the whole video sequence. These algorithms are more suitable for real-time applications in which the bit rate constraint is more important than the constant quality constraint. However, for off-line high-definition applications, constant quality video is more desirable and larger bit rate fluctuation of adjacent frames is tolerable. So researchers proposed to use two-pass VBR rate control techniques to optimize video quality for off-line video applications.

The basic principle of all the two-pass rate control algorithms is that in the whole or part of sequence, the first-pass encoding is employed with fixed quantization parameter value or CBR rate control method to obtain the information about the coding complexity of the frames. Then the model reflecting the coding complexity is built according to the first-pass statistics information and the quantization parameters can be adjusted correspondingly before or in the second-pass VBR encoding, so as to achieve higher coding performance or more consistent video quality. Westerink et al. [13] first introduced two-pass rate control algorithm for non-real-time applications based on MPEG-2 encoding. In their algorithm, the first-pass encoding gathers the quantization scale, spatial activity and temporal activity of each frame. Based on

these statistics and an R-Q model built by employing a perceptual experiment with several scenes, the target number of bits for each frame in the second-pass encoding is appropriately adjusted and quantization parameter can be obtained. Since considering the perceptual performance, this algorithm can achieve a constant visual quality. However, the visual perception experiments largely increase the complexity of the algorithm. For different resolution and frame-rate sequences, the new perceptual experiments could be needed to ensure accuracy of the R-Q model. In [14], for each frame of the first-pass encoding, the R-Q function is computed for all possible quantization factors. During the off-line processing, three optimal quantization factors for I-, P- and B-frames are first determined based on the MPEG2 TM5 rate control model. Then the quantization factor for each frame is adjusted based on the R-Q function of each frame. With adjusted quantization factors, the second-pass encoding is executed to generate the final VBR output bit stream. This algorithm can build efficient R-Q function and obtain more consistent video quality compared to TM5 CBR rate control scheme. However, this algorithm may cause quite large bit control error. Since all quantization factors for second-pass encoding are determined in advance, the encoder could not make adjustment according to the actual status of bits consumption in the second-pass encoding. So the large mismatch between the actual bit rate and the target bit rate would occur. Lie et al. [15] used a window-based two-pass rate control mechanism for the whole sequence. In each window segment which usually includes several frames, with the first-pass statistics, the content characteristics of each frame can be described by the models of rate vs. Lagrange multiplier and distortion vs. Lagrange multiplier. Then the second-pass encoding will adjust the QP value for each MB according to these models. This algorithm can obviously improve the coding efficiency and bit control accuracy, but only slightly smoother video quality can be achieved compared to JM CBR rate control. Since the two-pass encoding is employed in a limited window, only the content characteristics of several future encoded frames can be obtained for second-pass encoding, which makes the consistent video quality for the whole sequence is impossible because the encoder could not appropriately allocate the target coding bits according to the characteristic information of the whole sequence. Similarly, the GOP-level based two-pass rate control algorithm developed by Kwon et al. [16] can obtain higher coding efficiency, lower bit control error and smoother video quality compared to JM CBR rate control. But it could not realize the constant quality video coding for the whole sequence. In [17], Que et al. used the statistics information of first-pass CBR encoding to more appropriately allocate target bits and determine quantization parameters of all frames based on R-Q model before starting the second-pass encoding. Even though the experimental results show that this algorithm can achieve a constant video quality through the whole sequence for CIF sequences with comparable bit control accuracy to JM CBR rate control, it may suffer large bit control error for a higher resolution video encoded at a higher bit rate because it has no mechanism to monitor the bit control

status in the second-pass encoding and just uses the preset quantization parameter for each frame. The latest two-pass rate control algorithm, proposed by Huang et al. in [18], built the rate distortion (R-D) model based on the statistical analysis of the first-pass integer transform coefficients with generalized Gaussian distributions. Then the optimized quantization parameters for all frames would be determined by recursive calculation during the off-line processing. This algorithm improves video quality consistency of the whole sequence. However, it suffers from large bit rate mismatch, since it determines the quantization parameters for the second-pass with their R-D model obtained from the first-pass only and the bit budget constraint is neglected in the second-pass encoding. Therefore, the bit rate mismatch of each frame will propagate without any control mechanism implemented according to the actual status of the second-pass encoding.

In this paper, we propose a novel two-pass rate control algorithm for constant quality H.264/AVC high definition video coding based on our previous work [19,20]. Based on the prototype of our rate control method proposed in [19], we further optimized the algorithm including scene change detection, adaptive adjustment of the factor of R-D model according to the actual coding status of the second-pass encoding, the building of the D-Q model with lower computation complexity, and the parameter updating of the D-Q model at the scene change. In addition, a lot of comparison tests have been done to prove our two-pass algorithm is superior to the latest two-pass rate control algorithm [18] for H.264 HDTV encoding. In the first-pass encoding, all the frames are encoded with a fixed QP value. Then the rate and distortion information is collected to build the scene complexity model and detect the scene change frames. In the second-pass encoding, a GOP-level bit allocation scheme is first designed to adjust the target bit rate of each GOP based on the scene complexity of the GOP. Then with the built complexity model, the encoder can determine the expected distortion which could be achieved in the second-pass encoding under the target bit rate. According to the built linear distortion-quantizer (D-Q) model, before encoding one frame, the quantization parameter can be solved to realize constant quality encoding. Different from the latest constant quality two-pass rate control algorithm [18], our algorithm determines the second-pass quantization parameter of each frame according to the actual coding status. The encoder has a GOP-level-based rate control mechanism for second-pass VBR encoding to monitor the control error. When the GOP bit allocation is adaptively adjusted according to the overdue coding bits of the previous GOP, the expected distortion of this GOP, which is used to determine the QP value of each frame, is correspondingly adjusted. So the QP value of each frame in this GOP can be adjusted accordingly. As a result, the coding bit rate can be limited in a small error range. In addition, after encoding one frame, the D-Q model parameters will be updated with linear regression method to ensure the prediction accuracy of the quantization parameter of next encoded frame with the same coding type. The effect of scene change on the updating of D-Q

model is also considered. The model will be re-initialized at the scene change to minimize modeling error.

The rest of this paper is organized as follows. In Section 2, we describe the proposed two-pass rate control algorithm. The first-pass coding which is to build the scene complexity model and detect the scene complexity change is first presented. Then the GOP-level bit allocation and frame-level quantizer control of the second-pass coding are described. Section 3 presents and discusses the experimental results of comparison tests between our algorithm and the latest constant quality rate control algorithm [18]. Finally, Section 4 concludes this paper.

## 2. Two-pass rate control algorithm

Fig. 1 presents the flowchart of our proposed constant quality rate control algorithm. In the first-pass encoding, all the frames are first encoded with a fixed QP value. Then the characteristic information of each coded frame is collected to build its scene complexity model and the scene changes will be detected with the complexity information of the whole sequence. The complexity of each frame and the scene change information will be passed to the second-pass encoding. In the second-pass encoding, the GOP-level bit allocation is first adjusted according to the average scene complexity of each GOP. Then the expected average distortion of each GOP under the constraint of the target coding bits is determined with the first-pass complexity statistics. According to the built linear D-Q model, before encoding one frame, the quantization parameter can be solved to realize constant quality encoding. After encoding one P-/B-frame, the model parameters will be updated with linear regression method. In addition, in the process of updating the model parameters, the effect of scene change on the updating of D-Q model is considered. The model will be re-initialized at the scene change to minimize modeling error. The newly updated model is more accurate for providing the relationship between the quantization parameter and distortion. With the corresponding quantization parameter computed by the model, each frame is encoded to generate constant quality bit stream.

### 2.1. The first-pass encoding

#### 2.1.1. Information collection

In the first-pass encoding, the video frames are first encoded using a fixed QP value. The fixed QP value is recommended to be a value from 16 to 30. The coding parameters, such as GOP size and the number of reference frames, are set to be the same as those in the second-pass encoding. The statistics of each frame are collected, which include the rate (i.e., the coding bits for texture and motion information) and the distortion (as measured by MSE). These statistics will be used to model the scene complexity of each frame.

#### 2.1.2. Frame scene complexity modeling

Considering that the typical rate distortion characteristics can be modeled as the inverse-proportion

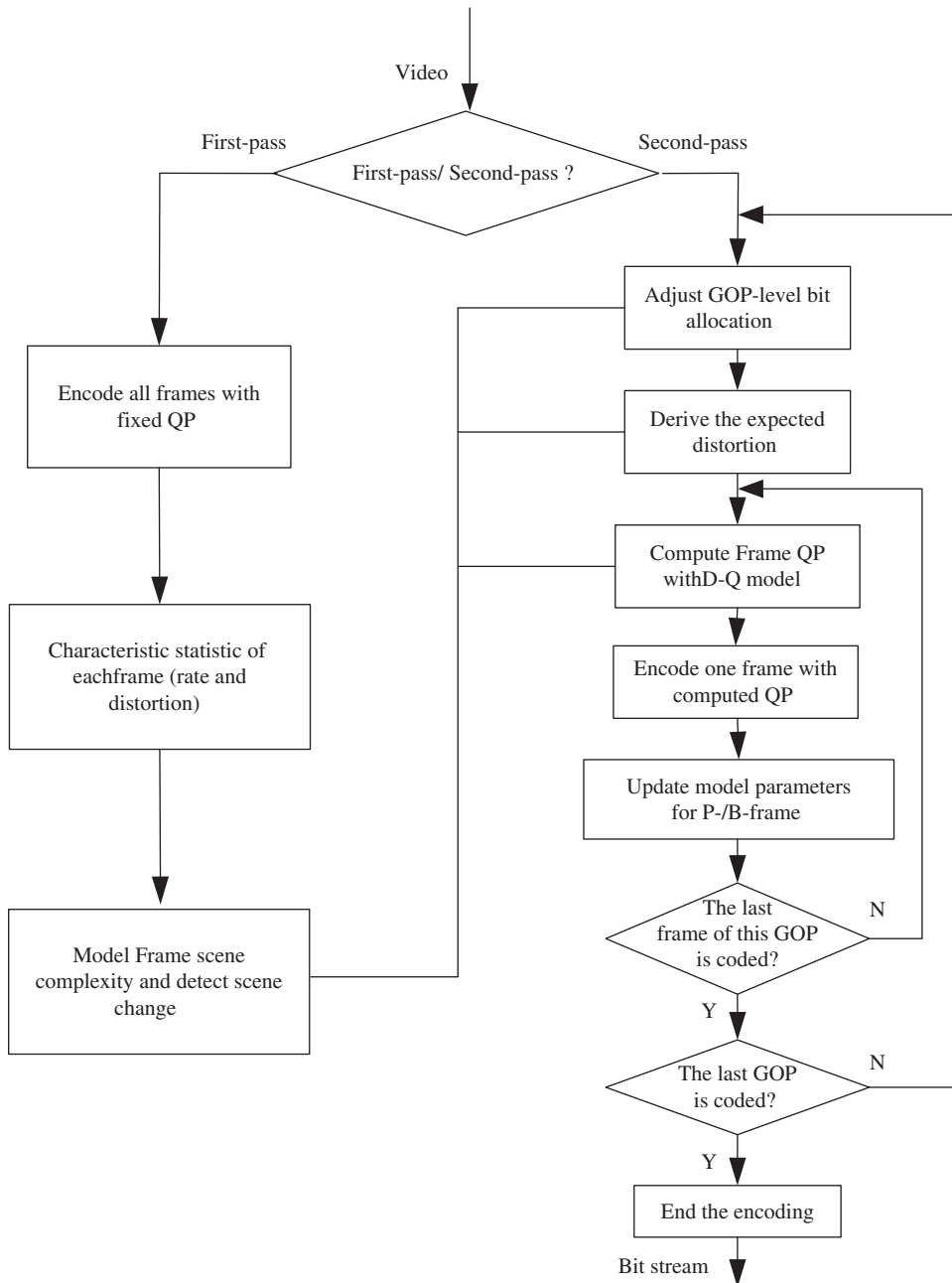


Fig. 1. The flowchart of proposed two-pass rate control algorithm.

relationship, we adopt the product of rate and distortion of each frame to approximately model the scene complexity of each frame as follows [21]:

$$C_i = D_i R_i \quad (1)$$

where,  $C_i$ ,  $D_i$  and  $R_i$  are the scene complexity, the MSE distortion and the bit rate of the  $i$ th frame, respectively. The larger the value of  $C_i$ , the higher is the scene complexity.

### 2.1.3. Scene change detection

With the scene complexity information, we can easily detect the scene changes of the whole sequence. Fig. 2 shows an example of the scene complexity distribution of the sequence named *Comb*, which is composed of three different sequences: 100 frames of *Bigship* (moderate motion), *Night* (fast Motion) and *Jet* (low motion). The data is in display order. The coding structure is IBBP and the GOP length is 30 frames. From this figure, we can find that there is quite a large scene complexity difference

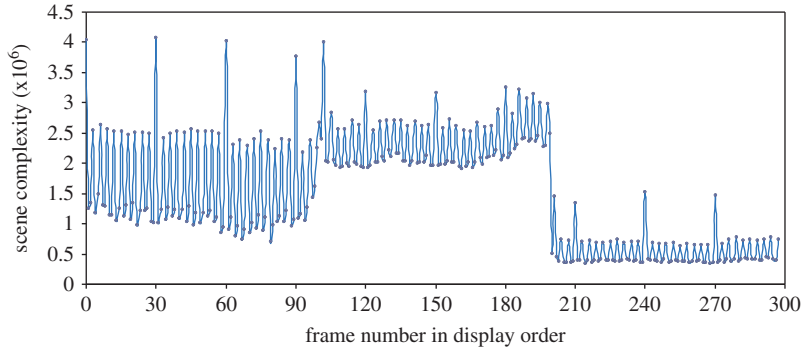


Fig. 2. The scene complexity of the *Comb* sequence with three scenes.

in I-frames, P-frames and B-frames among the different segments. So we can use the complexity difference of the frames with the same coding type to detect the scene changes. Since scene change has no effect on the coding of I-frame, we only need to judge whether a P/B- frame is a scene-change frame.

We define the complexity difference of the  $i$ th frame as

$$\Delta C_i = |C_i - C_{prev}| \quad (2)$$

In (2),  $C_{prev}$  is the scene complexity of the previous P-frame if the  $i$ th frame is a P-frame. If the  $i$ th frame is a B-frame,  $C_{prev}$  is the scene complexity of the previous B-frame at the similar coding position to the  $i$ th frame. That is, if there are two successive B-frames between adjacent P-frames, the complexity difference of these two B-frames should be computed with the scene complexity of the previous two successive B-frames. To detect the scene change frames, we simply use the following expression.

$$\Delta C_i > \theta \Delta C_{Aver} \quad (3)$$

where  $\Delta C_{Aver}$  is the average value of the complexity difference of all the coded frames with same coding type, and  $\theta$  is a constant with a typical value of 7.0. It is noted that the scene change detection is to detect the scene complexity change to achieve constant quality coding. Different scene with same coding complexity is not considered in the scene change detection since it will not affect the coding result.

## 2.2. The second-pass encoding

In the second-pass encoding, we first compute the average scene complexity of each GOP and adjust the GOP-level bit allocation accordingly. Then the expected distortion of each frame will be derived with the complexity of each frame, and the quantization step (Q-step) of each frame can be resolved according to the linear relationship between the distortion and the Q-step.

### 2.2.1. GOP-level bit allocation adjustment

In our algorithm, two aspects are first considered in the GOP-level bit allocation of the second-pass encoding. One aspect is that the number of coding bits for each GOP should be determined according to the scene complexity

of each GOP. If the scene complexity of one GOP is lower than the average scene complexity of all GOPs, its allocated bits should be less than the average number of bits allocated to each GOP. The other aspect to be considered is how to compensate for the difference between the actual coding bits and the budget bits of one encoded GOP. Here, we apportion the difference to all the future GOPs to be encoded.

Based on the above consideration, the target bits of each GOP can be adjusted as follows:

$$T_j = (C_j^{GOP} / C_{Aver}^{GOP}) T_j^a - \sum_{k=0}^{j-1} (B_k / (M - k - 1)) \quad (4)$$

The first part in the right side of the above expression means the number of bits that should be allocated to the  $j$ th GOP according to its scene complexity, where  $T_j^a$  is the number of average target bits of each GOP obtained by equal allocation. It can be computed with second-pass target bit rate  $R_t$ , frame rate  $f_r$ , and the number of frames in the GOP

$$T_j^a = (N_j - N_{j-1})(R_t / f_r) \quad (5)$$

here,  $N_j$  and  $N_{j-1}$  denote the number of coded frames after the  $j$ th and  $(j-1)$ th GOP encoding, respectively.  $C_j^{GOP}$  is the scene complexity of the  $j$ th GOP. It is defined as:

$$C_j^{GOP} = \frac{1}{N_j - N_{j-1}} \cdot \sum_{i=N_{j-1}}^{N_j-1} C_i \quad (6)$$

$C_{Aver}^{GOP}$  is the average GOP complexity of the whole sequence computed by

$$C_{Aver}^{GOP} = \frac{1}{M} \sum_{j=0}^{M-1} C_j^{GOP} \quad (7)$$

Here,  $M$  is the number of GOPs in the whole sequence.

The second part in the right side of (2) denotes the total number of bits that needs to be compensated in the  $j$ th GOP. Here,  $B_k$  is the difference between the actual coding bits and the target bits of the  $k$ th GOP. If the actual coding bits of the  $(j-1)$ th GOP is used as reference, the target bits of the  $j$ th GOP can be finally adjusted as

$$\tilde{T}_j = (1 - \alpha) \frac{C_j^{GOP}}{C_{j-1}^{GOP}} \frac{N_j - N_{j-1}}{N_{j-1} - N_{j-2}} T_{j-1}^c + \alpha T_j \quad (8)$$

Here,  $T_{j-1}^c$  is the number of the actual coding bits of the  $(j-1)$ th GOP.  $\alpha$  is a constant with the typical value of 0.8.

### 2.2.2. Frame-level QP calculation

For off-line video applications, at a given bit budget, the encoder is expected to provide constant quality except for the high coding efficiency. Since the scene complexity of each frame may be different, the QP of each frame should be adjusted frame by frame to achieve constant quality. In CBR JM rate control algorithm, the Q-step of P-frame is obtained by solving the quadratic R-Q model and the QP value of B-frame is interpolated by the QP values of its adjacent P-frames. This algorithm can effectively control the bit rate of each frame, but the quality fluctuation among frames is large since the quality consistency of adjacent frames is not considered. The JM algorithm is suitable for real-time applications in which the bit rate constraint is more important than the constant quality constraint. However, for off-line high-definition applications, constant quality video is more desirable and larger bit rate fluctuation of adjacent frames is tolerable. In this paper, we consider the two-pass VBR rate control method. In our algorithm, we try to obtain the expected average distortion of each GOP with the first-pass complexity information and the second-pass target bit rate. Then, we adopt linear D-Q model to compute the Q-step of a P/B-frame. Since the expected distortion of each frame in one GOP is same and the expected distortion among GOPs is closer so that encoding the frame with the computed Q-step can achieve constant quality among the whole sequence.

**2.2.2.1. Derivation of the expected distortion.** If the first-pass QP value is selected to make the actual bit rate as close as possible to that of the second-pass encoding, the scene complexity of each GOP for these two-pass coding can be assumed to be close. So we have

$$\sum_{i=N_{j-1}}^{N_j-1} R_i^{2nd} D_i^{2nd} = k_j \sum_{i=N_{j-1}}^{N_j-1} R_i^{1st} D_i^{1st}, \quad j = 0, \dots, M-1 \quad (9)$$

here,  $R_i^{2nd}$  and  $D_i^{2nd}$  are the expected coding bits and distortion of the  $i$ th frame in the second-pass encoding.  $R_i^{1st}$  and  $D_i^{1st}$  are the corresponding rate and distortion obtained by the first-pass encoding.  $N_j$  and  $N_{j-1}$  denote the number of coded frames after the  $j$ th and  $(j-1)$ th GOP encoding, respectively.  $k_j$  is a scale factor. The value of  $k_j$  is adaptively adjusted according to the ratio of the actual coding bit rate and the predicted one of the  $(j-1)$ th GOP.

$$k_j = \begin{cases} \min(k_{j-1}, k_0) + \Delta k_j, & \text{if } T_{j-1}^c > 1.05 \tilde{T}_{j-1} \\ \max(k_{j-1}, k_0) - \Delta k_j, & \text{if } T_{j-1}^c < 0.95 \tilde{T}_{j-1} \\ k_0, & \text{else} \end{cases} \quad (10)$$

where  $k_0$  is the initial value of  $k_j$  and is empirically set to be a value between 1.0 and 1.2.  $\Delta k_j$  is an adjustment constant with a typical value of 0.05.

We further assume that all the frames in the  $j$ th GOP achieve almost the same distortion  $D_c^j$  under the constraint of the target coding bits  $\tilde{T}_j$  in the second-pass

encoding. We have

$$D_c^j \sum_{i=N_{j-1}}^{N_j-1} R_i^{2nd} = D_c^j \tilde{T}_j = k_j \sum_{i=N_{j-1}}^{N_j-1} R_i^{1st} D_i^{1st} \quad (11)$$

So the expected average frame distortion of this GOP can be solved as follows:

$$D_c^j = k_j \sum_{i=N_{j-1}}^{N_j-1} R_i^{1st} D_i^{1st} / \tilde{T}_j \quad (12)$$

To control the potential quality fluctuation among different GOPs, the expected distortion of the previous GOP will be used to further adjust the expected distortion  $D_c^j$  as follows:

$$\tilde{D}_c^j = \beta D_c^j + (1 - \beta) D_c^{j-1} \quad (13)$$

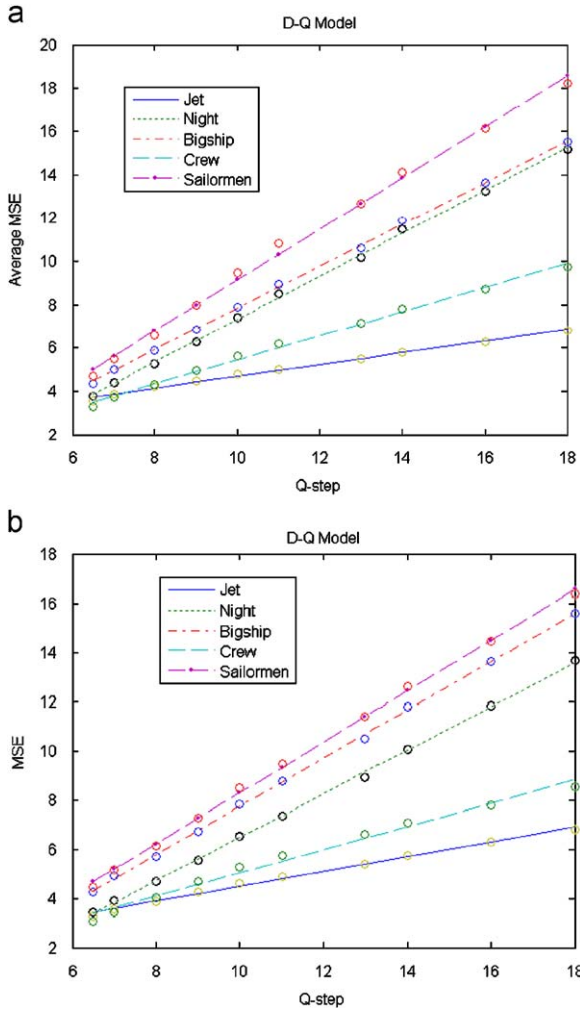
Here,  $\beta$  is a constant and its typical value is 0.6.

**2.2.2.2. The QP calculation of regular P- and B-frames.** To determine the QP value of one P/B-frame according to its expected distortion, we build the relationship between the distortion and the Q-step. Two types of D-Q models have been investigated to describe the relationship between the distortion and the Q-step. One is quadratic model [8,9] for CIF and QCIF sequences. The other is the linear model with intercept equal to zero for QCIF sequences [6]. However, according to the results of our extensive experiments for H.264/AVC high definition video, the relationship between the distortion and the Q-step fits well in a linear model in which the intercept usually is not equal to zero. Fig. 3(a) shows the examples of the relationship between the average MSE distortion of the whole sequence and the Q-step. Fig. 3(b) shows the examples of the relationship between the MSE distortion of the first P-frame and the Q-step. Here, each frame in the test sequence was encoded using a fixed QP value, and the chosen QP values were 20–29. After encoding, linear regression was applied to fit the obtained average MSE values of the whole sequence and Q-steps to study the relationship between them. For all the tests, the coding structure was IBBP, the motion estimation used the UMHExagonS mode with search range 48, the number of reference frames was 2, RD Optimization was enabled, and the number of coded frames was 300. From these two figures, we can find that the relationship between the MSE and the Q-step of one frame can be built with a linear model. The same relationship can also be found between the average MSE distortion of the whole sequence and the Q-step.

So we use the following linear D-Q model to calculate the Q-step of one B- or P-frame of the  $j$ th GOP

$$D_F = \tilde{D}_c^j = X_F \cdot Q_F + Y_F, \quad F \in (P, B) \quad (14)$$

After encoding a P/B-frame, the parameters  $X_P/X_B$  and  $Y_P/Y_B$  of model (14) are updated with linear regression method to ensure the prediction accuracy of the quantization parameter of next encoded frame with the same coding type.



**Fig. 3.** The relationship between the MSE distortion and Q-step: (a) the relationship between the average MSE distortion of the whole sequence and Q-step and (b) the relationship between the MSE distortion of the first P frame and Q-step.

To maintain the smoothness of the quality between the successive frames, the QP value derived from the corresponding Q-step is further adjusted by

$$\tilde{Q}P_F = \min\{QP_{prev} + 2, \max\{QP_{prev} - 2, QP_F\}\} \quad (15)$$

here,  $QP_{prev}$  is the QP of the previous frame which has the same coding type as current frame.

**2.2.2.3. The QP calculation of i-frames and the first P- and B-frames of each scene.** Since the distance between two I-frames is usually very large, their correlation will be very small. So we could not use a model like (14) to recursively obtain the QP value of each I-frame for the second-pass encoding. For the first P- and B-frames of each scene, the model of (14) is also not usable since we could not have information about the model parameters. Here, we decide the QP value of these frames with the following experi-

mental relationship between the first-pass QP and the second-pass one

$$QP^{2nd} = \left( \frac{\tilde{D}_c^j}{D^{1st}} \right)^{1/4} (QP^{1st} + \gamma) \quad (16)$$

here,  $QP^{1st}$  is the fixed QP value for the first-pass encoding.  $D^{1st}$  is the MSE distortion of each I-frame or the first P-/B-frame of each scene in the first-pass encoding and  $\gamma$  is a constant. When the first-pass actual bit rate is close to the second-pass one, the value of  $\gamma$  is set to 0. Otherwise, its value is experimentally set to be 1 or 2.

**2.2.2.4. The parameter update of the D-Q model at the beginning of a new scene.** When a scene change is detected, the window for updating the model parameters  $X_p/X_b$  and  $Y_p/Y_b$  needs to be re-initialized. If the update window is not re-initialized at the beginning of the new scene, the distortion and Q-step values of some last frames in the previous scene will be used to update the model parameters of the first frames in the new scene. This will make the linear regression method to exclude the data of the new scene as outliers when computing the update parameters of the first frames in the new scene. This is not right because the parameter update of a frame is generally more correlated to the future frames in the same scene. Fig. 4 shows the coding results of *Bigship* sequence, which has a scene change at the 202<sup>nd</sup> frame, for updating the model parameters with and without window re-initialization at the scene change with the same coding configurations. From this figure, we can see that there is obvious PSNR fluctuation up to about 2 dB at the scene change without the window re-initialization. In addition, for all the frames in the new scene, smoother PSNR performance can be achieved with the window re-initialization. The reason is that the encoder views the data of the previous scene as outliers and excludes them out of the update window by re-initializing the update window at the scene change so that the model can obtain more accurate updated parameters which makes the model converge faster than not re-initializing the window.

### 3. Experimental results

Based on JM11.0 encoder [22], we have compared our two-pass rate control algorithm against the latest two-pass rate control algorithm—Huang's method [18] with various sequences. The test sequences are 720p (1280 × 720) or 1080p (1920 × 1080) and encoded at frame rate 30 fps. The GOP size is 30 frames, the coding structure is IBBPBB; the motion estimation uses the UMHexagonS mode with the search range of 48 for 720p or 64 for 1080p; the number of reference frames is 2; and, the RD Optimization is enabled.

Table 1 shows the simulation results of various high definition video sequences, *Crew*, *Sailormen*, *Night*, *Jet*, *Bigship*, *Comb*, *Pedestrianarea* and *Rushhour*. Each sequence is encoded at four different bit rates. We use PSNR variance of the whole sequence to evaluate the quality consistency of one encoded sequence. Average PSNR of the whole sequence is used to measure the coding efficiency.

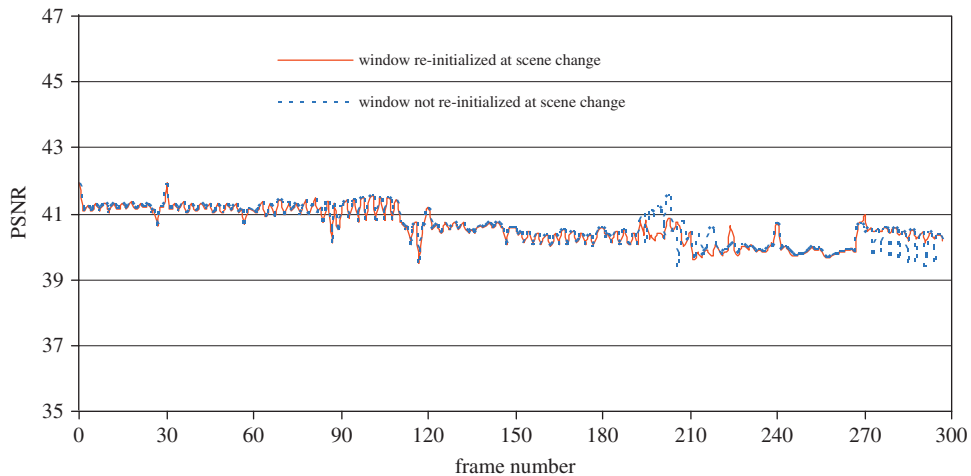


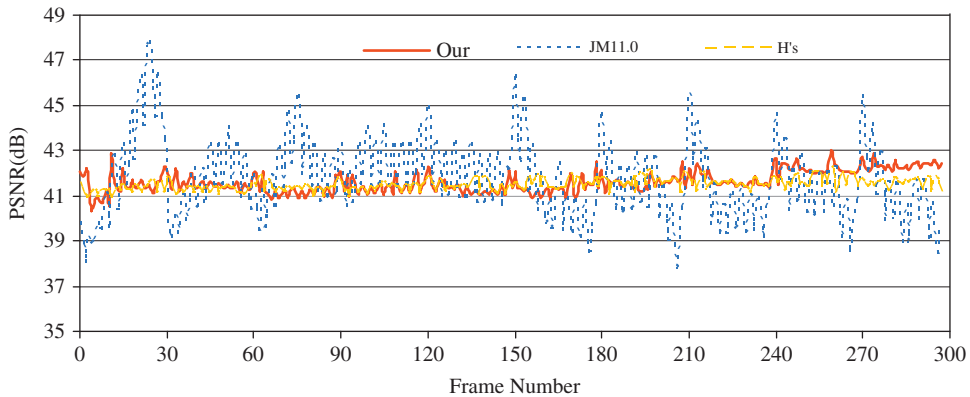
Fig. 4. The coding results for updating the model parameters with and without window re-initialization at the same coding configurations.

The bit control error denotes the bit control accuracy of each algorithm for different videos, which is computed by dividing the target bit rate into the difference between the target bit rate and the actual bit rate. A positive control error means that the encoder uses too many coding bits and the actual bit rate overruns the limit of the target bit rate. Whereas a negative error value expresses that the actual bit rate is less than the target bit rate and there are some surplus coding bits after encoding. The average control errors in the table are the average values of the absolute control errors of all the coded sequences for different algorithms, respectively. From this table, we can see that both our two-pass rate control algorithm (Ours) and Huang's algorithm (H's) can significantly reduce the PSNR variance, compared to JM11.0 rate control algorithm (JM). The highest improvement of our algorithm is 98.7% for *Bigship* sequence at bit rate 12 Mbps. The highest improvement of Huang's algorithm is 99.4% for *Bigship* sequence at bit rate 14 Mbps. On average, our algorithm and Huang's algorithm have 89.0% and 97.6% improvement of PSNR variance reduction, respectively. Even though our algorithm has marginal loss in terms of PSNR variance compared to Huang's algorithm, our algorithm can achieve higher bit control accuracy. By comparing the bit control errors of our algorithm with Huang's method, we found that among all the 32 encoded videos, only three encoded videos have absolute control errors over 1% for our algorithm and the highest control error is 1.94%. However, for Huang's method, only four encoded videos have bit control error less than 1%, and there are eleven encoded videos with absolute control error over 5%. The highest absolute control error is up to 25.9%. On average, the bit control error of our algorithm is 0.58%, and the bit control error of Huang's method is 5.73%. The improvement of bit control accuracy of our algorithm is about 90%. It is obvious that the Huang's method is not suitable for HDTV encoding due to its low bit rate control accuracy. The reason for the lower bit control accuracy of Huang's method is that it uses the preset QP for each frame encoding and has no mechanism to monitor the bit

control status in the second-pass encoding. However, in our algorithm, the encoder has a GOP-level-based rate control mechanism for second-pass VBR encoding to monitor the control error. The QP for each frame is determined according to the actual coding status in the second-pass encoding. In addition, after encoding one frame, the parameters of D-Q model, which is used to compute the QP value, will be updated with linear regression method to ensure the prediction accuracy of the quantization parameter of next encoded frame with the same coding type. These rate control modules make our algorithm to achieve higher bit control accuracy than Huang's algorithm. When comparing our algorithm to JM rate control algorithm, we can find that JM rate control algorithm has more accurate bit control than our algorithm. However, we can find that the bit rate control error of our algorithm is lower than 2% in the worst case, and on average the bit control errors of our algorithm and JM rate control algorithm are 0.58% and 0.16%, respectively, which are comparable. That is because our algorithm considers the GOP-level bit allocation adjustment in the second-pass encoding and further determines the QP value for each frame with the expected distortion derived under the GOP target bit rate. At the beginning of one GOP, the expected distortion of this GOP is adaptively adjusted according to the overdue coding bits of the previous GOP. So the QP value of each frame in this GOP can be adjusted accordingly. As a result, the coding bit rate can be limited in a small error range even though there is no direct frame-level bit control in our algorithm. More complex bit rate control of frame level could be included in the proposed algorithm to achieve more accurate bit rate control. However, it will introduce higher computational complexity and cause larger quality fluctuation so that it is not considered in this paper. Our rate control algorithm and Huang's algorithm achieve improved average PSNR over JM rate control algorithm on average by 0.05 dB and 0.09 dB, respectively. However, it is noted that some improvements of Huang's method are achieved by increased bit rate.

**Table 1**  
The simulation results.

| Sequence            | Target bit rate (Mbps) | PSNR variance |      |      |                            |      | Average PSNR (dB) |       |       |                 |                       | Actual bit rate (Mbps) |       |       | Bit control error (%) |        |       |
|---------------------|------------------------|---------------|------|------|----------------------------|------|-------------------|-------|-------|-----------------|-----------------------|------------------------|-------|-------|-----------------------|--------|-------|
|                     |                        | JM            | H's  | Ours | Improvement against JM (%) |      | JM                | H's   | Ours  | Gain against JM |                       | JM                     | H's   | Ours  | JM                    | H's    | Ours  |
|                     |                        |               |      |      | H's                        | Ours |                   |       |       | H's             | Ours                  |                        |       |       |                       |        |       |
|                     |                        |               |      |      |                            |      |                   |       |       |                 |                       |                        |       |       |                       |        |       |
| Crew                | 8                      | 2.84          | 0.07 | 0.23 | 97.5                       | 91.9 | 41.47             | 41.47 | 41.61 | 0               | 0.14                  | 8.02                   | 7.65  | 8.08  | 0.25                  | -4.38  | 1     |
|                     | 10                     | 3.32          | 0.06 | 0.10 | 98.2                       | 97.0 | 42.10             | 42.36 | 42.26 | 0.26            | 0.16                  | 10.04                  | 10.33 | 9.97  | 0.4                   | 3.3    | -0.3  |
|                     | 12                     | 4.13          | 0.05 | 0.20 | 98.8                       | 95.2 | 42.65             | 42.71 | 42.88 | 0.06            | 0.23                  | 12.04                  | 11.44 | 12.03 | 0.33                  | -4.67  | 0.25  |
|                     | 14                     | 4.86          | 0.03 | 0.16 | 99.4                       | 96.7 | 43.16             | 43.47 | 43.40 | 0.31            | 0.24                  | 14.06                  | 14.45 | 14.14 | 0.43                  | 3.21   | 1     |
| Sailormen           | 8                      | 1.84          | 0.05 | 0.14 | 97.3                       | 92.4 | 37.83             | 38.06 | 37.76 | 0.23            | -0.07                 | 8.00                   | 8.81  | 7.96  | 0                     | 10.13  | -0.5  |
|                     | 10                     | 2.27          | 0.09 | 0.17 | 96.0                       | 92.5 | 38.40             | 38.28 | 38.42 | -0.12           | 0.02                  | 10.01                  | 9.44  | 9.96  | 0.1                   | -5.6   | -0.4  |
|                     | 12                     | 2.58          | 0.07 | 0.11 | 97.3                       | 95.7 | 38.92             | 38.79 | 38.93 | -0.13           | 0.01                  | 12.00                  | 11.49 | 12.10 | 0                     | -4.25  | 0.83  |
|                     | 14                     | 2.86          | 0.04 | 0.21 | 98.6                       | 92.7 | 39.41             | 39.59 | 39.41 | 0.18            | 0                     | 14.01                  | 14.83 | 14.04 | 0.07                  | 5.93   | 0.29  |
| Night               | 8                      | 1.80          | 0.05 | 0.11 | 97.2                       | 93.9 | 38.99             | 39.06 | 38.72 | 0.07            | -0.27                 | 8.01                   | 8.54  | 8.02  | 0.13                  | 6.75   | 0.25  |
|                     | 10                     | 2.26          | 0.08 | 0.18 | 96.5                       | 92.0 | 39.70             | 39.55 | 39.67 | -0.15           | -0.03                 | 10.01                  | 9.74  | 10.01 | 0.1                   | -2.6   | 0.1   |
|                     | 12                     | 2.72          | 0.10 | 0.21 | 96.3                       | 92.3 | 40.33             | 40.34 | 40.22 | 0.01            | -0.11                 | 12.00                  | 12.48 | 11.95 | 0                     | 4      | -0.42 |
|                     | 14                     | 3.42          | 0.09 | 0.18 | 97.4                       | 94.7 | 40.89             | 40.97 | 40.91 | 0.08            | 0.02                  | 14.01                  | 14.32 | 14.01 | 0.07                  | 2.29   | 0.07  |
| Jet                 | 8                      | 2.36          | 0.03 | 0.10 | 98.7                       | 95.8 | 43.21             | 43.29 | 43.50 | 0.08            | 0.29                  | 8.01                   | 6.90  | 8.01  | 0.13                  | -13.75 | 0.13  |
|                     | 10                     | 3.05          | 0.04 | 0.09 | 98.7                       | 97.0 | 43.56             | 43.39 | 43.84 | -0.17           | 0.28                  | 10.02                  | 7.41  | 9.99  | 0.2                   | -25.9  | -0.1  |
|                     | 12                     | 4.26          | 0.05 | 0.15 | 98.8                       | 96.5 | 43.90             | 44.19 | 44.17 | 0.29            | 0.27                  | 12.01                  | 12.10 | 12.02 | 0.08                  | 0.83   | 0.17  |
|                     | 14                     | 5.13          | 0.05 | 0.16 | 99.0                       | 96.9 | 44.23             | 44.18 | 44.47 | -0.05           | 0.24                  | 14.00                  | 11.99 | 13.91 | 0                     | -14.36 | -0.64 |
| Bigship             | 8                      | 3.14          | 0.10 | 0.32 | 96.8                       | 89.8 | 40.17             | 40.06 | 39.89 | -0.11           | -0.28                 | 8.01                   | 8.31  | 7.96  | 0.13                  | 3.88   | -0.5  |
|                     | 10                     | 3.60          | 0.09 | 0.28 | 97.5                       | 92.2 | 40.76             | 40.92 | 40.60 | 0.16            | -0.16                 | 10.01                  | 11.00 | 9.94  | 0.1                   | 10     | -0.6  |
|                     | 12                     | 3.97          | 0.05 | 0.05 | 98.7                       | 98.7 | 41.30             | 41.29 | 41.30 | -0.01           | 0                     | 12.01                  | 11.88 | 11.91 | 0.08                  | -1     | -0.75 |
|                     | 14                     | 4.70          | 0.03 | 0.13 | 99.4                       | 97.2 | 41.76             | 42.06 | 41.79 | 0.3             | 0.03                  | 14.03                  | 14.98 | 13.99 | 0.21                  | 7      | -0.07 |
| Comb                | 8                      | 6.80          | 0.05 | 0.36 | 99.3                       | 94.7 | 40.69             | 40.81 | 40.77 | 0.12            | 0.08                  | 8.01                   | 8.25  | 8.04  | 0.13                  | 3.13   | 0.5   |
|                     | 10                     | 6.21          | 0.09 | 0.30 | 98.6                       | 95.2 | 41.31             | 41.66 | 41.50 | 0.35            | 0.19                  | 10.01                  | 10.84 | 10.06 | 0.1                   | 8.4    | 0.6   |
|                     | 12                     | 5.49          | 0.06 | 0.33 | 98.9                       | 94.0 | 41.86             | 42.04 | 42.11 | 0.18            | 0.25                  | 12.00                  | 11.95 | 12.11 | 0                     | -0.42  | 0.92  |
|                     | 14                     | 4.59          | 0.08 | 0.29 | 98.3                       | 93.7 | 42.41             | 42.59 | 42.60 | 0.18            | 0.19                  | 14.01                  | 13.91 | 14.02 | 0.07                  | -0.65  | 0.14  |
| Pedestrian area     | 12                     | 1.36          | 0.05 | 0.31 | 96.3                       | 77.2 | 42.00             | 42.20 | 41.89 | 0.2             | -0.11                 | 12.03                  | 12.75 | 12.06 | 0.25                  | 6.25   | 0.5   |
|                     | 14                     | 0.99          | 0.04 | 0.24 | 96.0                       | 75.8 | 42.26             | 42.54 | 42.37 | 0.28            | 0.11                  | 14.04                  | 15.82 | 13.86 | 0.29                  | 13     | -1    |
|                     | 16                     | 1.23          | 0.03 | 0.30 | 97.6                       | 75.6 | 42.61             | 42.57 | 42.53 | -0.04           | -0.08                 | 16.04                  | 16.19 | 16.05 | 0.25                  | 1.19   | 0.31  |
|                     | 18                     | 1.28          | 0.06 | 0.35 | 95.3                       | 72.7 | 42.80             | 42.77 | 42.80 | -0.03           | 0                     | 18.07                  | 17.90 | 18.35 | 0.39                  | -0.56  | 1.94  |
| Rushhour            | 12                     | 0.70          | 0.02 | 0.33 | 97.1                       | 52.9 | 42.74             | 42.97 | 42.78 | 0.23            | 0.04                  | 12.02                  | 12.21 | 11.81 | 0.17                  | 1.75   | -1.58 |
|                     | 14                     | 0.63          | 0.05 | 0.28 | 92.1                       | 55.6 | 42.97             | 43.12 | 43.00 | 0.15            | 0.03                  | 14.05                  | 14.45 | 13.96 | 0.36                  | 3.21   | -0.29 |
|                     | 16                     | 0.87          | 0.02 | 0.21 | 97.7                       | 75.9 | 43.13             | 43.29 | 43.19 | 0.16            | 0.06                  | 16.04                  | 16.52 | 15.86 | 0.25                  | 3.25   | -0.88 |
|                     | 18                     | 0.93          | 0.02 | 0.07 | 97.9                       | 92.5 | 43.30             | 43.29 | 43.40 | -0.01           | 0.10                  | 18.03                  | 16.62 | 17.73 | 0.17                  | -7.67  | -1.5  |
| Average improvement |                        |               |      |      | 97.6                       | 89.0 | Average gain      |       | 0.09  | 0.05            | Average control error |                        |       | 0.16  | 5.73                  | 0.58   |       |



**Fig. 5.** Frame PSNR comparison for Crew sequence at 8 Mbps bit rate.

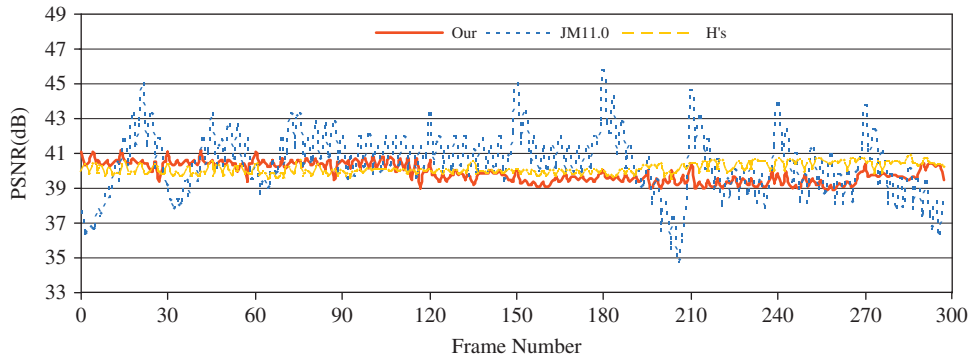


Fig. 6. Frame PSNR comparison for *Bigship* sequence at 8 Mbps bit rate.

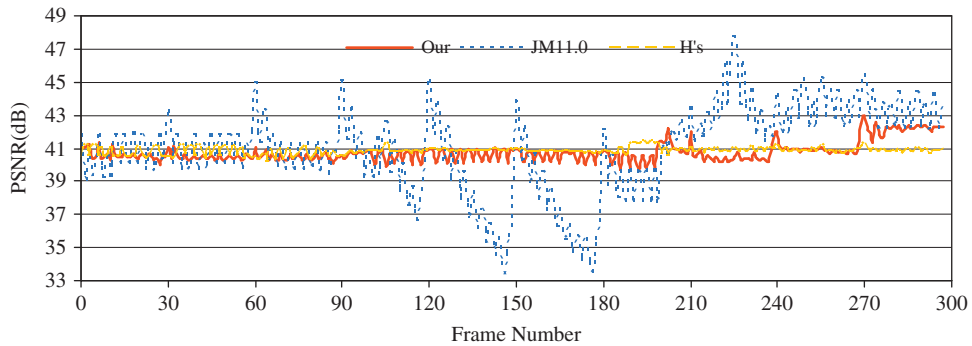


Fig. 7. Frame PSNR comparison for *Comb* sequence at 8 Mbps bit rate.

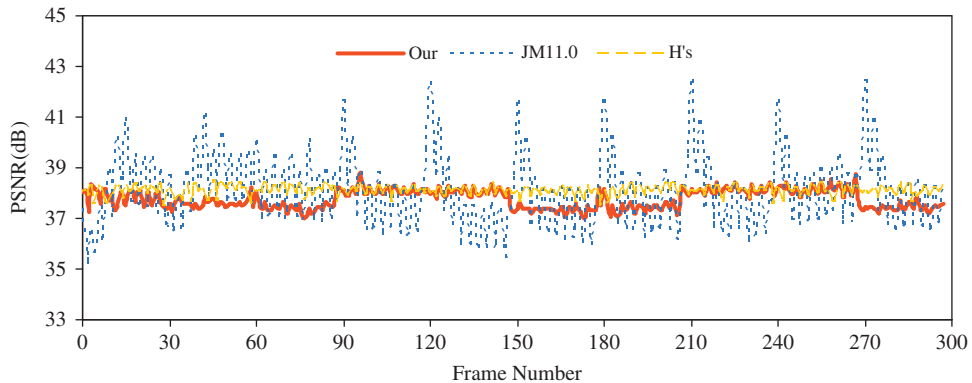


Fig. 8. Frame PSNR comparison for *Sailormen* sequence at 8 Mbps bit rate.

Figs. 5–8 compare the frame PSNRs of our rate control algorithm, Huang's rate control algorithm and JM rate control algorithm for *Crew*, *Bigship*, *Comb* and *Sailormen*, respectively. The significant reduction of the PSNR fluctuation can be seen from these figures for both our algorithm and Huang's algorithm. Since our algorithm and Huang's algorithm are two-pass rate control algorithm, the coding complexity information of the whole sequence can be obtained by first-pass encoding to adjust the QP values of the second-pass encoding and achieve constant quality video coding. These figures also show that Huang's

method has a little more consistent video quality than our algorithm. That is because there is a GOP-level bit allocation adjustment mechanism monitoring the bit rate control of each GOP in our algorithm. When the GOP bit allocation is adaptively adjusted according to the overdue coding bits of the previous GOP, the excepted distortion of this GOP, which is used to determine the QP value of each frame, is correspondingly adjusted, which may cause the fluctuation among the distortions of different GOPs. But it is negligible since the fluctuation among different GOPs is very small.

#### 4. Conclusions

A two-pass rate control algorithm has been proposed for constant quality H.264/AVC high definition video coding. With the first-pass characteristic statistics and the scene complexity model, the encoder can derive the expected distortion which could be achieved in the second-pass encoding under the adjusted target bit rate of each GOP. Further, when encoding one frame, the encoder can first determine the Q-step of the frame according to the built linear D-Q model. After one frame's encoding, the model parameters will be updated with linear regression method to ensure the prediction accuracy of the Q-step of next encoded frame with the same coding type. In addition, the D-Q model is re-initialized to minimize modeling error when scene change is detected. With the proposed two-pass rate control algorithm, a higher bit control accuracy than the latest two-pass rate control algorithm can be achieved at comparable constant quality and average PSNR.

#### Acknowledgements

The authors would like to thank Jianfei Huang and Jun Sun for providing the source code of their two-pass rate control algorithm for comparison tests in this paper. This work was partially supported by a grant from the Chinese University of Hong Kong Focused Investment Scheme (Project 1903003).

#### References

- [1] S. Ma, W. Gao, Y. Lu, Rate control on JVT standard, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-D030, January 2002.
- [2] Z.G. Li, F. Pan, K.P. Lim, G. Feng, X. Lin, S. Rahardja, Adaptive basic unit layer rate control for JVT, JVT of ISO/IEC125 MPEG and ITU-T VCEG Document JVT-G012r1, March 2003.
- [3] M. Jiang, N. Ling, Low-delay rate control for real-time H.264/AVC video coding, *IEEE Transactions on Circuits and Systems for Video Technology* 3 (2006) 467–477.
- [4] Y. Liu, Z.G. Li, Y.C. Soh, A novel rate control scheme for low delay video communication of H.264/AVC standard, *IEEE Transactions on Circuits and Systems for Video Technology* 1 (2007) 68–78.
- [5] S. Milani, L. Celetto, G.A. Mian, An accurate low-complexity rate control algorithm based on ( $\rho$ ; Eq)-domain, *IEEE Transactions on Circuits and Systems for Video Technology* 2 (2008) 257–262.
- [6] H. Wang, S. Kwong, Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination, *IEEE Transactions on Circuits and Systems for Video Technology* 1 (2008) 140–144.
- [7] X. Jing, L.-P. Chau, Improved frame level MAD prediction and bit allocation scheme for H.264/AVC rate control, in: *Proceedings of IEEE International Symposium on Circuits and Systems* (2007) 3639–3642.
- [8] Z. Chen, K.N. Ngan, Towards rate-distortion tradeoff in real-time color video coding, *IEEE Transactions on Circuits and Systems for Video Technology* 2 (2007) 158–167.
- [9] S. Ma, W. Gao, Y. Lu, Rate-distortion analysis for H.264/AVC video coding and its application to rate control, *IEEE Transactions on Circuits and Systems for Video Technology* 12 (2005) 1533–1544.
- [10] X.K. Yang, Y.M. Tan, N. Ling, Rate control for H.264 with two-step quantization parameter determination but single-pass encoding, in: *EURASIP Journal on Applied Signal Processing, Special Issue on Advanced Video Technologies and Applications for H.264/AVC and Beyond* (2006) 35–47.
- [11] D.-K. Kwon, M.-Y. Shen, C.-C. Jay Kuo, Rate control for H.264 video with enhanced rate and distortion models, *IEEE Transactions on Circuits and Systems for Video Technology* 5 (2007) 517–529.
- [12] Y.-K. Tu, J.-F. Yang, M.-T. Sun, Rate-distortion modeling for efficient H.264/AVC encoding, *IEEE Transactions on Circuits and Systems for Video Technology* 5 (2007) 530–543.
- [13] P.H. Westerink, R. Rajagopalan, C.A. Gonzales, Two-pass MPEG-2 variable-bit-rate encoding, *IBM Journal of Research and Development* 4 (1999) 471–488.
- [14] Y. Yu, J. Zhou, Y. Wang, C.W. Chen, A novel two-pass VBR coding algorithm for fixed-size storage application, *IEEE Transactions on Circuits and Systems for Video Technology* 3 (2001) 345–356.
- [15] W.-N. Lie, C.-F. Chen, T.C.-I. Lin, Two-pass rate-distortion optimized rate control technique for H.264/AVC video, in: *Proceedings of the SPIE Visual Communication and Image Processing* (2005) 1061–1070.
- [16] D.-K. Kwon, M.-Y. Shen, C.-C. Jay Kuo, R-D optimized frame-layer bit allocation for H.264, in: *Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing* 507–510.
- [17] C. Que, G. Chen, J. Liu, An efficient two-pass VBR encoding algorithm for H.264, in: *Proceedings of the 2006 International Conference on Communications, Circuits and Systems* 118–122.
- [18] J. Huang, J. Sun, W. Gao, A novel two-pass VBR coding algorithm for the H.264/AVC video coder based on a new analytical R-D model, in: *Proceedings of the 2007 Picture Coding Symposium*.
- [19] D. Zhang, Z. Chen, K.N. Ngan, Two-pass rate control for constant quality H.264/AVC high definition video coding, picture coding symposium, Lisbon, Portugal, November 2007, paper 1106.
- [20] D. Zhang, Z. Chen, K.N. Ngan, Constant distortion rate control for H.264/AVC high definition videos with scene change, *International Symposium on Circuits and Systems* (2008) Seattle, May.
- [21] A. Ortega, Variable bit-rate video coding, in: M.-T. Sun, A.R. Reibman (Eds.), *Compressed Video over Networks*, Marcel Dekker, New York, NY, 2000, pp. 343–382.
- [22] Joint Model 11.0 (JM11.0), [Online]. Available: <<http://bs.hhi.de/suehring/tml/download/jm11.0.zip>>.