Signal Processing: *Image Communication* 28 (2013) 20–33

Contents lists available at SciVerse ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Rate control for consistent visual quality of H.264/AVC encoding ☆

Long Xu [a,b,*], Sam Kwong [a], Hanli Wang [c,d], Debin Zhao [e], Wen Gao [f]

[a] *Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong*
[b] *School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 10083, China*
[c] *Department of Computer Science and Technology, Tongji University, Shanghai 201804, China*
[d] *Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200092, China*
[e] *Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China*
[f] *School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China*

ARTICLE INFO

Article history:
Received 7 October 2011
Accepted 7 October 2012
Available online 29 October 2012

Keywords:
Video coding
Sliding window
Rate control
Video streaming
Consistent/smooth visual quality

ABSTRACT

In this paper, a novel rate control scheme with sliding window basic unit is proposed to achieve consistent or smooth visual quality for H.264/AVC based video streaming. A sliding window consists of a group of successive frames and moves forward by one frame each time. To make the sliding window scheme possible for real-time video streaming, the initial encoder delay inherently in a video streaming system is utilized to generate all the bits of a window in advance, so that these bits for transmission are ready before their due time. The use of initial encoder delay does not introduce any additional delay in video streaming but benefits visual quality as compared to traditional one-pass rate control algorithms of H.264/AVC. Then, a Sliding Window Buffer Checking (SWBC) algorithm is proposed for buffer control at sliding window level and it accords with traditional buffer measurement of H.264/AVC. Extensive experimental results exhibit that higher coding performance, consistent visual quality and compliant buffer constraint can be achieved by the proposed algorithm.

1. Introduction

In recent years, we have witnessed an exponential increase of real-time multimedia applications. The real-time requirements of these applications such as video-phone, videoconference, Video On Demand (VOD), and TV broadcasting, pose a big challenge for their integration into IP networks. Due to limited network resources and buffer requirements, significant variations of picture quality may be observed when high motion object or scene change occurs. Thus, rate control should not only control bit rate, but also achieve perceptually consistent visual quality. Most of the popular rate control schemes of existing video coding standards are model-based, such as TM5 for MPEG-2 [1], TMN8 [2–4] for H.263 and VM8 for MPEG-4 [5]. For H.264/AVC, the rate control scheme [7] was adopted as the reference algorithm. The key of these model-based rate control schemes is to find the relation between rate and distortion, namely Rate-Distortion (*R-D*) model. Traditionally, both rate and distortion models are formulated as a function of Quantization Parameter (QP), so that traditional *R-D* models can be expressed as functions of QP, such as the conventional *R-D* model quadratic in D inverse [6]. Another novel *R-D* model is the ρ-domain *R-D* model [10], which directly models the relation between bit rate and the percentage of zero coefficients after quantization.

Traditional rate control schemes are concentrated on the accuracy of bit rate controlling between the target bit

☆ The work was partially supported by a grant for the Research Grants Council of the Hong Kong (Project CITYU115109), the Major State Basic Research Development Program of China (973 Program) under Grant 2009CB320905, the Shanghai Pujiang Program under Grant 11PJ1409400, the National Natural Science Foundation of China under Grants 61202242, 61102059, the Fundamental Research Funds for the Central Universities under Grant 0800219158.

* Corresponding author at: Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong. Tel.: +852 34425253.
E-mail addresses: lxu@jdl.ac.cn, xulong@ee.cuhk.edu.hk (L. Xu).

0923-5965/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.image.2012.10.001

rate and the final coded bit rate. In [7–9], the *R-D* model quadratic in D inverse and the linear prediction of Mean Absolute Distortion (MAD) are presented for H.264/AVC. Meanwhile, Hypothetical Reference Decoder (HRD) is being considered in the process of rate control. Thus, the match of bit rate controlling and compliant buffer constraint can be both satisfied. However, there is usually a significant picture quality fluctuation by such methods, the reason of which lies in the unavailability of the information of future frames for bit allocation, actual MAD values for calculating QPs before encoding the corresponding basic units, as well as of the existing buffer control strategy.

Conventionally, in addition to bit rate match and buffer constraint, rate control should also produce consistent visual quality [11–16] under given bit rate and buffer constraints. In [13], a new H.263+ rate control algorithm was proposed to support the Variable Bit Rate (VBR) encoding through frame rate adjustment. It changes the frame rate adaptively based on the motion information in a sliding window to reduce the image quality variation between adjacent frames. In [14], the sliding window concerned how many frames previously coded were used to update the rate-distortion model and model parameters. If the video content complexity changes significantly, a smaller window with more recent data would be used. In VBR rate control, the temporally consistent visual quality can be achieved by adjusting the output bits to the varying characteristics of video content. One of the typical VBR applications is the Digital Versatile Disk (DVD) [17,18]. Such an application relies on a VBR encoding system to achieve maximally consistent visual quality for the entire sequence, under the constraint of total storage capacity. It is a challenge because the video content may vary significantly from one frame to another. In practice, the researchers tried to solve this problem by two-pass or multi-pass encoding processes [17–20], where the first pass or several preceding passes are used to track the variation of characteristics of video content and thus the available bit resources can be distributed appropriately to various video segments, e.g., more bits are allocated to complex scenes or pictures. For the storage of video data without real-time requirement and buffer constraint, such as DVD, the coding scheme with multi-pass encoding is preferred for optimal coding efficiency and consistent visual quality. However, multi-pass encoding is not suitable for real-time encoding applications, so some one-pass VBR rate control algorithms are proposed [21–24]. On the other hand, in the client-server architecture of video streaming applications over networks, such as VOD and TV broadcasting, the videos may have been compressed already and are provided to users on demand, so there is no real-time requirement on encoding. But the buffer/time delay constraints should be considered due to the limited channel bandwidth, buffer capacity of network devices as well as the maximum tolerated time-delay requirement of terminal devices or end-users. The video streaming poses a big challenge to rate controls due to unstable networks and unexpected error of transmissions. And, rate controls for video streaming applications [25–28] have been becoming more and more popular in both research and industry fields.

In this paper, we first propose a new rate control scheme with sliding window basic unit for video streaming. A sliding window refers to a segment of video sequence which includes several successive frames or several Group-of-Pictures (GOPs). In addition, it moves forward by one frame each time along a video sequence. The initial encoder delay T_e inherently in video streaming systems is utilized to generate the encoding bits of a window for transmission purpose in advance, which makes the proposed rate control scheme possible for real-time video streaming applications. Then, a Sliding Window Buffer Checking (SWBC) method is proposed for elegant buffer control at sliding window level instead of frame level by traditional methods. It should be noted that the usage of sliding window in this work is different from those in [13,14]. The proposed algorithm in this work is also different from [21] which proposed a theoretical model handling the tradeoff between buffer constraint and picture quality fluctuation, but lacked a practical buffer control mechanism.

The rest of this paper is organized as follows. Section 2 gives a brief introduction of HRD and the Leaky Bucket (LK-B) model in H.264/AVC. In Section 3, the proposed SWBC and sliding window rate control algorithm are presented in detail. The simulation results of the proposed algorithm are given in Section 4, and the last section concludes the paper.

2. HRD in H.264/AVC

For video streaming applications, the buffer constraint is required, which is realized by HRD for H.264/AVC buffer management [29–31]. The goal of HRD is to ensure that the coded bitstream neither overflow nor underflow the decoder buffer under a given channel rate. As the heart of HRD, the LK-B [29] model is represented by a triple parameter (R,F,B), where F denotes the initial buffer fullness with F_d representing for decoder and F_e for encoder, B is the bucket capacity, and R is the leaking bit rate of bucket. The LK-B is actually a direct metaphor for the decoder's input buffer or encoder's output buffer, i.e., the queue between the decoder/encoder and the communication channel. At the encoder, we assume that the coded bits of the ith frame are poured into LK-B at the time s_i. After that, the coded bits are transmitted through communication channel with the given bit rate and enter into the decoder buffer. And then, at time $t_i=s_i+\Delta$ (Δ is the channel delay from server to client), the decoder removes the bits of the ith frame from the decoder buffer and decompresses it.

According to the LK-B model, the time scheduling including initial arrival time, removal time and decoding time of each frame can be deduced from the triple parameters (B,F,R) or (D,T_d,R). The relation between T_e/F_e, T_d/F_d and D/B is illustrated in Fig. 1, where the vertical axis shows the accumulated bits of encoder/decoder buffer and the horizontal axis represents time. The staircases in Fig. 1 are due to the behavior of bits arrival and removal of

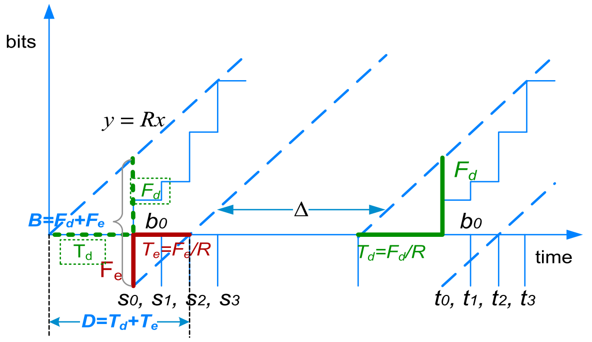

Fig. 1. The relation between F_e/T_e, F_d/T_d and B/D (CBR case: decoding schedule is the simple shift of encoding schedule by a constant delay Δ. $s_i = T_e + i \times (1/F_r)$ and $t_i = T_d + i \times (1/F_r)$ are the time-stamp of encoding and decoding, respectively).

a frame. The dotted lines denoted by "$y = Rx$" indicate the constant bit rate of communication channel between encoder and decoder. In Fig. 1, B is the distance between two dashed lines in which the slope equals to R. Given F_d and F_e (or T_d and T_e), then $B = F_d + F_e$ (or $D = T_d + T_e$) as shown in Fig. 1. Usually, the channel delay Δ is assumed to be constant and is omitted, so the E2E delay of a video streaming system is equal to D. In the Joint Model (JM) reference software of H.264/AVC [33], T_e is usually configured to be $0.8 \times D$. This indicates that the encoder buffer keeps the encoded bits until its fullness reaches at F_e, i.e., delay $T_e = F_e/R$ seconds.

There exist several LK-B models in the existing video coding standards, such as CAT-LB (Constraint Arrival Time-Leaky Bucket), EAT-LB (Earliest Arrival Time-Leaky Bucket) and LAT-LB (Latest Arrival Time-Leaky Bucket) [30–32]. In H.264/AVC, HRD employs a CAT-LB model for timing of bits arrival and removal associated with a frame from decoder buffer. CAT imposes a causality constraint on the arrival time of encoded bits at the decoder buffer. According to this constraint, the initial arrival time of each frame cannot be earlier than the difference of encoder processing time between that frame and the first frame, i.e., the initial arrival time of the ith frame $T_{ai}(i)$ is the latter of $T_{af}(i-1)$ and $T_{ei}(i)$ as

$$T_{ai}(i) = \max\{T_{af}(i-1), T_{ei}(i)\}$$
$$T_{ei}(i) = T_d + i \times (1/F_r); \quad T_{af}(-1) = 0; \quad i = 0,1,2,\ldots \quad (1)$$

where $T_{af}(i)$ is the time when the ith frame enters into decoder buffer completely, F_r represents frame rate and $1/F_r$ is the processing time of one frame which is the time interval between the removals of two successive frames from the decoder buffer. $1/F_r$ is also corresponding to the encoder processing time interval between two successive frames. $T_{ei}(i)$ is the earliest arrival time of the ith frame, which controls the earliest arrival time of the ith frame into decoder buffer due to CAT constraint.

3. Sliding window rate control with SWBC

Employing one-pass rate control [7–9], the temporally consistent visual quality cannot be ensured, especially for scene changes or high motion objects across frames. As mentioned above, two-pass and multi-pass rate control algorithms can provide consistent visual quality. However, they do not have any buffer control mechanism, so the encoded bitstream is not fit for video streaming which is the most popular usage of video coding nowadays. From Section 2, there is a certain capability for processing a group of frames together as the two-pass rate control in video streaming systems given the encoding delay T_e and computer power. To this end, a sequence is divided into windows which have the time length of T_e. Each window is encoded with the buffer constraint. Such a method is indicated as Window Rate Control (WRC) and the original one as Sequence-level Rate Control (SRC).

3.1. Feasibility of WRC in real-time coding system

The feasibility of performing WRC in real-time is validated by utilizing the initial encoder delay (i.e., T_e) as follows. H.264/AVC takes CAT-LB model to be the prototype of HRD. From (1), the earliest arrival time of the ith frame is T_{ei}. The time interval between each two successive frames entering into decoder buffer is $1/F_r$. According to symmetric theory of HRD [29] shown in Fig. 1, the time of the ith frame leaving encoder buffer for transmission is at $T_e + i(1/F_r)$. With T_e, the bits of a frame were delayed in the encoder buffer at least T_e seconds after they are generated. The time interval of two successive frames entering into channel is $1/F_r$, which is defined the processing time of a frame in HRD theory. Thus, we can accomplish the encoding of a window containing $T_e \times F_r$ frames theoretically before the due time of this window into channel by utilizing T_e inherently in a video streaming system. In the sequel, we are inspired to process up to $T_e \times F_r$ frames together as a basic unit in WRC, which does not introduce any additional delay but could benefits much smooth visual quality by using two-pass rate control algorithm at window level. WRC is competitive in video streaming applications. Note that T_e provides the possibility of WRC in real-time video encodings, such as Live TV programs. The computer power needs to be further considered in realization.

3.2. SWBC

In the traditional algorithm [7–9], the buffer status is checked for each frame for the HRD conformance. If the buffer violation occurs, the bits quota of current frame will be modified timely to prevent buffer overflow and underflow. Such a buffer strategy may result in the deterioration of picture quality when the frame contains complex content or high motion objects. In our work, the bits information of a window of frames is available before encoding, so all frames in a window are contributed to buffer control in the proposed SWBC if the buffer violation occurs. At the same time, the window slides forward one frame each time the buffer status of a frame is examined, so that SWBC conforms to HRD. This process is performed after the first-pass encoding and followed by the second-pass encoding.

The encoder buffer fullness is computed as

$$B_e(j) = \begin{cases} \sum_{i=j-S}^{j-1} R_1(i) + R_1(j), & j < T_e \times F_r; \\ \sum_{i=j-S}^{j-1} \max(0, R_1(i) - R_o(i)) + R_1(j), & j \geq T_e \times F_r; \end{cases} \quad (2)$$

where T_e is the initial delay of encoder, F_r is the frame rate, T_e represents the encoder initial delay and $S = F_r \times Te$. $R_1(i)$ represents the number of bits of the ith frame of the first-pass encoding. R_o is the throughput bits of encoder which is formulated as

$$R_o(j) = \begin{cases} 0 & j < T_e \times F_r; \\ \min\{R_c/F_r, B_e(j-1)\}, & j \geq T_e \times F_r; \end{cases} \quad (3)$$

where R_c is the average channel bit rate. It should be pointed out that only the decoder buffer underflow which corresponds to the encoder buffer overflow should be handled in VBR encoding. The decoder buffer overflow indicates that the channel is empty, which is allowable in VBR encoding. The decoder buffer occupancy is represented by

$$B_d(j) = \begin{cases} 0 & j < T_e \times F_r; \\ \sum_{i=j-S}^{j-1} R_o(i) + R_o(j) & T_e \times F_r \leq j < (T_e + T_d) \times F_r; \\ \sum_{i=j-S}^{j-1} R_o(i) + R_o(j) - \sum_{i=j-S-S'}^{j-1} R_1(i) & j \geq (T_e + T_d) \times F_r; \end{cases} \quad (4)$$

Here T_d denotes the decoder initial delay which is the minimum waiting time before decoder starts to work. S' represents the number of frames corresponding to the decoder initial delay T_d.

To avoid buffer underflow, the decoder buffer occupancy $B_d(j)$ should not be empty anytime, i.e., $B_d(j) > 0$. According to the symmetric characteristics of LK-B in Section 2, to avoid the decoder buffer underflow, the encoder buffer overflow should be forbidden, i.e., $B_e(j) < B$. After the first-pass encoding, the buffer status of each frame in a window can be computed by (2). When $B_e(j) < B$ does not hold for the jth frame, the encoding bits of the frames before the jth frame would be all modified in order to get as smooth as possible QPs. In this stage, the predictive bits instead of actual encoding bits are used to perform buffer checking. We mark it as $R_{expect}(j)$ for the jth frame which is obtained from the iteration of (8). Accordingly, the buffer checking formula is designed as

$$\sum_{i=j-S}^{j-1} \max(0, R_{expect}(i) - R_o(j)) + R_{expect}(j) < B, \quad (5)$$

where "max" is due to the CAT constraint, "expect" means the predicted bits from the statistics of the first pass encoding. We also update the buffer status with the actual coding bits during the second-pass encoding. Hence, there are actually two stages of buffer checking in the proposed SWBC method. By this means, the QPs of a group of frames, i.e., a window, are as smooth as possible. Meanwhile, the output bitstream is subject to the buffer constraint. In the second buffer checking stage, the buffer is updated with the actual coding bits of the preceding frames relative to the current frame as

$$\sum_{i=j-S}^{j-1} \max(0, R_{actual}(i) - R_o(j)) + R_{expect}(j) < B \quad (6)$$

Usually, the exception of (5) hardly occurs, so the revision is only at the current frame for simplicity if (6) fails.

3.3. Proposed sliding window rate control algorithm with SWBC

In SRC, the first-pass encoding yields the statistics of the entire video sequence such as bits usage profile, scene change, and quantization parameters. The second-pass encoding reallocates target bits for each frame to get consistent visual quality over the entire sequence according to the statistics above. In [19], a set of subjective experiments are performed on multiple video scenes, where the video scenes are encoded sequentially in CBR and VBR formats, respectively. The experimental results show that the bits usage profile of VBR is highly correlated with the distortion (or QP) profile of CBR, so an *R-Q* model for the second-pass VBR encoding is proposed as

$$R_{2,n} = kR_{1,n}(Q_{1,n})^p \quad (7)$$

where $R_{2,n}$ is the reallocated bits of frame n for the second-pass encoding, $R_{1,n}$ and $Q_{1,n}$ are the bits and quantization step obtained from the first-pass encoding, k is the model parameter and p is the scene-dependent factor. Based on our experiences, p equals 0.44, 0.45 and 0.5 for *I*, *P* and *B* frames, respectively. The model in [19] is based on the assumption that the coding complexity of the first-pass and second-pass are almost the same, i.e., $\sum_{n=0}^{N-1} R_{1,n}Q_{1,n} = \sum_{n=0}^{N-1} R_{2,n}Q_{2,n}$. In [34], Zhang et al. assumed that the frame bits and distortion conform to $\sum_{n=0}^{N-1} R_{1,n}D_{1,n} = \sum_{n=0}^{N-1} R_{2,n}D_{2,n}$, where $D_{i,n}$ represents the frame distortion in terms of Mean Squared Error (MSE) for the corresponding ith frame at nth-pass encoding. In [35], the relations between PSNR and bit rate, PSNR and QP are proposed according to the statistics of the first-pass encoding. This algorithm can provide very smooth PSNR on Common Interchange Format (CIF) sequences, but with degraded bit rate controlling accuracy and coding efficiency. In both [34,35], all the frames are quantized by a constant QP in the first-pass encoding, which assumes that their statistical *R-D* models can be well constructed with the bit rate of the first-pass encoding, but the accuracy of bit rate controlling is somewhat low when the target bit rate of the second-pass encoding is far from that of the first-pass encoding. In addition, these two algorithms are proposed specifically for VBR encoding.

In our proposal, the first-pass scheme employs the traditional CBR rate control of H.264/AVC [7–9]. The QPs of different frames change in a certain range when traditional CBR rate control is employed, so our *R-D* model for the second-pass encoding works well at a larger range of bit rates. In addition, the proposed algorithm can be applied to both CBR and VBR cases. For VBR encoding, the target bits of each window are allocated according to its complexity weight in the overall complexity of the whole sequence as described in [34]. For video streaming applications, the proposed rate control scheme can be regarded as the CBR method at window level and VBR method at frame level within each window. The total bits are distributed equally among windows due to the fixed window size given by $T_e \times F_r$ in our algorithm, so it is the window level CBR. While within a window, the bits budget is shared by all frames without quota limit on individual frames as the manner of VBR.

For the second-pass encoding, a new set of QPs are computed which is similar to the "bits reallocation" process as described in [19]. In the proposed algorithm, Eq. (7) is used to implement "bits reallocation". From (7), a new set of $\{R_{2,n}\}$ can be obtained when $Q_{1,n}$, k and p are provided. However, k is an unknown parameter which is calculated from the target bits of the second-pass encoding and $\{Q_{1,n}\}$, $\{R_{1,n}\}$ of the first-pass encoding [19]. Usually, an iterative process as shown in Table 1 is proposed to select the best k, and $R_{2,n}/Q_{2,n}$ is calculated according to (8). To be specific, the iterative process is performed on I, P and B frames separately as

$$\begin{cases} Q_{2,n}^{j+1} = f\left(k_j, R_{2,n}^{j}\right) \\ R_{2,n}^{j+1} = k_j R_{1,n}\left(Q_{2,n}^{j+1}\right)^{p} \end{cases}, \tag{8}$$

where j denotes the index of iteration, f is the inverse function of (7). Initially, $R_{2,n}^{0}$ and $Q_{2,n}^{0}$ ($j=0$) are given by the statistics of the first-pass encoding. In each iteration, two new sets of $\{R_{2,n}^{j}\}$ and $\{Q_{2,n}^{j}\}$ are calculated according to (8). After the whole process of iteration, a set of $\{Q_{2,n}^{j}\}$ can be obtained. In addition, the sum of $\{R_{2,n}^{j}\}$ corresponding to $\{Q_{2,n}^{j}\}$ matches the target bit rate of the second-pass encoding. We denote the final $\{R_{2,n}^{j}\}$ and $\{Q_{2,n}^{j}\}$ as $\{R_{2,n}\}$ and $\{Q_{2,n}\}$, respectively. Assuming that the boundaries of Binary Search as $[L, U]$, the total bits of the second-pass encoding as $R_{2,tot}$, and the total complexity of the first-pass encoding is computed by $C_{1,tot} = \sum_{n=0}^{n=S-1} R_{1,n} \times Q_{1,n}$, the Pseudo-code of the process is stated in Table 1.

Based on SWBC and Table 1, the proposed rate control algorithm is depicted as a flowchart shown in Fig. 2. Given the frame rate F_r, buffer delay D, the initial encoder delay is set to $T_e = 0.8 \times D$ and the window size is computed as $S = F_r \times T_e$. The whole encoding is divided into three parts in Fig. 3. In the first part, the first-pass CBR encoding is performed at the current window to gather the QP and bits usages $\{Q_{1,i}, R_{1,i}\}$ of a window of frames. In the second part, a new set of smooth QPs are produced by using the algorithm of Table 1. In addition, SWBC is performed to check buffer status and modify some of QPs in order to a compliant buffer of encoding. In the third part, the set of QPs from the second part are used to perform the second-pass encoding. However, the QPs would be additionally modified if the instant buffer status updated for each frame encoding. Thus, there are two stages for buffer checking in the proposed WRC. Generally, good buffer compliance can be obtained after the first stage of buffer checking. The second stage of buffer checking is only carried out for few of frames if buffer violation occurs during encoding.

3.4. Example of the sliding window rate control for video streaming system

Considering the application of the proposed rate control algorithm, a video streaming broadcasting system is taken as an example as shown in Fig. 3. The system is used for delivery of video stream over IP network by broadcasting, which is usually applied in the Local Area Network (LAN). As shown in Fig. 3, it consists of a video streaming server with an encoder, channel, and client with a decoder. Our proposed WRC and SWBC are employed in the encoder at the streaming sever. The output bitstreams meet the given buffer constraint and target bit rate. In video streaming system, E2E delay D from streaming server to client is the sum of T_e and T_d. We explain how to utilize them into our algorithm as follows.

As discussed in Section 2, the decoding behavior is determined by HRD whose mathematical model is CAT-LB in H.264/AVC. In the proposed SWBC, we do not change the syntax of HRD, but only utilize T_e of HRD to smooth visual quality. From Fig. 1, the ith frame is processed at the time interval from s_{i-1} to s_i ($s_i = T_e + i \times (1/F_r)$), and removal of the ith frame from encoder buffer is at the time of s_i. However, utilizing the startup delay T_e, the first window whose size equals $T_e \times F_r$ is processed together in a two-pass manner

Table 1
Proposed algorithm for searching a set of expected QPS.

initialize $\omega = R_{2,tot}/C_{1,tot}$, $L = 1\text{E}-7*\omega$, $U = 1\text{E}+4*\omega$, $k=0$, $j=0$;
for ($d=U$ to L, $d*=0.5$; j++)
  $k+=d$;
  Compute $Q_{2,n}^{j+1}$ according to the upper equation of (8) for each frame;
  Smooth $Q_{2,n}^{j+1}$ using Gaussian filter{$\omega(i)$}, i.e., $\omega=(i)\exp(-i_2/d^2)$ and $\sum_{i=-d}^{i=d}\omega_i = 1$
  Compute $R_{2,n}^{j+1}$ according to the lower equation of (8) for each frame;
  if $\sum R_{2,n}^{j+1} > R_{2,tot}$
    $k-=d$;
  end if
end for
$Q_{2,n} = Q_{2,n}^{j+1}$, which is satisfied with our requirement.

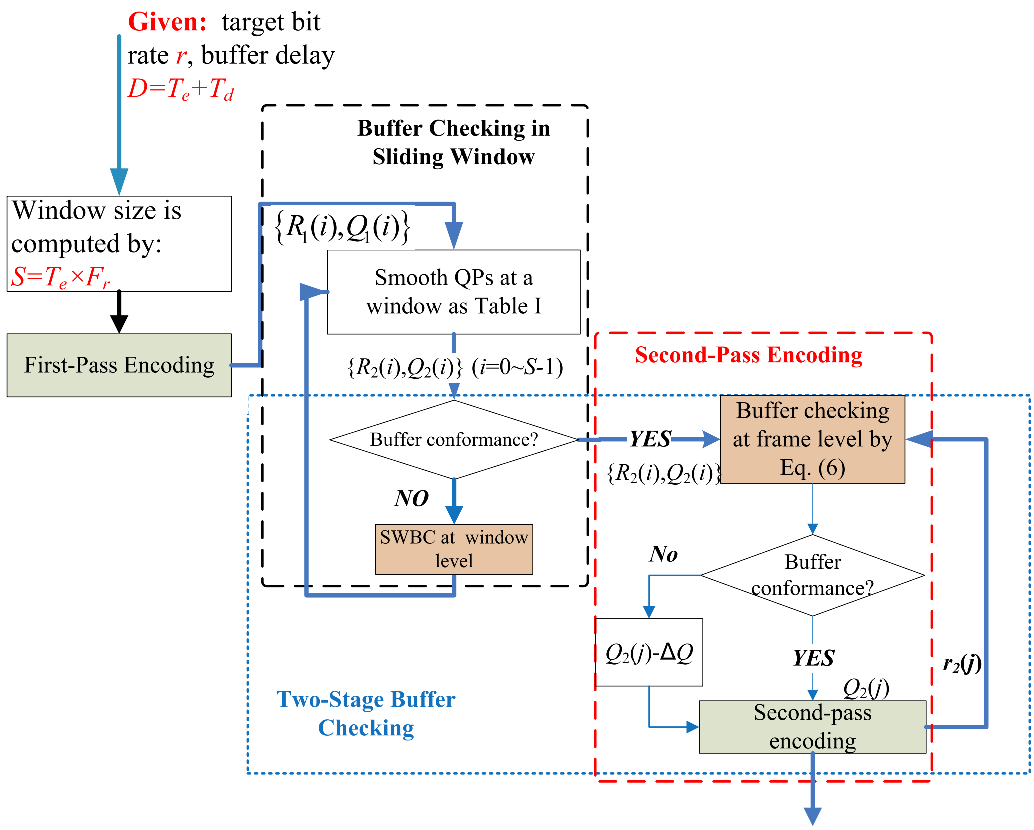

Fig. 2. The flowchart of the proposed rate control algorithm with SWBC.

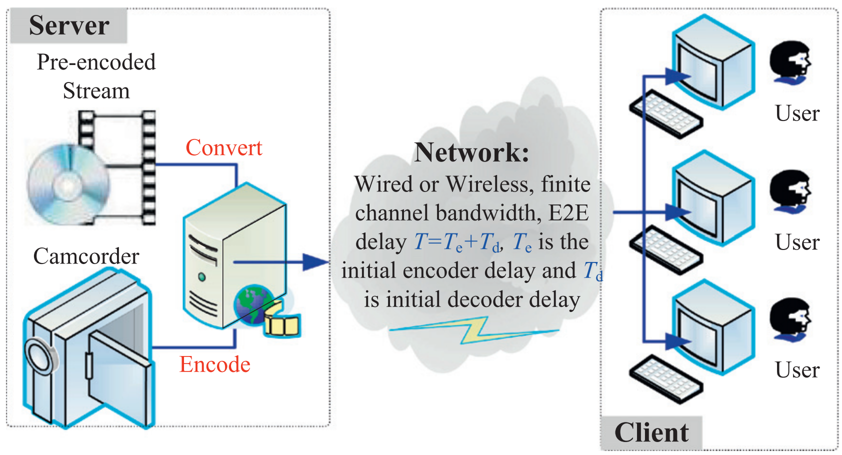

Fig. 3. Brief flowchart of a video streaming broadcasting system.

before time s_0, and then the encoded bits of all frames are ready at time of s_0. Thus, each frame of the first window can be ready at its due time. The same conclusion can be obtained for subsequent windows. The two-pass processing at window level is possible under the condition of enough computing power. In clients, the received bits are decoded successfully without buffer underflow and overflow by further introducing a startup delay of T_d to decoder.

4. Experimental results

We implement the proposed algorithm on JM14.0 of H.264/AVC [33] under the conditions: *Profile/Level: 100/40, IntraPeriod: 16, Reference frames*: *5, IPPP coding structure, Full search*, Search range: *16, FMO off, RDO on* and *CABAC*. The window size of WRC equals 16 by default unless specified otherwise. The experiments on the proposed

WRC and SWBC are arranged separately into two parts: the first part exhibits the efficiency of WRC from the aspects of visual quality smoothness and PSNR improvement, and the second part validates SWBC on buffer constraint encoding.

4.1. Coding efficiency of the proposed WRC

The experiments are performed with the traditional rate control algorithm [7–9] at frame level, SRC [19], the algorithm Zhang2009 [34], the algorithm Huang2007 [35] and the proposed WRC. The first one is a one-pass method, and the others are two-pass methods. SRC, Zhang2009 and Huang2007 are designed for VBR encoding, while the traditional algorithm [7–9] and WRC are used for CBR encoding. For a fair evaluation campaign among the CBR and VBR algorithms concerned, the buffer constraint is excluded from this part of experiments. Firstly, the comparisons between the traditional algorithm [7–9] and WRC are made with standard test sequences, including "Foreman", "Football", "Mobile" and "Silent" in CIF resolution, "Autumn" and "Crowds" in SD (720×576) resolution, "Night", "Crew" and "Harbour" in 720P (1280×720) HD resolution.

The frame QP/PSNR of WRC and the traditional algorithm are partially plotted in Figs. 4 and 5 for exhibiting the visual quality consistency. From both figures, WRC achieves much smoother QP/PSNR performances than the traditional one [7–9]. Moreover, the curves for window size of 64 denoted by "$L=64$" are smoother than those curves for window size of 16 denoted by "$L=16$" in Fig. 4, which indicates that the larger the window is, the smoother visual quality WRC can achieve. In Fig. 5, the curves are drawn for the two comparative algorithms with "Intraperiod" of 64 and window size of 64, where the difference between them can be seen much more clearly. Furthermore, WRC employs the traditional algorithm in its first-pass encoding for gathering statistics of frame bits and QP, so it would always benefit the performance improvement of traditional algorithm. The QP/PSNR variations represented by their Standard Deviations (STDs) are tabulated in Table 2 for all testing algorithms, where the less STDs of QP/PSNR can be observed for WRC as compared to the traditional one [9] in all tests. The other three algorithms are for VBR applications at sequence level, so they cannot be compared with WRC in a fair manner. The PSNR variations of Huang2007 are the smallest among all the test algorithms. However, it greatly degrades the RD performance. On the other hand, Zhang2009 is inferior on low-resolution videos. Considering both smooth quality and RD performance, SRC is better than the other two VBR algorithms.

For demonstrating the advantage of WRC in both *R-D* performance and bit control accuracy, the comparative results for all test algorithms are shown in Fig. 6 and summarized in Table 3. In Table 3, the upper number of each resolution in the column labeled "bit control error" is the absolute error of bit control, while the lower one is the average bit control error. The former exhibits the

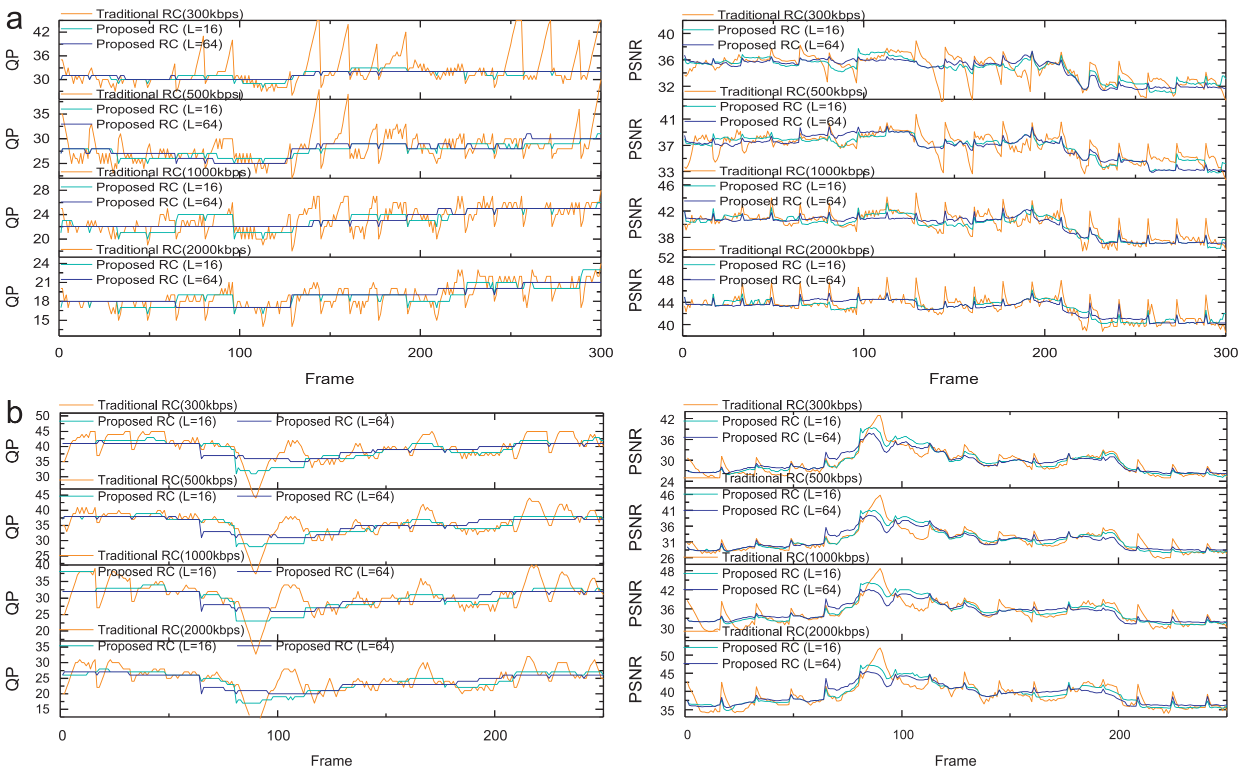

Fig. 4. Frame QP and PSNR curves of WRC and traditional algorithm. (a) "Foreman" and (b) "Football". (All curves are drawn with IntraPeriod of 16, "$L=16$" and "$L=64$" denote window size of 16 and 64, respectively.)

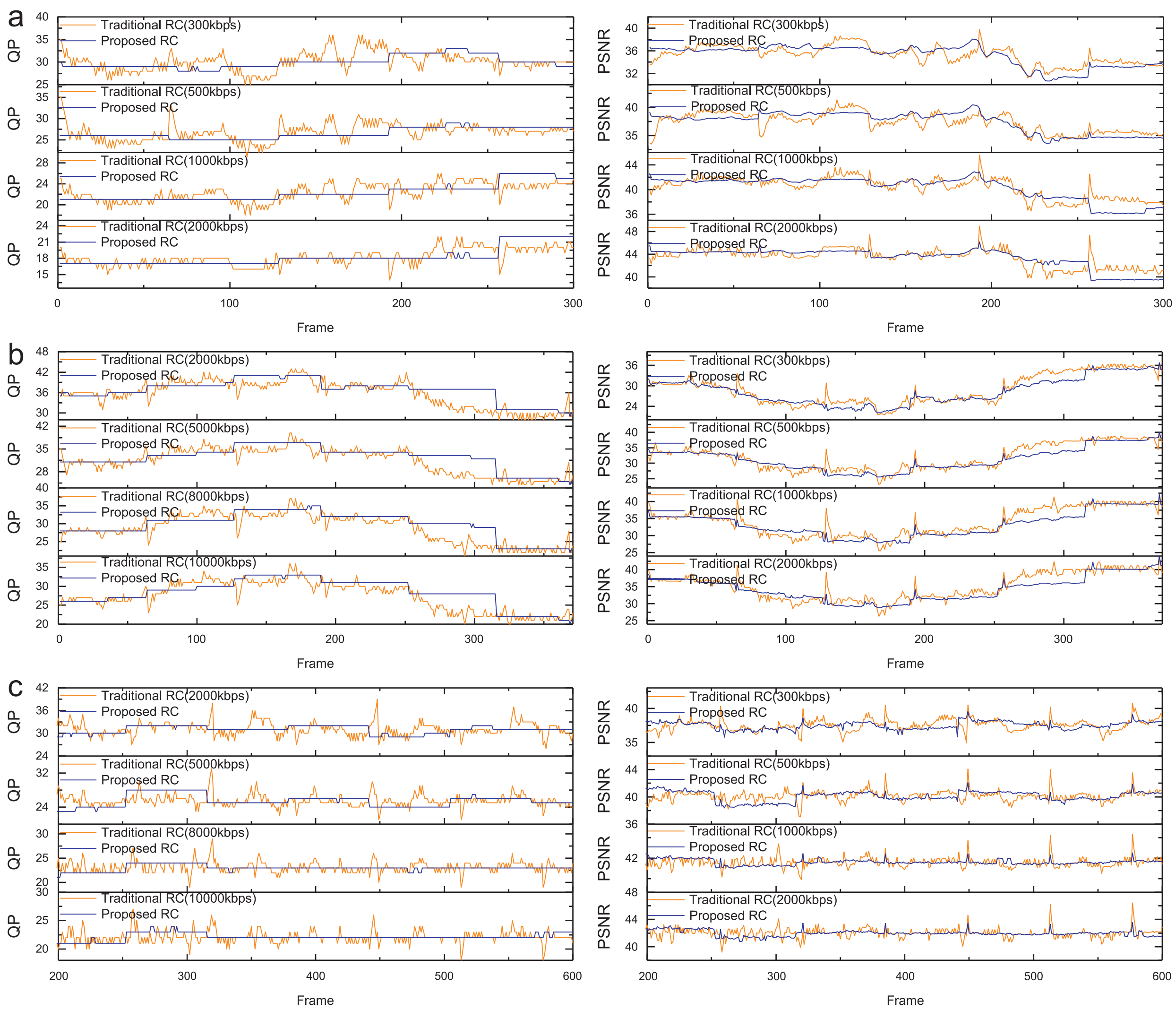

Fig. 5. Frame QP and PSNR curves of WRC and traditional algorithm (IntraPeriod = 64, window size = 64). (a) "Foreman", (b) "Autumn" and (c) "Crew".

Table 2
STD of QP/PSNR for all the test algorithms. (Each is the average STD of sequences with the resolution labeled by the first column, and only statistics data for intraperiod of 16.)

Resolution	Bit rate (kbps)	QP variation (STD)						PSNR variation (STD)					
		Traditional	WRC (L=16)	WRC (L=64)	Zhang2009	Huang2007	SRC	Traditional	WRC (L=16)	WRC (L=64)	Zhang2009	Huang2007	SRC
CIF	2000	2.22	1.63	1.32	1.84	1.27	0.24	1.97	1.78	1.69	1.41	0.35	0.75
	1000	2.50	1.77	1.47	2.65	1.44	0.24	1.95	1.86	1.76	1.67	0.34	0.82
	500	2.72	1.47	1.46	3.45	1.61	0.24	1.87	1.77	1.58	1.95	0.26	0.97
	300	3.71	1.47	1.35	3.81	1.69	0.24	1.87	1.74	1.53	2.04	0.26	1.02
SD	10,000	2.65	2.28	2.03	2.51	1.47	0.40	2.66	2.40	1.89	1.54	0.50	1.01
	8000	2.69	2.27	1.97	3.43	1.58	0.35	2.74	2.41	1.78	1.80	0.59	1.15
	5000	2.75	2.33	1.58	4.69	2.20	0.40	2.78	2.53	1.85	2.10	0.41	1.28
	2000	2.84	2.23	1.75	4.09	2.37	0.44	2.82	2.58	1.71	1.63	0.31	1.53
720P	10,000	1.66	0.92	0.53	0.86	0.62	0.35	1.15	0.81	0.54	0.75	0.18	0.48
	8000	1.50	0.89	0.78	0.98	0.70	0.24	1.06	0.77	0.71	0.79	0.26	0.39
	5000	1.32	0.81	0.76	1.18	0.74	0.65	0.86	0.67	0.50	0.77	0.27	0.68
	2000	1.57	1.22	0.48	2.02	0.60	0.84	0.85	0.78	0.49	0.96	0.21	0.76
Average		**2.34**	**1.61**	**1.29**	2.63	1.36	0.39	**1.88**	**1.68**	**1.34**	1.45	0.33	0.90

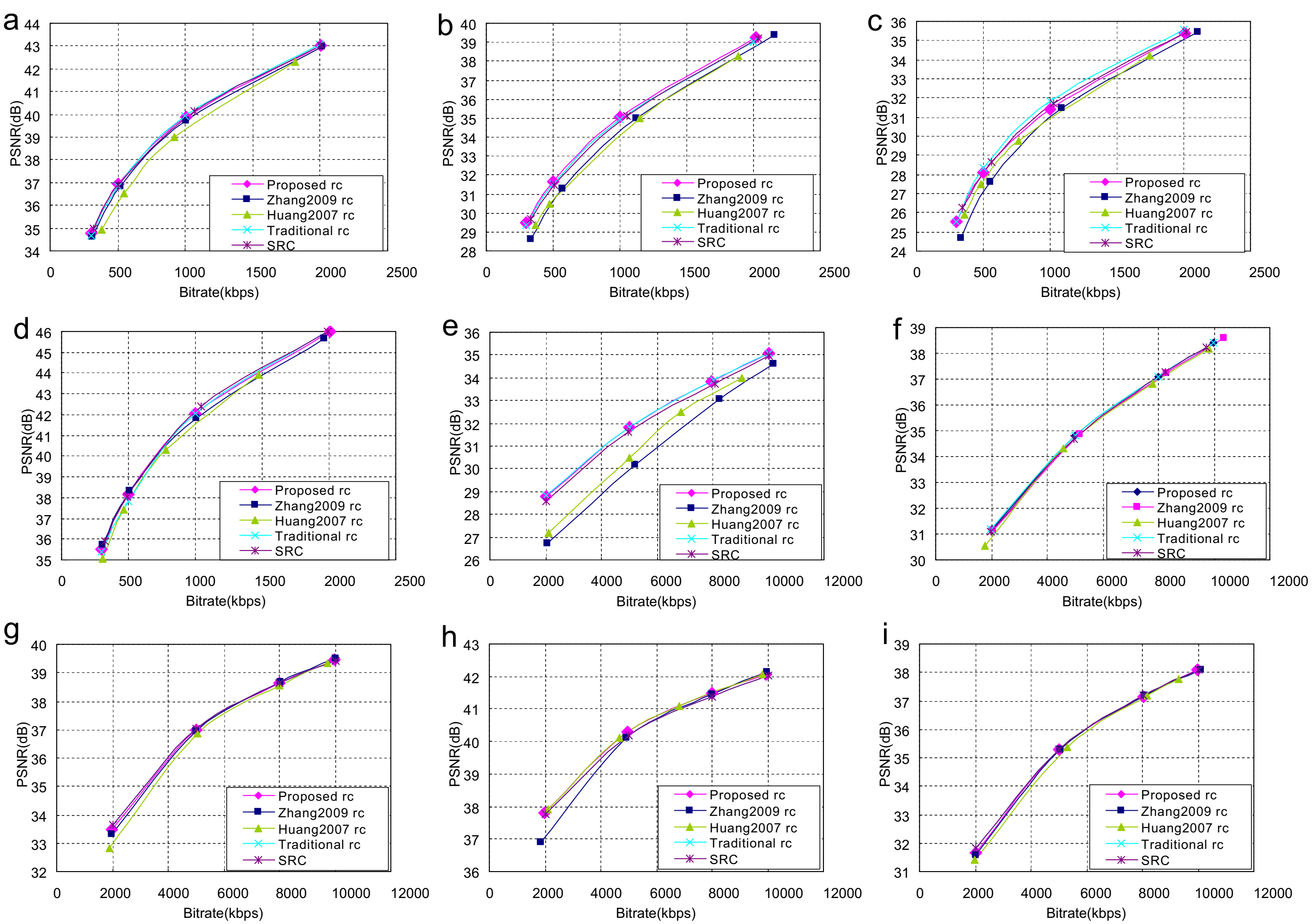

Fig. 6. Comparison of coding efficiency among all test algorithms. (a) “Foreman”, (b) “Football”, (c) “Mobile”, (d) “Silent”, (e) “Autumn”, (f) “Crowds”, (g) “Night”, (h) “Crew” and (i) “Harbour”.

Table 3
Comparison of coding efficiency of all the test algorithms.

Resolution	Sequence	Target bit rate (kbps)	Traditional			Zhang2009			Huang2007			SRC			Proposed WRC		
			Bit rate (kbps)	PSNR	Error (%)	Bit rate (kbps)	PSNR	Error (%)	Bit rate (kbps)	PSNR	Error (%)	Bit rate (kbps)	PSNR	Error (%)	Bit rate (kbps)	PSNR	Error (%)
CIF	Foreman	2000	2000.12	43.05	0.01	2015.69	42.96	0.78	1810.27	42.31	−9.49	2008.00	43.00	0.40	2003.63	43.00	0.18
		1000	1002.41	39.94	0.24	1007.67	39.72	0.77	918.79	38.99	−8.12	1069.18	40.14	6.92	1006.00	39.89	0.60
		500	502.96	36.82	0.59	517.31	36.85	3.46	549.55	36.54	9.91	503.68	36.91	0.74	501.98	36.92	0.40
		300	302.74	34.68	0.91	314.95	34.64	4.98	382.94	34.93	27.65	321.01	34.98	7.00	301.51	34.75	0.50
	Football	2000	2001.05	39.01	0.05	2159.76	39.41	7.99	1885.48	38.26	−5.73	2032.76	39.20	1.64	2014.11	39.26	0.71
		1000	1000.24	34.91	0.02	1125.70	34.96	12.57	1145.39	34.98	14.54	1048.60	35.11	4.86	1002.91	35.07	0.29
		500	500.17	31.52	0.03	568.30	31.29	13.66	477.05	30.45	−4.59	507.08	31.46	1.42	501.96	31.63	0.39
		300	300.16	29.38	0.05	334.72	28.60	11.57	369.20	29.37	23.07	330.48	29.67	10.16	300.20	29.50	0.07
	Mobile	2000	2001.90	35.60	0.10	2108.68	35.43	5.43	1749.97	34.21	−12.50	2018.00	35.46	0.90	2007.44	35.39	0.37
		1000	1001.65	31.81	0.16	1085.57	31.44	8.56	760.11	29.75	−23.99	1029.00	31.71	2.90	1003.29	31.38	0.33
		500	502.15	28.37	0.43	556.66	27.61	11.33	483.81	27.51	−3.24	557.10	28.67	11.42	501.15	28.12	0.23
		300	301.51	25.56	0.50	329.98	24.67	9.99	357.97	25.87	19.32	346.03	26.24	15.34	300.09	25.55	0.03
	Silent	2000	2002.72	46.07	0.14	1961.43	45.68	−1.93	1475.65	43.90	−26.22	1994.20	46.01	−0.29	2006.51	46.00	0.33
		1000	1002.11	42.07	0.21	1008.53	41.85	0.85	784.35	40.32	−21.57	1043.90	42.37	4.39	1003.16	42.07	0.32
		500	501.68	37.84	0.34	510.04	38.33	2.01	469.02	37.44	−6.20	488.04	38.07	−2.39	501.59	38.14	0.32
		300	301.27	35.37	0.42	309.82	35.74	3.27	303.80	35.07	1.27	325.93	35.91	8.64	300.74	35.51	0.25
Average					**0.26**			**6.20**			**13.59**			**4.96**			**0.33**
				35.75	**+0.26**		**35.57**	**+5.96**		**34.99**	**−1.62**		**35.93**	**+4.63**		**35.76**	**+0.33**
SD	Autumn	10,000	10,030.09	35.06	0.30	10,201.42	34.59	2.01	9041.00	34.00	−9.59	10,040.00	34.95	0.40	10,034.66	35.07	0.35
		8000	8032.12	33.88	0.40	8250.06	33.04	3.13	6846.55	32.47	−14.42	8100.00	33.75	1.25	7971.76	33.86	−0.35
		5000	5022.44	31.84	0.45	5208.46	30.17	4.17	5007.00	30.51	0.14	4980.00	31.60	−0.40	5018.97	31.84	0.38
		2000	2014.82	28.86	0.74	2075.67	26.73	3.78	2123.51	27.19	6.18	2020.00	28.60	1.00	2004.01	28.79	0.20
	Crowds	10,000	9999.00	38.39	−0.01	10,343.64	38.57	3.44	9802.45	38.20	−1.98	9728.57	38.25	−2.71	9995.70	38.40	−0.04
		8000	7999.08	37.10	−0.01	8299.90	37.25	3.75	7769.89	36.83	−2.88	8263.10	37.28	3.29	8009.49	37.11	0.12
		5000	4999.67	34.81	−0.01	5189.80	34.87	3.80	4586.81	34.32	−8.26	4970.00	34.67	−0.60	4997.14	34.81	−0.06
		2000	1999.82	31.19	−0.01	2067.18	31.15	3.36	1781.67	30.55	−10.92	1969.21	31.10	−1.54	2003.41	31.17	0.17
Average					**0.24**			**3.43**			**6.79**			**1.40**			**0.21**
				33.89	**+0.23**		**33.30**	**+0.34**		**33.01**	**−0.52**		**33.78**	**+0.09**		**33.88**	**+0.10**
720P	Night	10,000	10,017.36	39.44	0.17	9745.55	39.33	−2.54	10,030.94	39.50	0.31	10,017.36	39.44	0.17	9960.81	39.47	−0.39
		8000	8016.98	38.64	0.21	8008.56	38.57	0.11	8022.68	38.68	0.28	8016.98	38.64	0.21	7995.18	38.65	−0.06
		5000	5014.88	37.02	0.30	5056.02	36.85	1.12	4995.74	36.96	−0.09	5014.88	37.02	0.30	4992.98	36.97	−0.14
		2000	2007.90	33.63	0.40	1879.41	32.81	−6.03	1992.85	33.33	−0.36	2007.90	33.63	0.40	1999.28	33.49	−0.04
	Crew	10,000	10,018.28	42.01	0.18	9991.26	42.13	−0.09	9790.27	42.06	−2.10	10,018.28	42.01	0.18	9921.52	42.06	−0.78
		8000	8018.35	41.39	0.23	7996.91	41.46	−0.04	6838.43	41.07	−14.52	8018.35	41.39	0.23	7986.58	41.48	−0.17
		5000	5013.60	40.19	0.27	4932.05	40.11	−1.36	4673.69	40.10	−6.53	5013.60	40.19	0.27	4979.32	40.28	−0.41
		2000	2006.36	37.76	0.32	1852.54	36.91	−7.37	2114.68	37.90	5.73	2006.36	37.76	0.32	1995.61	37.82	−0.22
	Harbour	10,000	10,019.32	38.08	0.19	10,107.22	38.10	1.07	9315.00	37.77	−6.85	10,019.32	38.08	0.19	9997.00	38.11	−0.03
		8000	8027.41	37.18	0.34	8086.74	37.18	1.08	8154.35	37.18	1.93	8027.41	37.18	0.34	8022.52	37.16	0.28
		5000	5020.13	35.30	0.40	5036.64	35.29	0.73	5296.93	35.39	5.94	5020.13	35.29	0.40	5000.90	35.28	0.02
		2000	2012.36	31.81	0.62	2026.94	31.59	1.35	1978.86	31.43	−1.06	2012.36	31.80	0.62	2000.45	31.66	0.02
Average					**0.30**			**1.91**			**3.81**			**0.30**			**0.21**
				37.70	**+0.30**		**37.53**	**−1.00**		**37.61**	**−1.44**		**37.69**	**+0.30**		**37.71**	**−0.16**

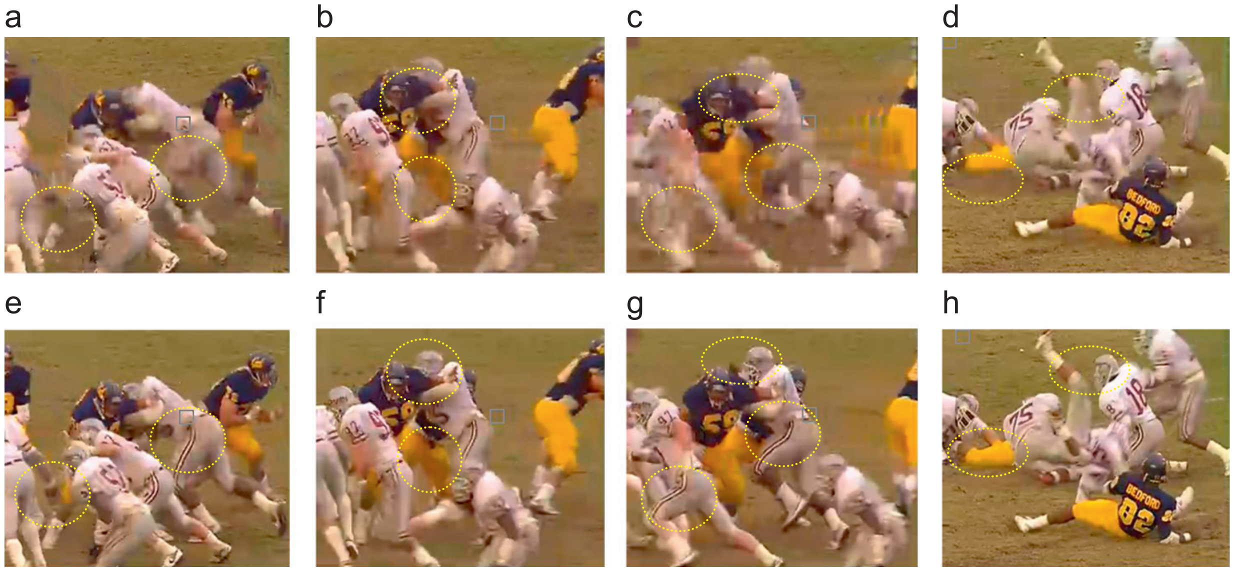

Fig. 7. Comparison of subjective visual quality between the traditional algorithm and WRC on "Football" at 300 kbps: (a)–(d) are decoded from the bitstreams of the traditional algorithm, (e)–h) are decoded from the bitstreams of WRC. (a) and (e) are 13th frame, (b) and (f) are 26th frame, (c) and (g) are 31st frame, (d) and (h) are 220th frame.

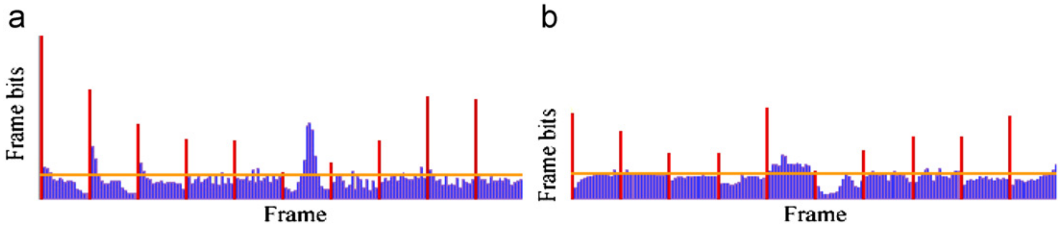

Fig. 8. Bits of each frame for the traditional algorithm and WRC on "Football" at 300 kbps. (a) Traditional algorithm and (b) Proposed WRC.

accuracy of bit rate control of a rate control algorithm. The latter is the average mismatch of bit control which is used to show *R-D* achievement by combining with average PSNR. From Fig. 6 and Table 3, the *R-D* performance of WRC is about the same as that of the traditional algorithm [7–9]. The highest PSNR degradations of Zhang2009 and Huang2007 are up to 2.13 dB and 1.65 dB (on "Autumn" at 2 Mbps), respectively, as compared with the traditional algorithm. Considering the bit control errors, the highest error is 13.66% and 27.65% for Zhang2009 and Huang2007, respectively. The average errors of those two algorithms are 6.2% and 13.59%, respectively, which are much worse than WRC. For HD sequences, the average bit control error of the proposed algorithm is 0.21%. And, the highest error of the proposed algorithm is 0.78% which is less than 7.37% and 14.52% of Zhang2009 and Huang2007, respectively.

To show the perceptual visual quality, a group of pictures decoding from the bitstream of "Football" are compared between the traditional algorithm and WRC in Fig. 7, where the blurring and blocking artifacts can be observed from the top-row pictures coded with the traditional algorithm, while the good visual quality of the bottom-row pictures are observed with WRC. From the results, WRC achieves consistent visual quality over the whole sequence in all test sequences. On the contrary, the traditional algorithm is prone to failure in cases of high motion object, fast scene switch and low bit rate. Usually, the traditional algorithm consumes too many bits on an I frame and its closely following frames, so the degradation of visual quality for subsequent frames of a GOP can be observed. The curves of frame bits of each frame for both of the two algorithms on "Football" are shown in Fig. 8. The smoother bits curve can be observed with WRC from Fig. 8(b), while for the traditional algorithm, most of bits are consumed for coding I frames and the following P frames as shown in Fig. 8(a).

For conventional CBR cases, the better performance of WRC than the benchmarks with respect to both *R-D* performance and visual quality smoothness can be observed. In addition, the proposed algorithm is also competitive in the situation where the available channel bandwidth is time varying. Such a channel can be simulated by a classical wireless transmission with packet loss [36]. The experiments show that the much more smooth visual quality and a certain *R-D* improvement also can be achieved by the proposed algorithm. As far as computational complexity, the first-pass encoding is only needed for the raw YUV data input in two-pass encoding, which is now unusual in actual applications. Most of video contents were already compressed for storage. Thus, there is no need of the first-pass encoding to collect bits and QP usages in the proposed algorithm.

4.2. Validation of SWBC on buffer constraint encoding

To demonstrate the efficiency of SWBC, the traditional rate control algorithm with the HRD constraint and the proposed one with SWBC are compared. The buffer delay is assumed to be 0.5 s for CIF sequences. The curves of buffer fullness for CIF sequences "Foreman" and "Football" are shown in Fig. 9, where violation of buffer fullness can be observed for both traditional and WRC without SWBC. By integrating SWBC into WRC, the compliant buffer fullness can be obtained. From the top curves of Fig. 9, the traditional algorithm may fail to control buffer although there is a buffer control method at frame-level The QP/PSNR curves shown in Fig. 10 indicate that the

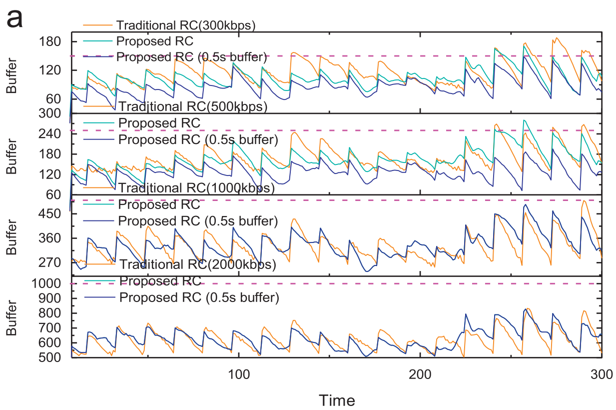

Fig. 9. Comparisons of encoder buffer fullness for WRC, WRC with SWBC and traditional algorithms. (a) "Foreman" and (b) "Football".

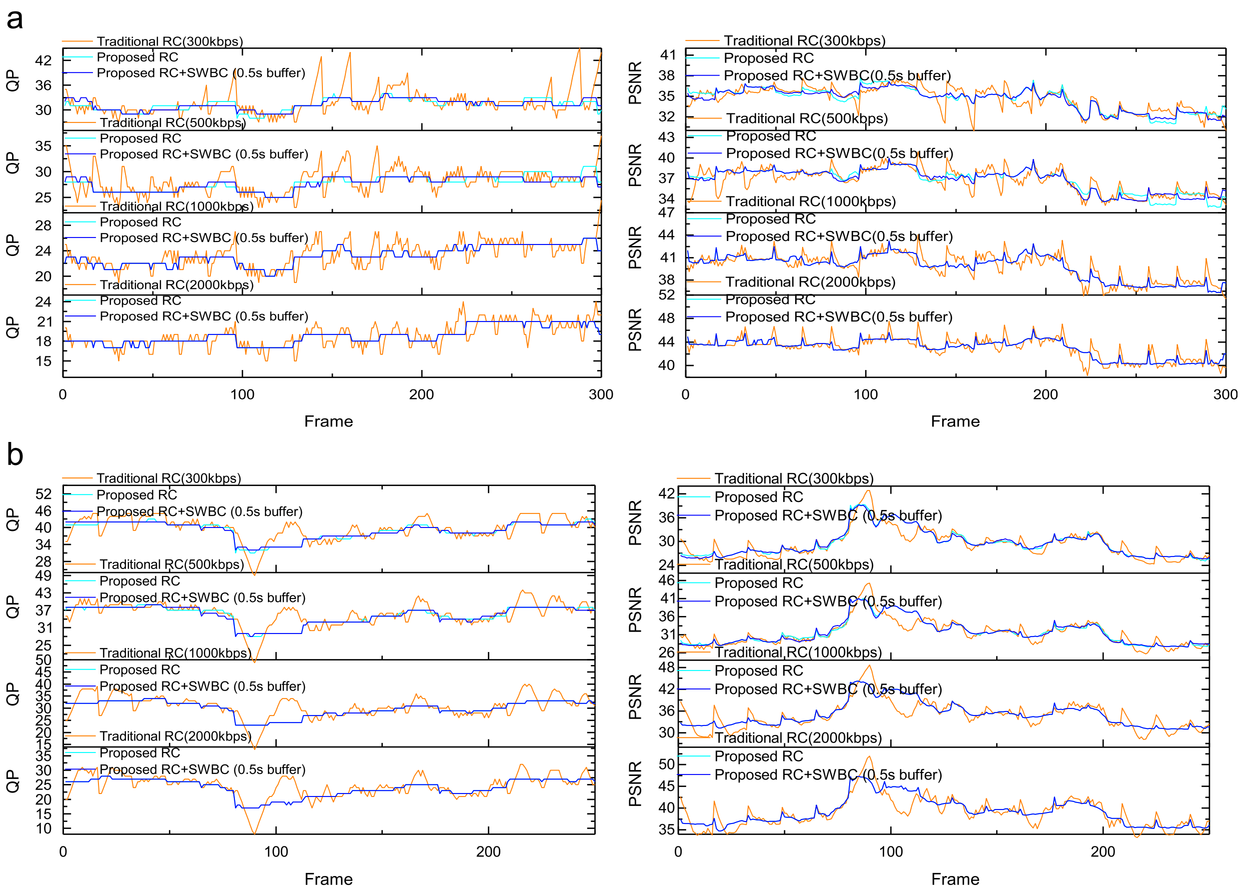

Fig. 10. Frame QP/PSNR curves for WRC, WRC with SWBC and traditional algorithms. (a) "Foreman" and (b) "Football".

Table 4
Buffer delay for constant QP encoding, traditional CBR encoding and our proposed SWBC.

Sequence	Fixed QP encoding	Traditional CBR encoding	The proposed WRC with SWBC
Die Another Day	17.74	1.96	8.32
Kill Bill	10.70	1.00	7.77
Sports	40.17	3.63	10.09
Lilo and Stitch	15.41	3.94	7.93
Entertainment	47.37	2.93	10.06

proposed algorithm can achieve significantly smoother picture quality than the others.

To verify SWBC on the proposed video streaming system in Subsection 3 (D), we further perform the experiments on three movie sequences "Die Another Day", "Kill Bill" and "Lilo and Stitch" (D1, i.e., 720×480), 2 video clips (QVGA, i.e., 320×240) recorded from a sports TV program and an entertainment TV program denoted as "Sports" and "Entertainment", respectively. The performance of the proposed algorithm can be validated from the perceptual experience of viewers, including (1) smooth visual quality, i.e., no obvious visual quality flicker; and (2) compliant buffer constraint, i.e., the bitstream can be decoded normally under the given buffer constraint without decoder crush and jitter. From the results, it is observed that the most consistent visual quality is obtained for the fixed QP encoding. However, the output bit rate varies very much, so there is a necessity of large buffer/long delay to transmit the coded bitstream over a constant bit rate channel. The delay is usually too large to most of the applications. The traditional CBR encoding is with the smallest initial delay due to the constant output bit rate. However, the serious fluctuation of visual quality may occur as the scene changes frequently. As a solution, the proposed algorithm is target at long-term CBR and VBR quality is also provided in short time interval. To demonstrate the performance of SWBC, the buffer delay of test movies is tabulated in Table 4, where the results of the fixed QP encoding are beyond the requirement of general video streaming applications (which is assumed to be 10 s in this work), and the resultant buffer delay by SWBC is less than 10 s.

5. Conclusions

In this paper, a novel rate control scheme at sliding window is proposed for video streaming applications. With the proposed algorithm, both smooth picture quality and buffer delay requirements can be satisfied. For smooth buffer controlling, a sliding window is utilized to regulate the output bits so that the coded bitstream is subject to the buffer requirement. The experimental results demonstrate that both high picture quality and compliant buffer constraint are achieved by the proposed algorithm. In the future, we will investigate the proposed algorithm for transcoding-based video streaming applications.

References

[1] ISO/IEC 13818-2, Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video, Second edition, December 2000.

[2] J.R. Corbera, S. Lei, Rate control in DCT video coding for low-delay communications, IEEE Transactions on Circuits and Systems for Video Technology 9 (1) (1999) 172–185.

[3] G. Côté, B. Erol, M. Gallant, F. Kossentini, H.263+: video coding at low bit rates, IEEE Transactions on Circuits and Systems for Video Technology 8 (7) (1998) 849–866.

[4] P. Yin, J. Boyce, A new rate control scheme for H.264 video coding, in: Proceedings of the IEEE ICIP 2004, Singapore, October 2004, pp. 449–452.

[5] H.J. Lee, T. Chiang, Y.Q. Zhang, Scalable rate control for MPEG-4 video, IEEE Transactions on Circuits and Systems for Video Technology 10 (6) (2000) 878–894.

[6] T. Chiang, Y.Q. Zhang, A new rate control scheme using quadratic rate distortion model, IEEE Transactions on Circuits and Systems for Video Technology 7 (1) (1997) 287–311.

[7] Z.G. Li, F. Pan, K.P. Lim, G. Feng, X. Lin, S. Rahardja, Adaptive basic unit layer rate control for JVT, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 7th Meeting, Doc. JVT-G012r1, Thailand, March 7–14, 2003.

[8] Y. Liu, Z.G. Li, A novel rate control scheme for low delay video communication of H.264/AVC standard, IEEE Transactions on Circuits and Systems for Video Technology 17 (1) (2007) 68–78.

[9] S.W. Ma, Z.G. Li, F. Wu, Proposed draft adaptive rate control, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 8th Meeting, Doc. JVT-H017r3, Geneva, May 20–26, 2003.

[10] Z. He, Y.K. Kim, S.K. Mitra, Low-delay rate control for DCT video coding via ρ-domain source modeling, IEEE Transactions on Circuits and Systems for Video Technology 11 (8) (2001) 928–940.

[11] K.–L. Huang, H.–M. Hang, Consistent picture quality control strategy for dependent video coding,, IEEE Transactions on Image Processing 18 (5) (2009) 1004–1014.

[12] K. Wang, J.W. Woods, MPEG motion picture coding with long-term constraint on distortion variation, IEEE Transactions on Circuits and Systems for Video Technology 18 (3) (2008) 294–304.

[13] H.J. Song, J.W. Kim, J. Kuo, A sliding window approach to real-time H.263+ frame rate adjustment, in: Proceedings of the Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems & Computers, vol. 1, November 1998, pp. 858–862.

[14] H.J. Lee, T. Chiang, Y.Q. Zhang, Scalable rate control for very low bit rate (VLBR) video, in: Proceeding of International Conference on Image Processing, vol. 2, 1997, pp. 768–771.

[15] B. Xie, W. Zeng, A sequence-based rate control framework for consistent quality real-time video, IEEE Transactions on Circuits and Systems for Video Technology 16 (1) (2006) 56–71.

[16] L. Overmeire, L. Nachtergaele, F. Verdicchio, J. Barbarien, P. Schelkens, Constant quality video coding using video content analysis, Signal Processing: Image Communication 20 (4) (2005) 343–369.

[17] 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 11 (3) (2001) 345–356.

[18] H.B. Yin, X.Z. Fang, L. Chen, J. Hou, A practical consistent-quality two-pass VBR video coding algorithm for digital storage application, IEEE Transactions on Consumer Electron 50 (4) (2004) 1142–1150.

[19] P.H. Westerink, R. Rajagopalan, C.A. Gonzales, Two-pass MPEG-2 variable-bit-rate encoding, IBM Journal of Research and Development 43 (4) (1999) 471–488.

[20] J.F. Cai, Z.H. He, C.W. Chen, A novel frame-level bit allocation based on two-pass video encoding for low bit rate video streaming applications, Journal of Visual Communication and Image Representation 17 (4) (2006) 783–798.

[21] L. Xu, D.B. Zhao, X.Y. Ji, L. Deng, S. Kwong, W. Gao, Window-level rate control for smooth picture quality and smooth buffer occupancy, IEEE Transactions on Image Processing 20 (3) (2010) 723–734.

[22] D. Bagni, B. Biffi, R. Ramalho, A constant-quality single-pass VBR control for DVD recoders, IEEE Transactions on Consumer Electronics 49 (3) (2003) 653–662.

[23] B.C. Song, K.W. Chun, A one-pass variable video coding for storage media,, IEEE Transactions on Consumer Electronics 49 (3) (2003) 689–692.

[24] N. Mohsenian, R. Rajagopalan, C.A. Gonzales, Single-pass constant- and variable-bit-rate MPEG-2 video compression, IBM Journal of Research and Development 43 (1999) 4.

[25] P. Zhu, W.J. Zeng, C.W. Li, Joint design of source rate control and QoS-aware congestion control for video streaming over the Internet, IEEE Transactions on Multimedia 9 (2) (2007) 366–376.

[26] Z.H. He, D.P. Wu, Linear rate control and optimum statistical multiplexing for H.264 video broadcast, IEEE Transactions on Multimedia 10 (7) (2009) 1237–1241.

[27] Y.S. Huang, S.W. Mao, S.F. Midkiff, A control-theoretic approach to rate control for streaming videos, IEEE Transactions on Multimedia 11 (6) (2009) 1072–1081.

[28] J. Viéron, C. Guillemot, Real-time constrained TCP-compatible rate control for video over the Internet, IEEE Transactions on Multimedia 6 (4) (2004) 634–646.

[29] J.R. Corbera, P.A. Chou, S.L. Regunathan, A generalized hypothetical reference decoder for H.264/AVC, IEEE Transactions on Circuits and Systems for Video Technology 13 (7) (2003) 674–687.

[30] N. PeterFreund, Time-Shift Causality Constraint on the CAT-LB HRD, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 5th Meeting, Doc. JVT-E133, Geneva, October7–14, 2002.

[31] E. Viscito, HRD and Related Issues, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 4th Meeting, Doc. JVT-D131, Klagenfurt, Austria, July 2002.

[32] L.J. Yuan, W. Gao, Y. Lu, Latest arrival time leaky bucket for HRD constrained video coding, in: Proceedings of the IEEE ICME 2003, Baltimore, MD, USA, July 2003, pp. 773–776.

[33] Joint Video Team (JVT), H.264/AVC reference software version 14.0, downloadable at 〈http://iphome.hhi.de/suehring/tml/download/old_jm〉.

[34] D.D. Zhang, K.N. Ngan, Z.Z. Chen, A two-pass rate control algorithm for H.264/AVC high definition video coding, Signal Processing: Image Communication 24 (5) (2009) 357–367.

[35] 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 IEEE PCS 2007, November 2007.

[36] H.S. Wang, N. Moayeri, Finite-state Markov channel-a useful model for radio communication channels,, IEEE Transactions on Vehicular Technology 44 (1) (1995) 163–171.