

Content-Aware Adaptive Media Playout Controls for Wireless Video Streaming

Hsiao-Chiang Chuang, *Member, IEEE*, ChingYao Huang, *Member, IEEE*, and Tihao Chiang, *Senior Member, IEEE*

Abstract—Video streaming is one of the killer applications for cellular communications. The MPEG-4 fine-granularity scalability video coding technique can adapt to bandwidth variation and random packet errors. In this paper, to explore the impacts of cellular channel characteristics on the tolerance of buffer performance and quality of service, a novel statistical model-based adaptive media playout (AMP) is proposed by utilizing the statistical assumptions of both arrival and departure processes for a better decision on the dynamic threshold adjustment and frame-rate adjustment. Based on third-generation cellular transmission environment, simulation results will demonstrate that as compared to other AMP schemes, the proposed AMP control provides better visual quality with lower complexity.

Index Terms—Adaptive media playout (AMP), buffers, cdma2000 1×-RTT, scalable video streaming.

I. INTRODUCTION

THE third-generation wireless communication technologies have been developed for years. The high data throughput of third-generation networks enables several data applications, such as web browsing, network games, background download of e-mails and files, and multimedia streaming. Among these applications, multimedia streaming has the most stringent requirements in terms of the transmission bandwidth and delay. To ensure the transmission quality of service (QoS), especially when the radio frequency (RF) condition changes rapidly, a dynamic control is required.

Wu *et al.* [1] addressed the challenges in designs for the rapid changes of the effective transmission characteristics, such as available channel rates and error rates. The rapid variations of the channel condition could cause serious buffer outage at the receiving side. Stockhammer *et al.* [2] explored the required pre-roll time and buffer size for video streaming via variable wireless bit-rate channels. They proposed a buffer design strategy which pre-determines the buffer size and pre-roll time

for a specific bounded reception process. However, the method of choosing a fixed pre-roll time is insufficient for certain streaming scenarios such as live-content streaming. Several channel-adaptive streaming techniques [3] including adaptive media playout (AMP), rate-distortion optimized packet scheduling, and a channel-adaptive packet dependency control are proposed to solve this problem. The AMP is a receiver-based buffer control technique that adjusts the playback frame rate to minimize the probability of buffer outage. Informal subjective tests have shown that the reduction of the playback rate up to 25% is unnoticeable [4]. Typically, the AMP-based buffer control is performed in two steps. The first step is to determine the threshold for control activation. It chooses a suitable threshold of buffer fullness to prevent buffer outage. The second step is to compute the playout rate based on the relationship between current buffer fullness and the pre-determined threshold.

Yuang *et al.* [5] proposed a video smoother based on a threshold selection strategy. The selected threshold is plugged into an exponentially-distributed service time scheduled to determine the next time for video playout. Laoutaris and Stavrakakis [6] addressed the issue of visual quality with adaptive video playout. The receiving buffer is formulated as M/G/1 queue and the state of buffer occupancy can be modeled as a Markov chain. The impact from both the buffer outage and buffer control is merged into a metric, called Variance of distortion of playout (VDoP), to dynamically adjust the playout rate. Kalman *et al.* gave an analytical result for various streaming environments such as archived streaming and live-content streaming [7]. The effect of packet retransmission caused by an error-prone channel was also considered in the system model. Liang and Huang [8] proposed a content-based adaptive media player based on perceived motion energy (PME) computed from the motion activity of a video sequence. The associated AMP-based control uses a distortion function combined both the distortion caused by the control and buffer outage. In [9], a playout buffer algorithm was proposed based on the utilization of the mean opinion score (MOS) function. Yang *et al.* [10] proposed the AMP based on the channel quality to have an early adoption of the playout rate.

Only a few literatures discussed specific algorithm designs for mobile devices while taking the overall system structure into consideration. The motivation of this paper is to design an appropriate algorithm of a playout control for mobile devices such that the smoothness of playback is preserved. This paper proposes a statistical model-based AMP control mechanism by utilizing the statistical assumptions of both arrival and departure processes to decide the buffer threshold and playout rate adjustments. The threshold and playout rates are determined by con-

Manuscript received February 26, 2005; revised November 14, 2005 and March 21, 2007. This work was supported in part by Taiwan MOEA under Grant 95-EC-17-A-01-S1-031, Taiwan MOE ATU Program 95W803C and ZyXEL Corporation, Taiwan. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Qing Li.

H.-C. Chuang was with the Department and Institute of Electronics Engineering, National Chiao Tung University, Hsinchu 30050, Taiwan, R.O.C.. He is now with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: chuangh@purdue.edu).

C. Y. Huang and T. Chiang are with the Department and Institute of Electronics Engineering, National Chiao Tung University, Hsinchu 30050, Taiwan (e-mail: cyhuang@mail.nctu.edu.tw; tchiang@mail.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2007.902884

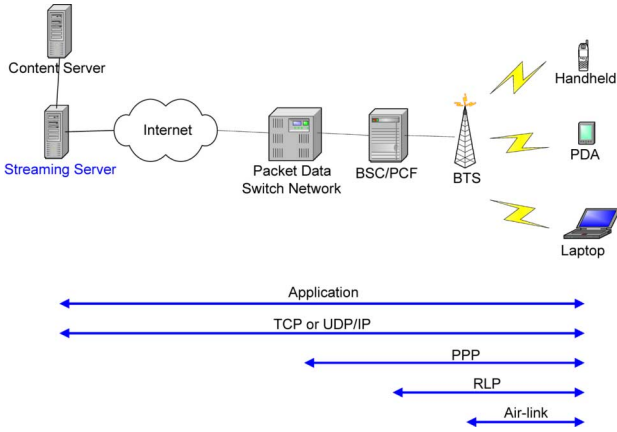


Fig. 1. General scenario for wireless multimedia streaming.

sidering both the elimination of buffer outage and the smoothness of video playback. Besides, the proposed control is acted timely in the unit of a radio frame, which is adequately responsive for the prevention of buffer outage.

The remaining of the paper, which describes the mathematical fundamental and the technical details of the proposed algorithm, is organized as follows. Section II investigates the characteristics of streaming services in a wireless system. The model-based AMP is proposed in Section III. The simulation system model and simulation results are discussed in Section IV. Finally, conclusions are drawn in Section V.

II. AMP-BASED BUFFER CONTROL FOR SCALABLE VIDEO STREAMING OVER WIRELESS SYSTEMS

A. System Models

Fig. 1 illustrates a general scenario for multimedia streaming over a wireless network. The associated protocol stacks are shown in the lower part of the figure. It is evident that a multimedia streaming service requires not only that the wireless connection between the mobile station (MS) and base transceiver station (BTS) is maintained but also that the access to the Internet is supported. The uncertainty of the Internet connection would complicate the discussion of the multimedia streaming. In this paper, our focus is on the last-mile transmission by assuming the streaming server is located in front of the Packet Data Switch Network (PDSN), a data serving node used in the cdma2000 wireless network.

Typically, there are two rate controllers in the transmission path for video streaming: one is the source rate controller and the other is the air-link rate controller (rate assignment), as illustrated in Fig. 2. The source rate controller should send out data based on the channel condition of the network components including the Internet, base-station queue (BSQ), and the available data rate provided by the base station controller (BSC). If there is no in-time negotiation protocol between source and air-link rate controllers, or if the roundtrip delay is relatively large as compared to the required response time, the behavior of the source rate controller can be assumed to be independent from that of the air-link rate controller. Typically, these two rate controllers are equipped in different modules. Thus, we assume

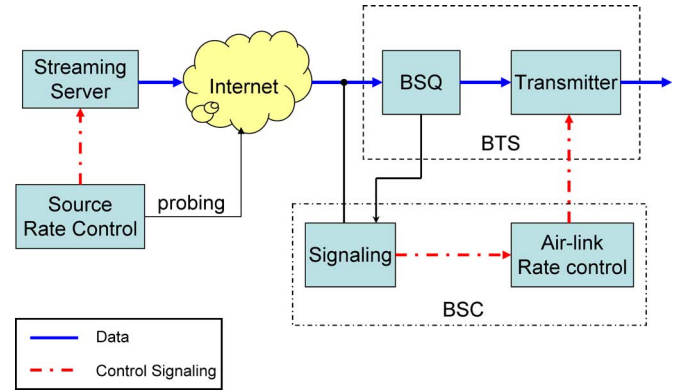


Fig. 2. Two rate controllers in the transmission path for video streaming application.

TABLE I
PARAMETERS THAT INFLUENCE THE PROBABILITIES OF BUFFER OUTAGE

Departure rate			
$\mu(t) \uparrow$	$P(\text{underflow}) \uparrow$ $P(\text{overflow}) \downarrow$	$\mu(t) \downarrow$	$P(\text{underflow}) \downarrow$ $P(\text{overflow}) \uparrow$
Arrival rate			
$\lambda(t) \uparrow$	$P(\text{underflow}) \downarrow$ $P(\text{overflow}) \uparrow$	$\lambda(t) \downarrow$	$P(\text{underflow}) \uparrow$ $P(\text{overflow}) \downarrow$
3. MSB size			
$B \uparrow$	$P(\text{underflow})$ $P(\text{overflow}) \downarrow$	$B \downarrow$	$P(\text{underflow})$ $P(\text{overflow}) \uparrow$
4. Pre-roll ratio			
$Pr \uparrow$	$P(\text{underflow}) \downarrow$ $P(\text{overflow}) \uparrow$	$Pr \downarrow$	$P(\text{underflow}) \uparrow$ $P(\text{overflow}) \downarrow$

that the source rate controller and air-link rate controller are independent in this paper.

B. Buffer Sizes

There are mainly two buffers in the transmission path for video streaming. In this paper, it is assumed that the BSQ assigned for each video-streaming user is large enough such that BSQ never overflows. As for the mobile-station buffer (MSB), the internal application memories in the state-of-art 3G handsets could be ranged from several megabytes to hundred megabytes. In this paper, we assume that the case of buffer overflow would not occur during the streaming process.

The allocation of a larger dynamic memory for MSB reduces the probability of buffer outage and overflow but at the same time, it could cause a longer bitstream pre-rolling time. Here, we assume that an effective buffer size of 200 kB for video streaming application will be used in a mobile station. The issue of the required buffer size has been studied in [11].

C. Exponentially-Distributed Buffer Underflow

To observe the probability of underflow, we consider parameters that would introduce buffer underflow. Table I shows four parameters that would cause buffer outage. In the following texts, we will investigate how these parameters influence the dynamics of buffer fullness.

Fig. 3 shows the probabilities of underflow versus the buffer fullness at various source rates. Intuitively, if the buffer fullness

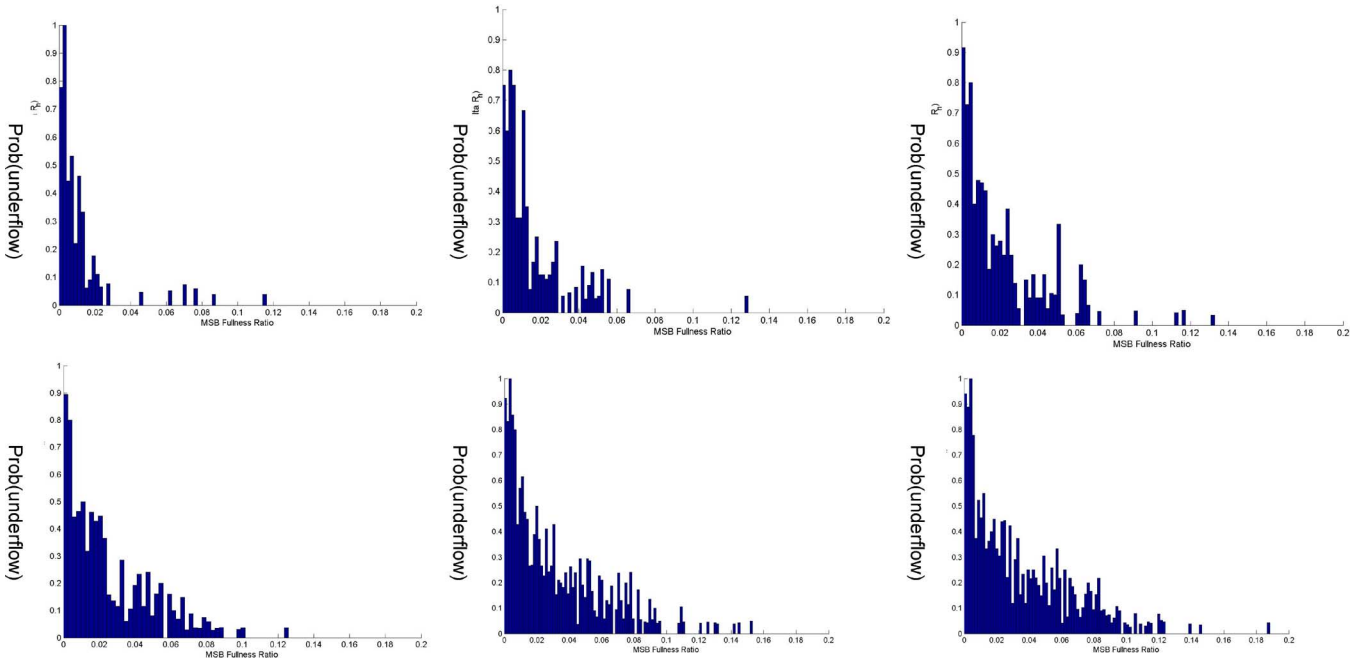


Fig. 3. Probabilities of underflow versus buffer fullness at various source rates (average departure rate): (a) 30 kbps; (b) 40 kbps; (c) 50 kbps; (d) 60 kbps; (e) 70 kbps; and (f) 80 kbps.

is lower, there is a higher probability that the buffer will be underflow. Usually, the handling procedure of buffer underflow is to halt the playback and waits for the duration of the pre-roll time. Here, we suppose that the buffer underflow occurs when the virtual decoder could not retrieve a complete video frame from MSB at a specific time instant. From both Fig. 3 and the cdma2000 transmission model [12], we could model the probabilities of underflow at different fullness by using exponential distribution curves and are formulated as

$$P(\text{underflow, fullness} = F) \cong \alpha \exp(-\alpha F). \quad (1)$$

The value of α represents the probability of underflow when the buffer fullness is equal to zero

$$\alpha = P(\text{underflow, fullness} = 0). \quad (2)$$

This exponential curve could model the probability distribution of the underflow event successfully, but it requires a numerical scaling factor β such that the unit of buffer fullness F can be in bits. Therefore, (1) would become

$$P(\text{underflow, fullness} = F') \cong \alpha \cdot \exp(-\beta \cdot \alpha \cdot F') \quad (3)$$

where the unit of F' can be arbitrary since its unit can be compensated by β .

Let $\lambda(t)$ and $\mu(t)$ be the arrival and departure processes respectively. The probability of underflow can be expressed as

$$\begin{aligned} P(\text{underflow, fullness} = F(t)) &= P(\Delta R(t) < -F(t)) \\ &= P(\lambda(t) < \mu(t) - F(t)) \end{aligned} \quad (4)$$

where $\Delta R(t) = \lambda(t) - \mu(t)$. Because the arrival rate is given by the air-link rate controller and the departure rate is controlled by

the source rate controller, we could assume that the arrival and departure processes are independent. The probability defined in (4) can be simplified and be re-written as

$$\begin{aligned} &P(\lambda(t) < \mu(t) - F(t)) \\ &= \sum_{i=1}^{\infty} P(\lambda(t) < \gamma_{i-1}, \mu(t) - F(t) = \gamma_{i-1}) \\ &= \sum_{i=1}^{\infty} P(\lambda(t) < (i-1) \cdot d\gamma, (i-1) \cdot d\gamma < \mu(t) - F(t) < i \cdot \gamma) \\ &\approx \sum_{i=1}^N P(\lambda(t) < \gamma_{i-1}, \gamma_{i-1} < \mu(t) - F(t) < \gamma_i) \end{aligned} \quad (5)$$

where $d\gamma$ is an infinitesimal interval of the arrival or departure rate, and γ_i is the boundary value for segmenting the arrival or departure curve into finite intervals. The approximation may lose some accuracy in the estimation for the probability of underflow, but the loss would be negligible when N is large enough. Thus, to improve the accuracy in estimation, the number of interval, N , should be set larger. In theory, the probability of underflow will be changed whenever the fullness is changed. Nevertheless, the evaluation of the distribution curve would cost a lot in computational complexity when N is large. As a result, in order to keep both estimation accuracy high and computational complexity low, the frequent updating of the distribution-curve should be avoided.

To reduce the computational complexity of estimating the distribution curve for buffer underflow, we utilize the approximation function for the underflow probability model defined in (3). Let the probability of underflow equal $P_u(F)$. The associated probability density function $pdf_u(F)$ will therefore be

$$pdf_u(F) = \frac{P_u(F)}{\int_0^{\infty} P_f(F) dF} = \alpha \beta \cdot e^{-\alpha \beta \cdot F}. \quad (6)$$

To estimate these two parameters in the probability density function defined in (6), we can exploit the statistical relationship between the arrival and departure processes. By comparing (4) with (3), we find that there are two boundary conditions for evaluating the parameters α and β in (3)

$$\begin{aligned} \alpha &= P(\text{underflow, fullness} = 0) \\ &= P(\lambda(t) < \mu(t)) \\ &\approx \sum_{i=1}^N P(\lambda(t) < \gamma_{i-1}, \gamma_{i-1} < \mu(t) < \gamma_i) \\ &= \sum_{i=1}^N P(\lambda(t) < \gamma_{i-1}) \\ &\quad \times P(\gamma_{i-1} < \mu(t) < \gamma_i) \end{aligned} \quad (7)$$

$$\begin{aligned} \int_{\mu_{\max}}^{\infty} pdf_u(f) df &= \int_{\mu_{\max}}^{\infty} \alpha \beta \cdot e^{-\alpha \beta \cdot f} df = \delta \\ \Rightarrow \beta &= \frac{-1}{\alpha \cdot \mu_{\max}} \ln(\delta) \end{aligned} \quad (8)$$

where δ represents the controlled probability of underflow while the buffer fullness exceeds the maximum departure rate μ_{\max} ever recorded. It is reasonable to set δ as a small number. As long as (7) and (8) are both solved, we can obtain the exponential function to approximate the probability distribution curve of buffer underflow.

III. PROPOSED STATISTICAL MODEL-BASED ADAPTIVE MEDIA PLAYOUT (SM-AMP) [13]

The SM-AMP algorithm is based on a simple idea: adjust the frame rate such that the probability of buffer outage is eliminated. However, the variation of the playback video frame rate causes the degradation on the subjective quality. The challenge of the buffer control is to optimize the visual quality in a time-varying channel environment. The proposed statistical model-based AMP resolves the problem of buffer underflow statistically and takes the smoothness of video playback into account at the same time. In the following subsections, we will detail the proposed buffer control algorithm and some relevant AMP-based control algorithms.

A. Mathematical Model for Buffer Controls

The statistical model-based AMP buffer control mainly utilizes the parameters estimated by (7) and (8). The approximation function with these two parameters would be depicted in Fig. 4. To perform this buffer control, a quality factor P_T should be chosen in advance to ensure that the probability of underflow can be controlled at P_T and below. Once this quality parameter has been determined, the threshold of fullness, F_{th} , could be computed from the (6) by simply taking the integration to fullness F . Note that the update of the approximation function should be done only when some of the boundary conditions are changed. For example, once the maximum departure rate μ_{\max} is changed, the departure distribution function needs to be modified accordingly. Therefore, we should re-compute

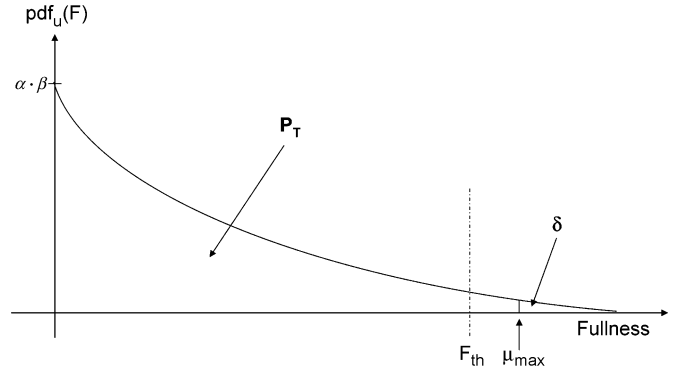


Fig. 4. Approximation function for probability of buffer underflow.

the parameters of the approximation function for the underflow probability. Moreover, the number of segments N would also affect the modeling accuracy of buffer outage, which results in a tradeoff between the computational complexity and the modeling precision.

Based on the similar assumptions made in the previous section, the steady-state arrival process could be analogous to a truncated Gaussian process, which, in the cdma2000 system, is bounded by the minimum and maximum air-link channel throughput. For the departure rate, in MPEG-4 FGS video-coding structure, the base-layer bitstream has more chance to be transmitted than the enhancement-layer bitstream. Thus, it is more likely for the receiver to obtain low-rate video content, and hence the departure process should be modeled as a triangular distribution rather than a uniform distribution.

The arrival process is bounded by the maximum and minimum channel throughput supported by the system over a period of time. Take the cdma2000 system as an example: the maximum channel throughput in 100 ms would be 163.2 kbps ($= 153.6 + 9.6) * 0.1 s = 16.3$ kbits, while the minimum channel throughput would be $9.6 \text{ kbps} * 0.1 s = 960$ bits. We could utilize the predetermined system parameters of channel throughput to formulate the arrival distribution. In the meantime, the departure process is bounded by the source transmission rate, which could be determined by the detection of the backhaul network.

B. Statistical Model-Based AMP Buffer Control

An AMP-based buffer control could in general be split into two parts. The first part is the threshold adjustment, and the second part is the frame-rate adjustment. The objective of the threshold adjustment is to determine a proper activation threshold to which the associated frame rate adjustment is applied. The frame-rate adjustment is to decide when the next video frame should be retrieved and displayed. The proposed architecture of the buffer control mechanism is depicted in Fig. 5. The upper half of this figure is the normal path of uncontrolled playback, and the lower half shows the control mechanism. In order to adjust the frame rate, the buffer monitor keeps tracks of the buffer-related information, such as buffer fullness, arrival statistics, and departure statistics. The output of the control entity is the estimated playtime of the next video frame. The control is accomplished by enabling/disabling the

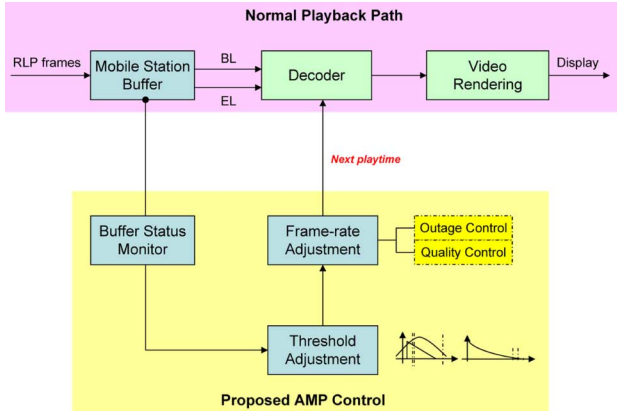


Fig. 5. Proposed architecture of buffer control mechanism.

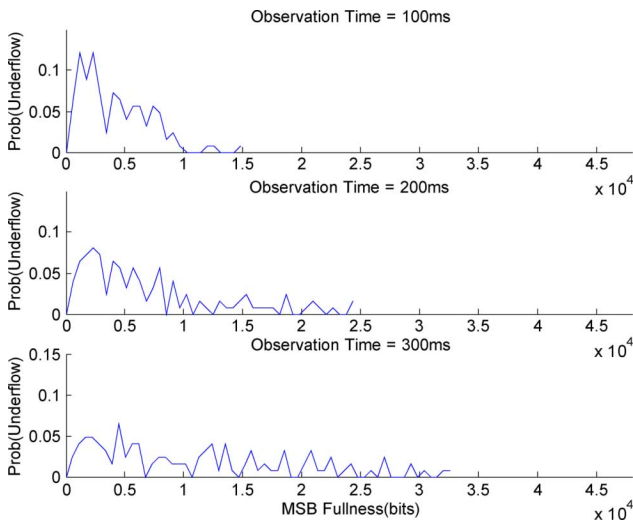


Fig. 6. Illustration of different observation time.

decoding process. In the following subsections, we will discuss the details of the proposed control algorithm.

To apply the control algorithm to the buffer controller, we need to determine the frequency of control activation in advance. There are, in general, two states for the control mechanism to operate. The first one is the normal state, in which the control algorithm checks the current buffer fullness periodically. The threshold will be updated according to the statistics. The second state is the potential outage state, in which the control process is activated more frequently to prevent the buffer from underflow. Fig. 6 shows the impact from different observation time periods. If the observation time period is too long, the probability of underflow does not follow an exponential function but a uniform distribution. To have an effective control, we like to have the probability of underflow followed by an exponential function. Besides, we set the period of the observation time equal to the duration of two consecutive video frames (on the scale of hundreds of millisecond).

C. Outage Controls

Threshold Adjustment: As mentioned previously, the threshold should be re-calculated when parameters for the hypothetical model have been updated. This serves as a re-syn-

TABLE II
IMPACT OF THRESHOLD ON VISUAL QUALITY

Threshold	32k	48k	64k	80k	96k	112k
Q1	1.1592	1.1592	1.1951	1.1951	1.2308	1.2661
Q2	1.8003	1.7306	1.6569	1.5813	1.595	1.5382
Q3	0.7949	0.6981	0.68643	0.62084	0.57663	0.57125

Q1: Mean discrepancy of frame rate
 Q2: Standard deviation of frame rate
 Q3: Short-term standard deviation of frame rate

chronization of a primary threshold adjustment. However, the resulting threshold may last for a long time because the associated traffic parameters, such as μ_{max} , change infrequently during the overall streaming process. In addition to the calculation of the primary threshold, the dynamic threshold adjustment is needed to minimize the impairment of AMP control on the visual quality.

For the dynamic threshold adjustment, we adopt a heuristic scheme based on the observation of the static threshold solution. Let us consider the influence of threshold values on the visual quality first. As shown in Table II, a high threshold causes more unnecessary control activation and lengthens the overall duration of the streaming process (like slow-motion playback). So, we have to determine a proper value for the threshold such that we could obtain good tradeoffs among all quality factors.

We take the trend of buffer fullness as a fundamental reference for the prediction of underflow. The basic idea of this prediction is as follows: once the trend of buffer fullness goes downward, the probability of underflow should arise, and vice versa. Suppose that the primary threshold is a low value, the corresponding variation of the frame rate will be large. This is because the control would not be activated in time. Hence, the threshold should arise when the trend of buffer fullness is decreasing. In that case, the AMP control could be activated earlier to improve the margin. The increase/decrease of threshold would be similar to a scaled “mirror” version of the buffer fullness. Fig. 7 illustrates the flowchart of the proposed threshold adjustment, where the ΔF is the difference of buffer fullness between the current and the last control check time instant, and k is a scaling constant.

Frame-Rate Adjustment: In this subsection, we propose two types of frame-rate adjustment for computing the next playlist. The first type is the stochastic-approaching adjustment (SAA). The basic idea of SAA is that the adjustment of the frame rate is equivalent to the update of the departure distribution. As shown in Fig. 8, the probability density functions of various departure rates imply that the associated threshold should be lower if the maximum departure bit rate, μ_{max} , is reduced. With a fixed control quality parameter P_T , SAA searches for the departure rate such that the corresponding threshold is just lower than current buffer fullness. Moreover, after finding that the desired departure rate should be located between, say, M fps and $M - 1$ fps, the fractional frame rate could be achieved by linear interpolation. The disadvantage of SAA lies in its computational complexity. For example, if the original frame rate is set to 10 fps, each model update will have to compute ten distributions instead of one, which is a large burden due to the power constraint within a mobile station.

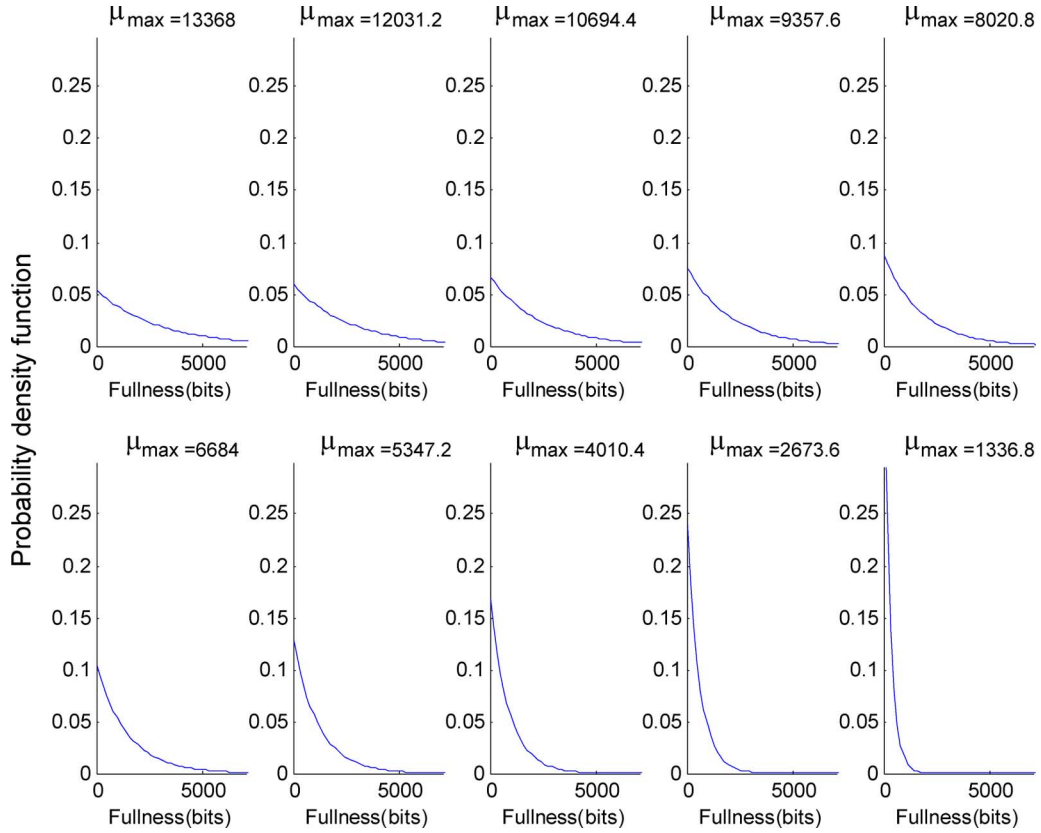


Fig. 8. Distributions of various maximum departure rates at corresponding frame rates.

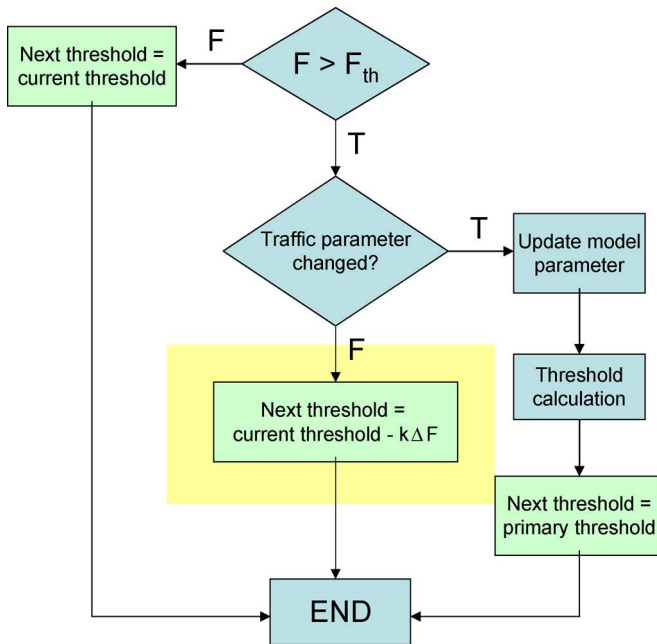


Fig. 7. Flowchart of the proposed threshold adjustment.

The second type of proposed frame-rate adjustment is the content-aware adjustment (CAA). This control utilizes the syntax-level information about the video frame size. Due to the pre-rolling mechanism, the size of each video frame could be recorded while they are pre-rolled into the MSB. This

information is valuable for the estimation of the buffer fullness in the next control checkpoint, which will be

$$F_{\text{next}} = F_{\text{now}} + \left[A_{\text{next}} \cdot \frac{fd_{\text{now}}}{td} - D_{\text{next}} \right] \quad (9)$$

where the capital F represents the buffer fullness. A and D express the amount of arrival (in bits/radio-frame) and departure data (in bits), respectively. The fd is the interval between the current time and the time for outputting the next video frame, and td is the inter-arrival time of radio frames (e.g. radio link protocol (RLP) frame). The value of D_{next} could be acquired from some syntax-level information. Note that the only unknown variable in (9) is the amount of arrival data, A_{next} , which could be predicted using simple estimators such as mean, median, or minimum. If estimated buffer fullness is less than a threshold F_{th} , the CAA will compute the suitable playtime for next video frames by using the following inequality:

$$F_{\text{next}} \geq F_{th} \Leftrightarrow fd_{\text{now}} \geq \frac{F_{th} - F_{\text{now}} + D_{\text{next}}}{A_{\text{next}}} \times td. \quad (10)$$

Once the frame duration is determined, we perform the staircase frame duration adjustment to mitigate the effect of the variation of the frame rate. As shown in Fig. 9, the resumption of the frame rate should be smooth enough to reduce the variation caused by the associated control. One important factor to this quality control lies in the decision of the step size of each control. A large step causes the variation of the playout rate; conversely, a small step makes the resumption of the playout rate slower, which displays the slow-motion playback of received

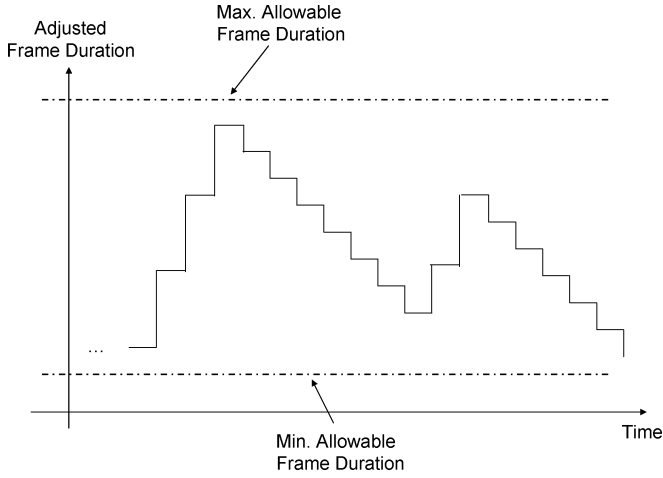


Fig. 9. Smooth adjustment of frame duration.

video frames. Therefore, an inappropriate step incurs the loss in subjective visual quality.

We perform a dynamic step adjustment for the quality control. The step size could be computed based on the following derivation. Let the video frame duration of the current and the next seconds be fd_C and fd_N , a scaling factor for the tolerable variation of the video frame duration, S_T , and the number of the video frames that will be played in the next second, L . There will be a simple relationship between fd_C , fd_N , and S_T

$$fd_N = fd_c \cdot S_T. \quad (11)$$

For a smooth playback, the step size (step) and the number of the steps within 2 s (L) could be computed as follows:

$$\begin{aligned} \text{step} &= \frac{fd_N - fd_c}{L} \\ L &= \left\lfloor \frac{2000}{fd_N + fd_c} \right\rfloor. \end{aligned} \quad (12)$$

By combining (11) and (12), we could obtain the step size from the following formula:

$$\text{step} = \frac{(1 - S_T) \cdot fd_c}{\left\lfloor \frac{2000}{(1 + S_T) \cdot fd_c} \right\rfloor} \quad (13)$$

where the unit of fd_C is in milliseconds, corresponding to the value of 2000 (2 s). Moreover, the value of S_T could be empirically determined for various types of the sequences. For example, motion jitter would be more easily perceived in a high-motion sequence than in a slow-motion one. Hence, the value of S_T should be lower for a high-motion sequence. A more proper selection of the value for S_T could be done by automatically taking the motion information of a sequence into consideration. For example, it is unnoticeable if the reduction in the playout speed is within 25% [4]. Hence, the value of S_T may be set at 0.75 for normal sequences.

There are three states for the buffer control. The state zero is used for updating the distribution curve when either the arrival

process or departure process changes its statistical characteristics. State one is the state for the normal operation of the decoder. State two handles underflow events with a proper choice of the playout frame rate. Each state transition may adjust the frame rate accordingly. The proposed buffer control is based on the probability models of arrival and departure processes. With appropriate approximations for these two probability distributions, the resulting estimation of the probability of buffer underflow would be relatively precise.

IV. SIMULATION MODEL AND RESULTS

In this section, we explore the performance of the proposed statistical AMP-based control scheme. These experiments employ Monte-Carlo method with 50 times of simulation. Furthermore, we assume that the rate assignment for each RLP frame is done every 20 ms, i.e., the transition between two rates could take place once every two consecutive RLP frames. We support an air-link channel with an average throughput of 40 kbps which is typically achievable performance in cdma2000 systems [14]. Before comparing the AMP-based buffer control algorithm, it is important to define the visual quality that will be considered in this study.

A. QoS—Visual Quality

The AMP-based buffer control suffers annoyance from the variation of the playback speed. Hence, we need to address visual quality issues while handling buffer-underflow problems.

Mean Discrepancy From Original Frame Rate: This metric is primarily used to indicate the gap of average frame rate between the original streaming and the distorted streaming, which may be caused by overflow, underflow, or inadequate buffer control. Higher discrepancy implies that the playback process resembles slow-motion version of the original one (for underflow), in the average sense. This metric could be formulated as follow:

$$\text{MDFr} = \frac{1}{N} \sum_{i=0}^{N-1} |OFr(t) - CFr(t)| \quad (14)$$

where $OFr(t)$ is the original frame rate, and the $CFr(t)$ represents the current frame rate.

Standard Deviation of Frame Rate: The variation of the frame rate describes the smoothness of the playback process. Large variation often causes inferior perceptual quality. This metric points to the long-term variation of the entire playback and implies one of the norms for quality-of-service. This metric could be expressed as the following formula:

$$\text{FrStd} = \sqrt{\frac{N \sum_{t=0}^{N-1} [CFr(t)]^2 - \left[\sum_{t=0}^{N-1} CFr(t) \right]^2}{N \times (N - 1)}} \quad (15)$$

Short-Term Standard Deviation of Frame Rate: Human perception prefers to perceive a smooth video playback in a period of time, i.e., a stable frame rate for some time periods. This metric reflects the requirement of stable playback for the human

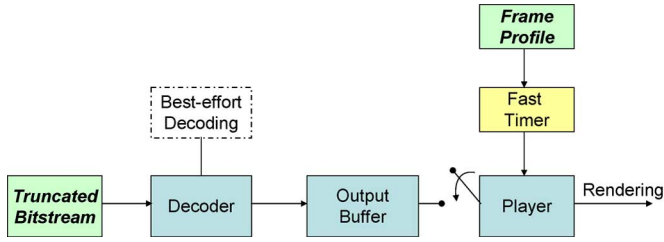


Fig. 10. Architecture of the prototyping system for visualization of AMP-based control algorithms.

TABLE III
ENVIRONMENTAL PARAMETERS FOR THE SUBSEQUENT EXPERIMENTS

Parameter	Value
P_T	0.99
Header compression ratio	5
Buffer size	200kB
Length of video clip	216 sec.
Video format	QCIF (176x144)
Source frame rate	10 fps
Pre-roll time	1 sec.
Avg. source rate	30 kbps
Avg. channel throughput	40 kbps

perceptual quality. The metric can be written as the following expression:

$$Std = \sum_{i=0}^{M-1} Std \{CFr(i \times W), \dots, CFr(i \times W + (w-1))\} \quad (16)$$

$$M = \lceil N/W \rceil$$

where the function Std computes the standard deviation of the frame rate inside a sliding window of time, the symbol w is the size of the sliding window, N is the overall playtime for the associated video streaming, and the value of M is the total number of the computational sliding window.

B. Prototyping System for Visualization of AMP Algorithm

Fig. 10 depicts the modified client-side architecture from the MPEG-21 Testbed [15] to visualize the effect of AMP-based control algorithm. The inputs are the truncated bitstream from the streamer of MPEG-21 Testbed and the frame profile from the cdma2000 1x-RTT simulator. Without modification, the decoder operates at its full speed of a thread to decode the incoming bitstream. The reconstructed video frames are then put into the output buffer and waiting to be fetched by the player. Moreover, the fetch time of the player is controlled by a fast timer, which reads the frame profile and performs an accurate timing control. The timing resolution of the fast timer could be in milliseconds, which is sufficiently precise to support the proposed control mechanism. Some common environmental parameters for the subsequent experiments are defined in Table III.

C. Comparison of Frame-Rate Adjustment Schemes

Besides what we have proposed, SAA and CAA, to compare the performance of different controls, we also adopt other two types of frame-rate adjustment schemes. The first type is a simple staircase adjustment, and the second type is the linear

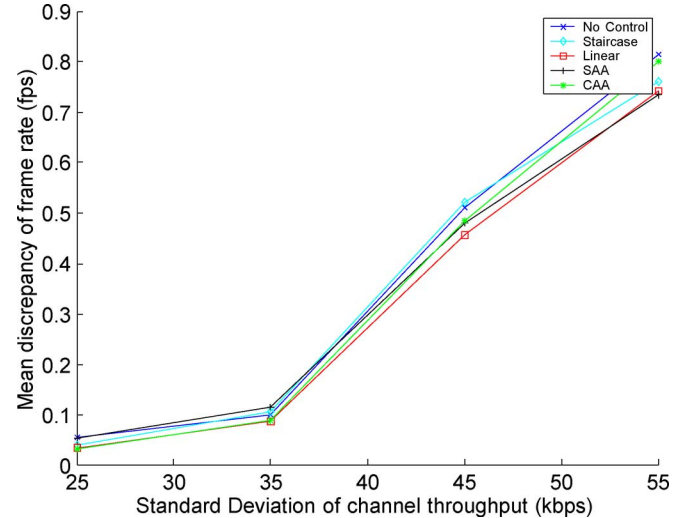


Fig. 11. Mean discrepancy among various frame-rate adjustment schemes.

adjustment. Specifically, the staircase adjustment increases or decreases the frame rate when the current buffer fullness is less than a specific threshold. The linear adjustment simply maps the current fullness into a corresponding frame rate in a linear sense. This adjustment can be described as follows:

$$Fr(t) = \frac{F(t)}{Th(t)} OFr(t) \quad (17)$$

where $Fr(t)$ is the resulting frame rate, $F(t)$ is the current buffer fullness, and $Th(t)$ is the current threshold.

First, we compare the performance of the raw frame-rate adjustment schemes, i.e., we turn off the use of dynamic threshold adjustment and the quality control to observe the nature of the associated schemes. The short-term standard deviation is calculated with the window size of 3 s. Fig. 11 shows the mean discrepancy of the frame rate, which implies the mean latency of the playback using some control schemes. The horizontal axis of the plot is the standard deviation of the channel throughput (abbreviated as channel_STD in the following text) with the mean throughput defined in Table III. As expected, a higher channel variation will result in the higher probability of buffer outage. Based on the average performance, we have observed that all the curves are nearly overlapped, and hence their latencies are almost the same during the whole streaming process. This indicates that the system occupancy time is similar among these control algorithms and the system capacity would not be reduced.

To actually compare the performance of the control in smooth video playback, we choose the long-term and short-term standard deviation of the frame rate as an indication of smoothness. In general, a larger value of the standard deviation implies more motion jitters in the whole playback process, and vice versa. Fig. 12 depicts the corresponding performance for different control schemes. We can see that the content-aware frame-rate adjustment could provide comparable quality to the sophisticated SAA frame-rate adjustment, because the content-aware scheme could prevent the fullness below the threshold at each control checkpoint by the use of syntax-level information (size of each

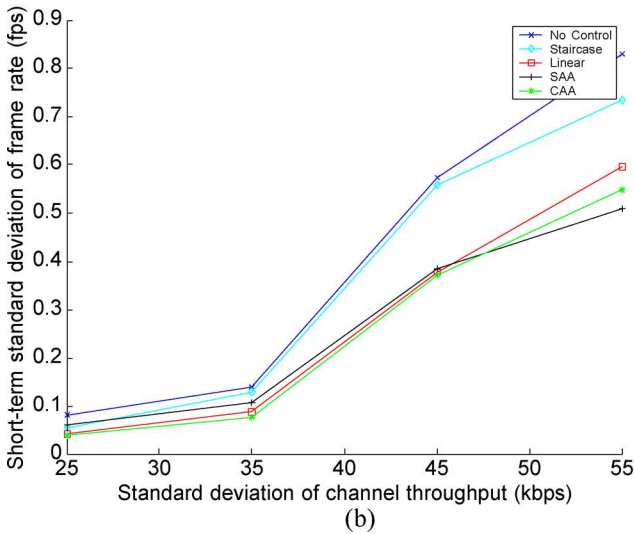
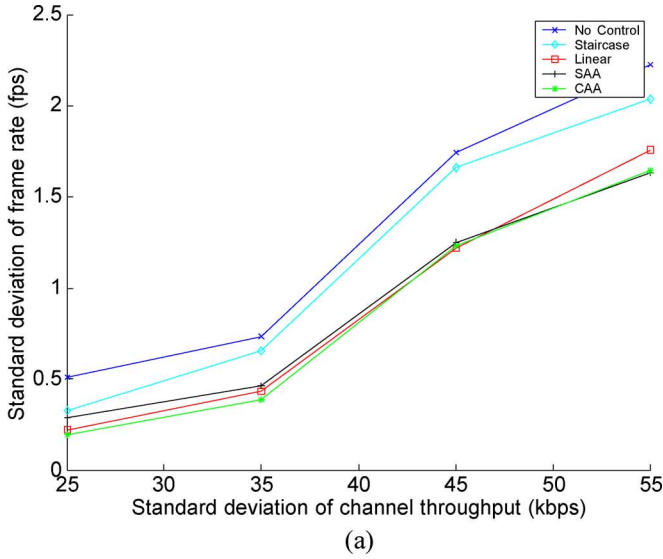


Fig. 12. Comparison of (a) long-term and (b) short-term standard deviation among various frame-rate adjustment schemes.

video frame). This makes the departure process become a deterministic process. This prior knowledge simplifies the process of buffer fullness into a single random process (arrival process), which is easier to estimate. The accurate estimation of buffer fullness gives a more precise result of the frame-rate adjustment, as represented in (10). Hence, the content-aware control can benefit from low complexity and outperforms other control algorithms.

Fig. 13 shows the associated number of underflow among these control algorithms. It is obvious that all algorithms resolve the buffer outage successfully even the associated visual quality might be different among these control schemes. The merit of a good frame-rate adjustment scheme lies in its treatment to potential buffer underflow, since they share an identical threshold value.

D. Performance of Pre-Processing and Post-Processing Tools

We have seen that the content-aware frame-rate adjustment (CAA) can provide good visual quality with reasonable complexity. Here, as shown in Fig. 14, the pre-processing tool,

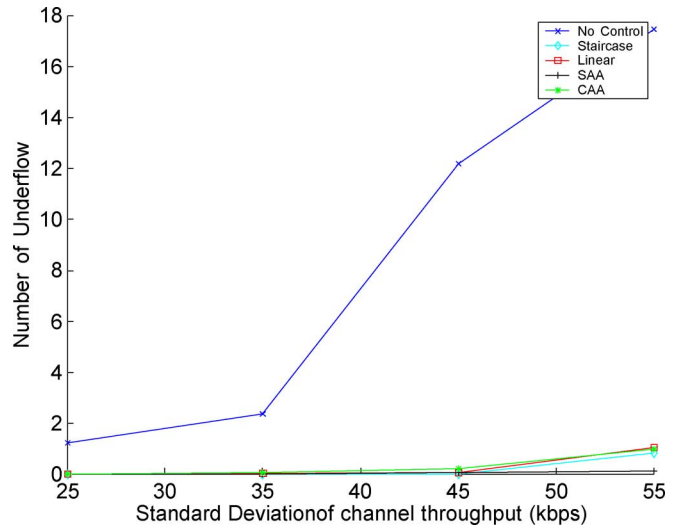


Fig. 13. Comparison of underflow events among various frame-rate adjustment schemes.

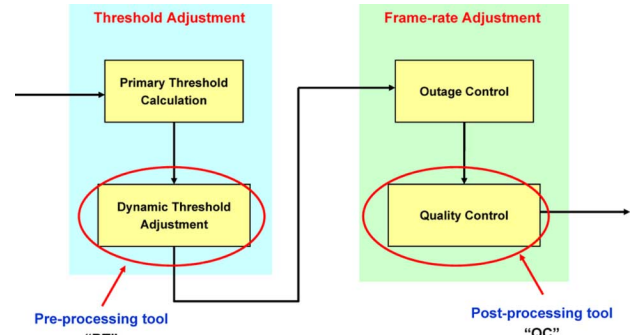


Fig. 14. Framework of the proposed control mechanism.

called dynamic threshold adjustment, and the post-processing tool, called temporal visual quality control, are added on top of the baseline CAA to further improve the visual quality. As discussed, the primary objective of the proposed control mechanism aims to eliminate buffer outage, the term “pre-” and “post-” are used to modify the additional fine-tuning before and after the outage control.

As to some parameters related to the post-processing tools, we choose the scaling factor k as 0.2, and the size of the window for the buffer trend is set as the length of pre-roll time (1 s, in this case). Similarly, for the post-processing tool, the maximum allowable value of the increased frame-duration (to prevent buffer from going underflow) is set to 40 ms, and the value of S_T is set to 0.75. The effects of the pre-processing and post-processing tools are shown in Fig. 14. There are four symbols defined in Fig. 15: the “Original” means that both tools are turned off, the “QC Only” means we only turn on the quality control, the “DT Only” stands for the activated dynamic threshold adjustment, and the “DT+QC” represents that both tools are applied. As we can see from Fig. 15, when the channel STD equals to 55 kbps, the original CAA frame-rate adjustment suffers about one buffer underflow in average, while the other three control scenarios incur buffer underflow below 0.5 times per

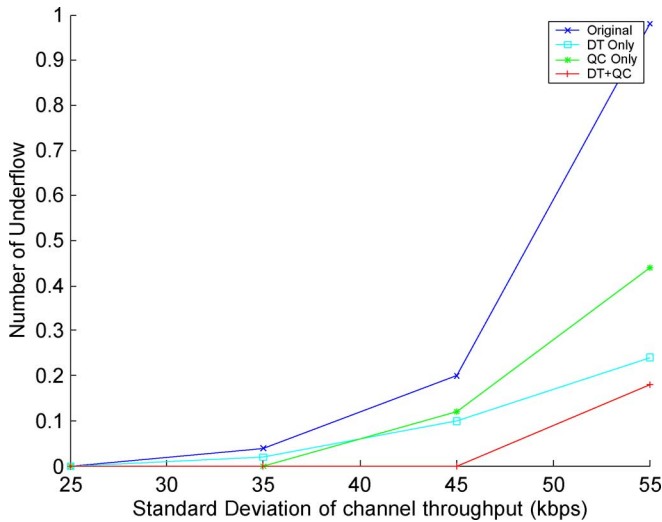


Fig. 15. Comparison of underflow events among various control scenarios.

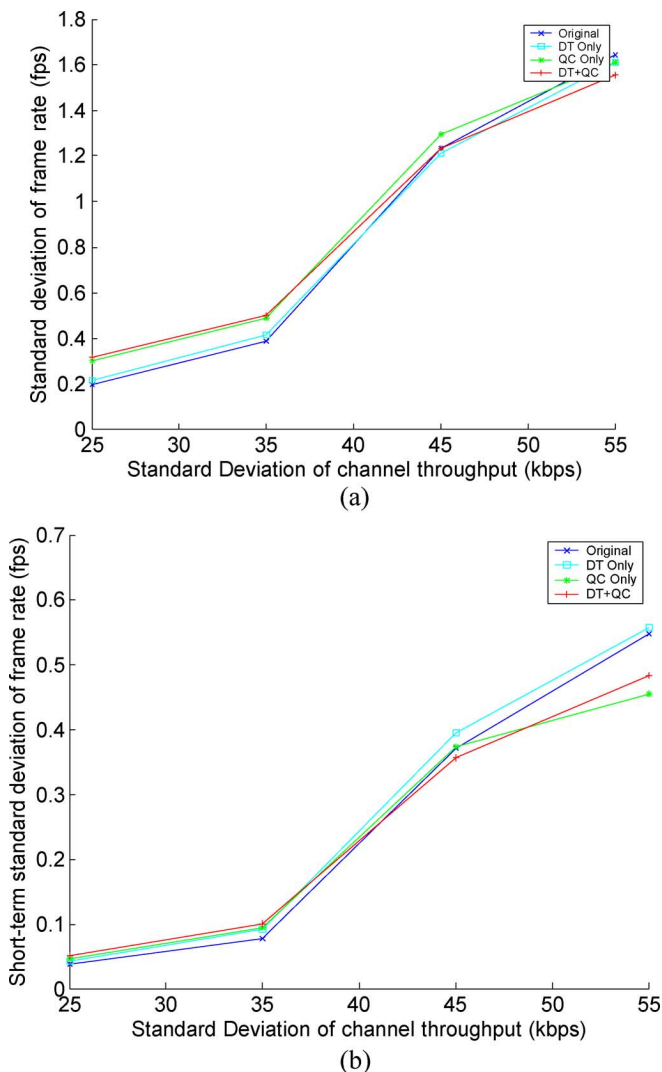


Fig. 16. Comparison of (a) long-term and (b) short-term standard deviation among various control scenarios.

streaming. The control with both DT and OC has the best control of the buffer underflow. Fig. 16 shows both the long-term and short-term variation of the frame rate. Obviously, the processing

tools suffer more quality degradation when the channel STD is low (25~35 kbps), since there are more ineffective control activated when either QC or DT is turned on. Strictly speaking, the QC would suffer more quality degradation, in terms of standard deviation, than DT does. This is because DT could be released from the control activation more quickly when the buffer fullness arises. On the other hand, the short-term variations of the frame rate are not far off since the QC operates in a smooth way for frame-rate adjustment. In this case, QC can outperform DT. This is because the frequent on-and-off control activation caused by DT increases the effective short-term variation of the frame rate in the case of frequent buffer underflow.

V. CONCLUSION

In this paper, we derive an analytical formula to estimate the underflow probability of various buffer fullness. The derivation is based on statistical assumptions for the arrival and departure processes. With the associated traffic statistics and the defined probability of the buffer underflow, the primary threshold can be calculated. To further ensure the visual quality with lower complexity in designs, a dynamic threshold control and quality controls are proposed. The dynamic threshold is calculated based on the buffer fullness while the quality control is to smooth the visual perception. Results shows that with the proposed controls, the perceived visual quality can be improved in terms of the long-term and short-term variation of the video frame rate of streaming. The flexible combination of the pre-processing (dynamic threshold calculation) and the post-processing (quality control) tools can be decided based on an accurate probe of channel characteristics and user's preferences.

REFERENCES

- [1] D. Wu, T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the internet: Challenges and approaches," *Proc. IEEE*, vol. 88, no. 12, pp. 1855–1875, Dec. 2000.
- [2] T. Stockhammer, H. Jenkač, and G. Kuhn, "Streaming video over variable bit-rate wireless channels," *IEEE Trans. Multimedia*, vol. 6, no. 2, pp. 268–277, Apr. 2004.
- [3] B. Girod, M. Kalman, Y. J. Liang, and R. Zhang, "Advances in channel-adaptive video streaming," in *Proc. IEEE ICIP'02*, 2002, vol. 1, pp. 9–12.
- [4] M. Kalman, E. Steinbach, and B. Girod, "Adaptive playout for real-time media streaming," in *Proc. IEEE ISCAS'02*, 2002, vol. 1, pp. 45–48.
- [5] M. C. Yuang, S. T. Liang, and Y. G. Chen, "Dynamic video playout smoothing method for multimedia application," in *Proc. IEEE ICC'96*, 1996, vol. 3, pp. 1365–1369.
- [6] N. Laoutaris and I. Stavrakakis, "Adaptive playout strategies for packet video receivers with finite buffer capacity," in *Proc. IEEE ICC'01*, 2001, vol. 3, pp. 969–973.
- [7] M. Kalman, E. Steinbach, and B. Girod, "Adaptive media playout for low-delay streaming over error-prone channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 6, pp. 841–851, Jun. 2004.
- [8] C. H. Liang and C. L. Huang, "Content-based adaptive media player for network video," in *Proc. IEEE ISCAS'04*, 2004, vol. 3, pp. 749–752.
- [9] K. Fujimotol, S. Ata, and M. Murata, "Adaptive playout buffer algorithm for enhancing perceived quality of streaming applications," *ACM Telecommun. Syst.*, vol. 25, no. 3, pp. 259–271, Mar. 4, 2004.
- [10] Y.-H. Yang, M.-T. Lu, and H. H. Chen, "Smooth playout control for video streaming over error-prone channels," in *Proc. 8th IEEE Int. Symp. Multimedia*, San Diego, CA, 2006, pp. 415–418.
- [11] S.-H. Lee, K.-Y. Whang, Y.-S. Moon, W.-S. Han, and I.-Y. Song, "Dynamic buffer allocation in video-on-demand systems," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1535–1551, 2003.
- [12] H.-C. Chuang, C. Y. Huang, and T. Chiang, "On the buffer dynamics of scalable video streaming over wireless network," in *Proc. IEEE VTC'04A*, fall 2004, vol. 4, pp. 2582–2586.

- [13] H.-C. Chuang, C. Y. Huang, and T. Chiang, "A novel adaptive video playout control for video streaming over mobile cellular environment," in *Proc. IEEE ISCAS'05*, 2005, vol. 4, pp. 3267–3270.
- [14] J. Huang, R. Y. Yao, Y. Bai, and S.-W. Wang, "Performance of a mixed-traffic cdma2000 wireless network with scalable streaming video," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 13, no. 10, pp. 973–981, Oct. 2003.
- [15] C.-N. Wang *et al.*, ISO/IEC JTC1/SC 29/WG 11 M11117: Scalable Multimedia Streaming Test Bed for Media Coding and Testing in Streaming Environments Jul. 2004.

Hsiao-Chiang Chuang (M'02) received the B.S. and M.S. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 2002 and 2004, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN.

During 2002 to 2004, he joined the development of the MPEG-21 part-12 Testbed project. Since 2006, he has been a Research Assistant in the School of Electrical and Computer Engineering at Purdue University. His research interests include video streaming, video coding, audio coding, and image processing.

ChingYao Huang (M'00) received the B.S. degree in physics from National Taiwan University, Taipei, Taiwan, R.O.C., in 1987 and the Master's and Ph.D. degrees in electrical and computer engineering from New Jersey Institute of Technology, University Heights, Newark, NJ, and Rutgers University (WINLAB), Newark, in 1991 and 1996, respectively.

He joined AT&T, Whippany, NJ (later Lucent Technologies) in 1996, as a System Engineer (Member of Technical Staff) for the AMPS/PCS Base Station System Engineering Department. In 2001 and 2002, he was an Adjunct Professor at Rutgers University, Newark, NJ, and New Jersey Institute of Technology, Newark. Since 2002, he has been with the Department of Electronics Engineering, National Chiao Tung University (NCTU), Taiwan, as an Assistant Professor and then Director of the NCTU Technology Licensing Office since 2003. His research areas include wireless medium access controls, radio resource management, and scheduler control algorithms for wireless high-speed data systems. He has published more than 50 technical memorandums, journal papers, and conference papers and is the chapter author of the book "Handbook

of CDMA System Design, Engineering and Optimization". He currently has 12 patents and 20 pending patents.

Dr. Huang was the recipient of the Bell Labs Team Award from Lucent in 2002 and the Best Paper Award from IEEE Vehicular Technology Conference in the fall of 2004. He has served as Editor for ACM WINET and is Symposium Co-Chair and Technical Chair for the Symposium on Multimedia over Wireless 2007 and The First International Conference on Ambient Media and Systems (Ambi-sys) 2008.

Tihao Chiang (S'91–M'95–SM'99) was born in Chiayi, Taiwan, R.O.C., in 1965. He received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Columbia University, New York, in 1991 and 1995, respectively.

In 1995, he joined David Sarnoff Research Center (formerly RCA Laboratory), Princeton, NJ, as a Member of Technical Staff. Later, he was promoted to Technology Leader and Program Manager at Sarnoff. While at Sarnoff, he led a team of researchers and developed an optimized MPEG-2 software encoder. In September 1999, he joined the faculty at National Chiao Tung University, Taiwan, R.O.C. On his sabbatical leave from 2004, he worked with Ambarella USA and initiated its R&D operation in Taiwan. Since 1992, he has actively participated in ISO's Moving Picture Experts Group (MPEG) digital video coding standardization process with particular focus on the scalability/compatibility issue. He is currently the co-editor of Part 7 of the MPEG-4 committee. He has made more than 100 contributions to the MPEG committee over the past ten years and has published over 70 technical journal and conference papers in the field of video and signal processing and holds over 40 patents.. His main research interests are compatible/scalable video compression, stereoscopic video coding, and motion estimation.

Dr. Chiang was a co-recipient of the 2001 Best Paper Award from the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. For his work in the encoder and MPEG-4 areas, he received two Sarnoff achievement awards and three Sarnoff team awards. He was an Associate Editor for IEEE TRANSACTION ON MULTIMEDIA and a Guest Editor for IEEE TRANSACTION ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and the *International Journal of Imaging Systems and Technology*. He is Chairman for the Taipei Chapter of the IEEE Signal Processing and IEEE Consumer Electronics societies.