Towards Rate-Distortion Tradeoff in Real-Time Color Video Coding

Zhenzhong Chen, *Student Member, IEEE*, and King Ngi Ngan, *Fellow, IEEE*

***Abstract*—In this paper, we address the key problem in real-time video coding, the rate-distortion (R-D) tradeoff. As most video coding applications employ color images, we analyze the R-D modeling problem of the color image sequence. We investigate that the R-D characteristics of color video signal in traditional transform-based video coding systems should be modeled for the luminance and chrominance components separately. So we propose separable R-D models for color video coding. The feedback from the encoder buffer is analyzed by a control-theoretic adaptation approach to avoid buffer overflow and underflow. To achieve smooth video quality and satisfy the delay constraints in real-time applications, a novel R-D tradeoff controller is designed. In the proposed R-D tradeoff framework, both the quality variation and buffer safety are considered. Extensive simulation results demonstrate that the proposed approach can achieve smooth video quality without sacrificing overall quality while efficiently utilize the bandwidth without buffer overflow or underflow.**

***Index Terms*—Buffer control, H.264, rate control, rate-distortion (R-D) model, R-D tradeoff, video coding.**

I. Introduction

VARIOUS video coding standards have been developed, e.g., ITU-T H.261 [1], H.263 [2], ISO/IEC MPEG-1 [3], MPEG-2 [4], MPEG-4 [5], and H.264/AVC [6], and deployed in the various multimedia applications such as video conferencing, storage video, video-on-demand (VOD), digital television broadcasting, and Internet video streaming. These video coding standards employ efficient compression techniques to remove the spatial and temporal redundancy within and between frames. Real-time video coding becomes more and more important in recent years with the evolution of the advanced video coding standards, e.g., H.264 [6], which provide high compression ratio for the video source information and further improve the efficiency of the communication systems. However, real-time video coding faces an inherent challenge since both the bit rate and picture quality in compressed video sequences are variable. The tradeoff between the rate and distortion is a key issue in these lossy video compression applications.

Variable bit rate (VBR) video coding is employed in some applications in which natural video frames need to be represented with more stable quality. VBR video can also be neatly incorporated in a VBR transmission networking infrastructure [7], [8]. For VBR video where delay or rate constraint is not as strict as real-time video coding, The function of rate control algorithm in VBR video coding aims to determine the quantization parameter to maximize consistent video quality. Besides one-pass rate control algorithms [9] which are suitable for live VBR video encoding, two-pass VBR rate control algorithms can be developed for storage applications [10], [11] which are strictly under budget constraint instead of delay constraint. Using the two-pass approach, the encoder can control quality smoothly and the bit budget precisely. To guarantee the near-constant or constant video quality, the minimum maximum (MINMAX) criteria [12] and minimum distortion variation (MINVAR) criteria [13] are often adopted. However, most of these approaches, such as the bisection method [12], the multistage algorithm [14], [15], the ratio algorithm [14], and the pseudo I-frame methods [14], can be applied in off-line applications only since they need to generate the rate-distortion (R-D) curves of the sequence or GOP first before employing optimization techniques such as the minimum average (MINAVE) distortion based optimization [16]. Yang *et al.* proposed the fast suboptimal algorithms to achieve either MINMAX [17] or MINAVE [18] for embedded wavelet coder which are applicable for real-time video coding. But they are not applicable to the current block-based standard video coders such as MPEG-4 and H.264 in which exploring the operational R-D curves is very time consuming. In practical video coding applications, it is essential to consider the rate/buffer/delay constraints. In this paper, we focus on applicable method for low-delay constant bit rate (CBR) real-time applications.

Rate control always evolves into constrained problems in practical applications [19]. Since the amount of information in compressed video sequences is inherently variable, a buffer is placed between the video encoder and the transmission channel to smooth out the rate variation to satisfy the delay requirement in real-time environment. Larger buffer corresponds to longer end-to-end delay. The encoder buffer is placed between the video encoder and the transmission channel to monitor the generated bit rate and dictate the amount of delay in transmission systems. The quantization parameters are adjusted based on the feedback from the buffer status to match the source bit rate to the transmission bit rate. Many CBR rate control mechanisms have been proposed to provide a satisfactory bit rate control [9], [20]–[27]. In these approaches, the R-D relationships are modeled by the R-D models. From the classical R-D functions [28], many mathematical expressions of R-D characteristics have been developed. Many rate control algorithms assume that the source statistics are stationary and belong to some

Manuscript received January 20, 2006; revised June 15, 2006 and September 16, 2006. This work is supported in part by the CUHK Focused Investment Scheme and affiliated with the Microsoft-CUHK Joint Laboratory for Human-Centric Computing and Interface Technologies. This paper was recommended by Associate Editor R. C. Lancini.

The authors are with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: zchen@ee.cuhk.edu.hk; knngan@ee.cuhk.edu.hk).

Digital Object Identifier 10.1109/TCSVT.2006.888022

probabilities such as Gaussian [29] or Laplacian [21] and then derive the R-D models based on the R-D theory [30]. Ding *et al.* [20] proposed a generic rate-quantizer model according to the changes in picture activity and a feedback-based bit allocation. A re-encoding method was developed with the rate-quantization model. In [29], a source model is derived from the classical R-D theory to describe the relationships between the rate, distortion and quantization parameters whilst in [31] an adaptive model-driven bit allocation method based on a parametric R-D model was proposed which incorporated a region classification scheme. Ribas-Corbera *et al.* [22] presented a logarithmic model for bit rate and distortion and used Lagrange optimization to minimize the distortion subject to the bit rate constraint. A spline method was reported in [13] and a quadratic rate-quantizer model was proposed in [21], respectively. The quadratic rate-quantizer model was further extended into a scalable rate control algorithm in [32] and optimized for its accuracy in [26]. Cheng *et al.* [33] studied the linear relationship between the activity measure and bit rate and derived an empirical first-order bits model. He *et al.* [25] proposed a linear ρ-domain R-D model and the corresponding R-D analysis framework, where ρ is measured by the percentage of zeros of the discrete cosine transform (DCT) coefficients. The mapping between the quantization parameter Q and ρ is calculated based on the distribution of the DCT coefficients.

However, the priority in above approaches is to control the target bit rate. These strict bit rate control schemes result in large quality fluctuations due to the varying content in natural visual scenes. In contrast to these rate control algorithms, a low-pass filter based MPEG compression scheme was proposed by He *et al.* [34] to smooth the video quality for higher bit rate applications. They introduced the smoothed rate control framework by using a geometric averaging filter. By combining the buffer control, the low-pass filtering based approach can achieve smooth quality with limited delay. Xie *et al.* [35] proposed a sequence based rate control by using the rate-complexity model to achieve constant quality in real-time MPEG-4 video coding.

In this paper, we address the R-D tradeoff problems in real-time H.264 video coding to achieve smooth video quality. As most video coding applications employ color images, we analyze the R-D modeling problem of the color image sequences. We found that the distribution of the integer cosine transform coefficients in H.264 color video coding can be approximated by Laplacian distribution but the R-D characteristics of color video signal in traditional transform-based video coding systems should be modeled separately for the luminance and chrominance components. So we propose separable R-D models for color video coding. This separable R-D modeling approach outperforms the traditional non-separable R-D modeling ones. Based on the R-D models and a novel buffer adaptation approach, we propose R-D tradeoff framework by considering both the quality variation and buffer constraint. The R-D tradeoff is utilized via the quality smoothing cost function with reference to the buffer constraint. We conduct extensive tests to justify the performance of our algorithm.

The rest of this paper is organized as follows. Section II provides the analysis on R-D characteristics of color video signals in traditional transform-based video coding systems. We present a novel R-D control algorithm in Section III. The feedback from the encoder buffer is analyzed by a control-theoretic adaptation approach. The adaptation output is used by the R-D controller to achieve tradeoffs between the encoder delay and picture quality fluctuation. Simulation results in Section IV demonstrate that the proposed approach can achieve a smooth video quality without sacrificing overall quality. We conclude in Section V.

(a) (b) (c)

Fig. 1. Luminance and chrominance components of test sequence *Mobile* (QCIF, 4:2:0). (a) Y. (b) U. (c) V.

II. Color Video R-D Modeling

In most video coding techniques, the color videos consist of luminance and chrominance components. Fig. 1 shows the luminance (Y) and chrominance (U, V) components of a sample color video, *Mobile*. The luminance component provides more texture details while the chrominance components are more uniform. Other color video sequences also present similar characteristics. Integer cosine transform (ICT) is used in H.264 which is a close approximation of the traditional DCT transform. In this paper, we analyze the ac coefficients of the 4×4 ICT [36]

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \tag{1}$$

and the 8×8 ICT [37]

$$\mathbf{H} = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix} \tag{2}$$

adopted in H.264 and examine the different distribution characteristics of the luminance and chrominance components. Figs. 2 and 3 show the distribution of the ac coefficients of the ICT for intraframes and interframes of the test video sequence *Mobile*, respectively. Here we also apply the 8×8 ICT for the chrominance components for comparisons. The distributions for the luminance (Y), and the chrominance (U and V) components are plotted separately and approximated by the Laplacian distribution with different λ

$$p(x) = \frac{\lambda}{2} e^{-\lambda |x|}. \tag{3}$$

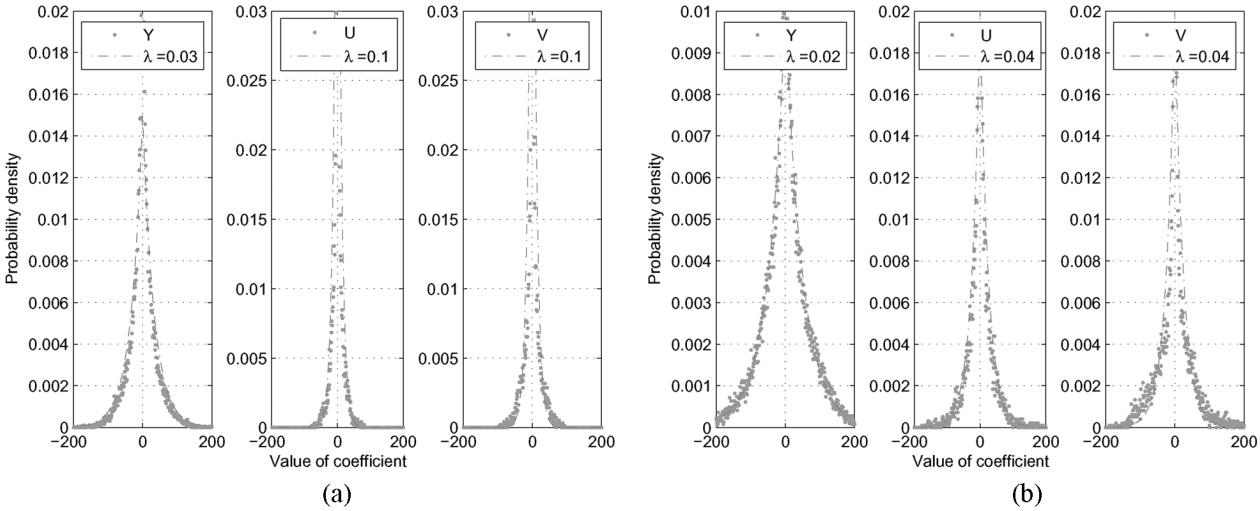


Fig. 2. PDF of integer transform coefficients (1,1) of Y, U, V components of intraframe of test sequence *Mobile* (QCIF, 4:2:0), and approximation by *Laplacian* distribution. (a) 4 × 4 transform. (b) 8 × 8 transform.

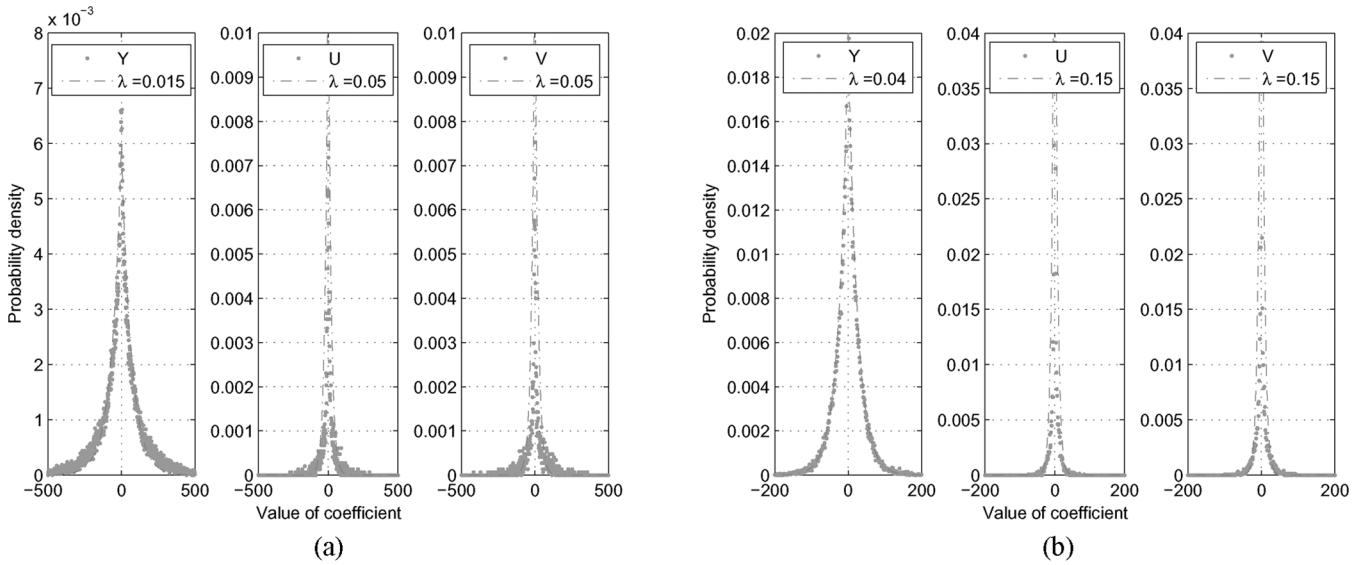


Fig. 3. PDF of integer transform coefficients (1,1) of Y, U, V components of interframe of test sequence *Mobile* (QCIF, 4:2:0), and approximation by *Laplacian* distribution. (a) 4 × 4 transform. (b) 8 × 8 transform.

As shown in Figs. 2 and 3, the distributions of the ac transform coefficients for the two chrominance components (U and V) are close and are different to the luminance component (Y). So we can describe the ac transform coefficients of U and V by the same distribution in further analysis. It is noted that this paper is not to conclude the distribution to approximate the transform coefficients of the transform, but to examine the different characteristics for the distribution of the luminance and chrominance components. Same observation is found for DCT transform based video coding systems, e. g., MPEG-4. So for the DCT-based video coding systems, the conclusion, separately modeling the luminance and chrominance components of color video signal, is valid as well. For the distribution expression of transform coefficient problem, extensive of study has been done in the literature. The readers may refer to these state-of-the-art works [38]–[40] among which the Laplacian distribution based principle has been widely accepted in video coding area.

To further analyze the effects of these different components on video coding, we compress the color video sequence by H.264 codec [41] and plot the R-D curves in Fig. 4(a). As shown in Fig. 4(b), we can find that the shapes of R-D curves of different components are different due to the different representations of luminance and chrominance textures. Fig. 4(b) shows the bit rate curves of different components at a fixed quantization parameter $Q = 32$. Although the bit rate of the chrominance components is lower than that of the luminance component, the performance of the chrominance part is different and this difference should not be ignored. It should also be noted that we use a 4:2:0 format color video as shown in Fig. 1. For other typical color video format such as 4:4:4, the ratio of the the chrominance components in the bit rate is higher. It is noted that this phenomenon can also be found in DCT transform based video coding systems.

As the R-D characteristics of the luminance and chrominance components are different, we build the R-D models for color

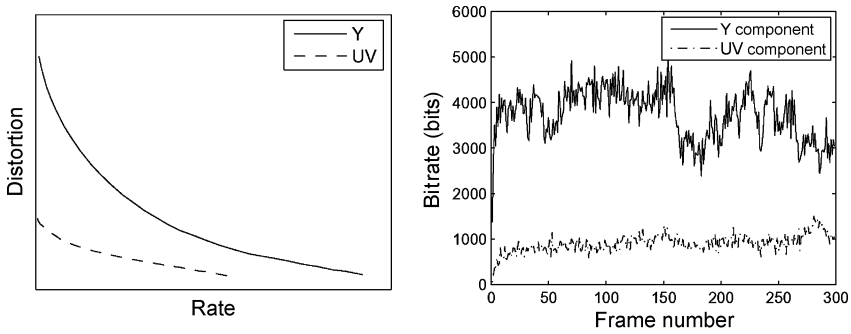


Fig. 4. Test results from test sequence *Mobile* (QCIF, 4:2:0). (a) R-D curves of Y, UV components. (b) Bit rates of Y, UV components at $Q = 32$.

video, separately. Assuming that the source information statistics are the Laplacian distribution, the quadratic R-D models can be derived based on R-D theory [21], [28], [42] for the luminance and chrominance components separately as

$$R_{\text{texture},c} = S_c\left(\frac{a_{1,c}}{Q} + \frac{a_{2,c}}{Q^2}\right) \tag{4}$$

$$D_{\text{texture},c} = b_{1,c}Q + b_{2,c}Q^2 \tag{5}$$

and the R-D models for the color video can be summarized as

$$\begin{aligned} R_{\text{YUV}} &= \sum_{c=1}^{2} R_{\text{texture},c} \\ &= \frac{a_{1,1}S_1 + a_{1,2}}{Q} + \frac{a_{2,1}S_1 + a_{2,2}}{Q^2} \end{aligned} \tag{6}$$

$$\begin{aligned} D_{\text{YUV}} &= \sqrt{\frac{1}{\delta}\sum_{c=1}^{2}\gamma_c (D_{\text{texture},c})^2} \\ &= \sqrt{\frac{1}{\delta}\left((b_{1,1} + \gamma_2 b_{1,2})Q + (b_{2,1} + \gamma_2 b_{2,2})Q^2\right)} \end{aligned} \tag{7}$$

where $R_{\text{texture},c}$ is the bit rate for the luminance texture information $(c = 1)$ or the chrominance texture information $(c = 2)$, $D_{\text{texture},c}$ is the estimated distortion (root mean square error, RMSE) for the luminance texture information $(c = 1)$ or the chrominance texture information $(c = 2)$, R_{YUV} is the bit rate for the color video signal, D_{YUV} is the distortion for the color video signal, S is the luminance texture complexity $(c = 1)$ denoted by MAD (mean absolute difference) which is predicted in H.264 or 1 $(c = 2)$, Q denotes the quantization parameter, $\delta = 1.5, \gamma_1 = 1$, and $\gamma_2 = 0.5$ for 4:2:0 color video. We use the combined-channel peak signal-to-noise ration PSNR_{YUV} [43] for color video signal in this paper. a_1, a_2, b_1, and b_2 are the model parameters updated by linear regression method from the previous coded results. From the N previously encoded frames, we have N data sets $(R_{\text{texture},c,i}, S_{c,i}, Q_i)$ and $(D_{\text{texture},c,i}, Q_i)$. So we can minimize the quadratic error functions

$$\begin{cases} f_{\text{R}}(a_{1,c}, a_{2,c}) = \sum\limits_{i=1}^{N}\left(a_{1,c} + \frac{a_{2,c}}{Q_i} - \frac{R_{\text{texture},c,i}Q_i}{S_{c,i}}\right)^2 \\ f_{\text{D}}(b_{1,c}, b_{2,c}) = \sum\limits_{i=1}^{N}\left(b_{1,c} + b_{2,c}Q_i - \frac{D_{\text{texture},c,i}}{Q_i}\right)^2 \end{cases} \tag{8}$$

where N is the number of selected past frames, Q_i, $R_{\text{texture},c,i}$, $D_{\text{texture},c,i}$, and $S_{c,i}$ are the quantization levels, the actual bit rates, the actual distortion, and the complexities of the luminance or chrominance components in the past frames, respectively. So these parameters can be obtained as

$$\begin{cases} a_{1,c} = \dfrac{\sum\limits_{i=1}^{N} Q_i \frac{R_{\text{texture},c,i}}{S_{c,i}} - a_{2,c}Q_i^{-1}}{N} \\ a_{2,c} = \dfrac{N\sum\limits_{i=1}^{N}\frac{R_{\text{texture},c,i}}{S_{c,i}} - \left(\sum\limits_{i=1}^{N} Q_i^{-1}\right)\left(\sum\limits_{i=1}^{N} Q_i \frac{R_{\text{texture},c,i}}{S_{c,i}}\right)}{N\sum\limits_{i=1}^{N} Q_i^{-2} - \left(\sum\limits_{i=1}^{N} Q_i^{-1}\right)^2} \\ b_{1,c} = \dfrac{\sum\limits_{i=1}^{N} Q_i^{-1} D_{\text{texture},c,i} - b_{2,c}Q_i}{N} \\ b_{2,c} = \dfrac{N\sum\limits_{i=1}^{N} D_{\text{texture},c,i} - \left(\sum\limits_{i=1}^{N} Q_i\right)\left(\sum\limits_{i=1}^{N} Q_i^{-1} D_{\text{texture},c,i}\right)}{N\sum\limits_{i=1}^{N} Q_i^2 - \left(\sum\limits_{i=1}^{N} Q_i\right)^2}. \end{cases} \tag{9}$$

To evaluate the effectiveness of the proposed R-D models in (6) and (7), we compared the proposed separable color video R-D modeling method with traditional nonseparable R-D models. As the quadratic expression based models have been used in both MPEG-4 and H.264, we used frame-based MPEG-4 coder, object-based MPEG-4 coder, and H.264 coder for evaluation. We conducted exhaustive simulations using different quantization parameters for the rectangular frame level video sequences, *Mobile*, *Foreman*, *Akiyo*, *Bus*, *Flower*, and *Football*, and object level video sequences, *Akiyo*, *Foreman*, *Stefan*, *Children*, *Bream*, and *Dancer*. We encoded the first frame as intraframe and the rest frames as interframes. We used MPEG-4 reference software [44] which supports both rectangular frame based video coding and object based video coding with full search motion estimation. For object coding, the shapes were lossless coded. The H.264 reference software [41] was configured with one reference frame full search motion estimation (search range 16), CABAC entropy coding, and R-D optimized mode decision. In all the tests, the quantization parameters were predefined to remove the effects

TABLE I
ESTIMATION ERROR

	Bit rate			Distortion		
	Trad.	Prop.	Gain	Trad.	Prop.	Gain
H.264	29.2%	19.9%	31.8%	5.6%	1.9%	66.1%
MPEG-4 (Frame)	20.0%	12.0%	40.0%	3.6%	2.4%	33.3%
MPEG-4 (Object)	20.2%	12.9%	36.1%	5.2%	2.5%	51.9%

of rate control algorithms during encoding. The corresponding average relative estimation errors are defined as

$$\epsilon = \frac{1}{M}\sum_{i=1}^{M}\frac{|E_i - \tilde{E}_i|}{E_i} \tag{10}$$

where M denotes the number of the estimated frames, E_i and $\tilde{E}_i$ are the actual and estimated results for ith frame, respectively. Table I provides the average estimate results of these simulation tests. It demonstrates from the simulation results that the proposed color video R-D modeling method using separated R-D models outperforms traditional non-separable model based method which assume the same distribution function for luminance and chrominance components. It is noted if we only need to consider the distortion of the luminance component, the distortion estimation results will be the same whereas the rate estimation still needs to consider both luminance and chrominance bit rates for color video. It should be emphasized that our separable R-D modeling principle can be applicable for various expressions of R-D models, not restricted on the quadratic one.

III. R-D TRADEOFF FOR REAL-TIME VIDEO CODING

After deriving the R-D models for color video, we apply these models to our R-D tradeoff framework. Based on the feedback from the encoder buffer, the target bit rate for current encoding frame will be decided by our control-theoretic adaptation approach. The R-D tradeoff is utilized via the quality smoothing cost function regarding to the R-D models and buffer constraints. The buffer constraint is a reflection of the delay constraint as analyzed in following subsection.

A. Buffer Feedback Control

Henceforth, we assume that the encoder delay of the nth frame T_n is bounded by $0 \le T_n \le T_{\max}$. The encoder delay can be expressed according to the buffer status B_n and buffer size $B_{\rm s}$ can be decided by the maximum accumulated delay $T_{\max}$

$$T_n = \frac{B_n}{R_{\rm bitrate}} \tag{11}$$

$$B_{\rm s} = R_{\rm bitrate} \times T_{\max} \tag{12}$$

where R_n denotes the generated bits, $R_{\rm drain} = R_{\rm bitrate}/f$ is the number of bits removed from the buffer per frame, $R_{\rm bitrate}$ is the encoding bit rate and f is the frame rate. So the buffer status B_n is

$$B_n = B_{n-1} + R_n - R_{\rm drain}. \tag{13}$$

After obtaining the buffer status, we use this feedback to determine the target bit rate for incoming encoding frame. To design the feedback based controller, various combinations of proportional, integral, and derivative control actions are widely used [45], such as proportional-derivative (PD) controller for video adaptation [51] and proportional-integral-derivative (PID) controller for rate control in packet network [46] and video coding [47]. The combination of different controllers is based on the design goals. The PID controller is more flexible but more complicated as it contains more parameters of these three controllers [48]. Meanwhile, the derivative controller is sensitive to noise. We aim to provide advanced R-D tradeoff function by guaranteeing smoothing video quality and allowing larger buffer deviation. So in this paper, we consider the proportional-integral (PI) controller. We set the target encoder buffer level at half of the buffer size $B_{\rm s}$ which allows maximum bit rate fluctuation while we smooth the coded video. The error signal is defined as

$$\varepsilon_n = \frac{B_n - B_{\rm s}/2}{B_{\rm s}/2} \tag{14}$$

and the PI compensator is expressed as

$$\xi_n = K_{\rm P}\varepsilon_{n-1} + K_{\rm I}\sum_{i=1}^{n-1}\varepsilon_i \tag{15}$$

where $K_{\rm P}$ and $K_{\rm I}$ are control parameters for the proportional and integral components, respectively. In this paper, $K_{\rm P} = 0.1$ and $K_{\rm I} = 0.05$ are used empirically after exhaustive tests in H.264 real-time video coding environment.

The initial target bit rate $R'_{T,n}$ for the incoming frame n is realized by

$$R'_{T,n} = (1-\eta)\frac{1}{n-1}\sum_{i=1}^{n-1}R_i + \eta(B_{n-1} - B_{\rm target} + R_{\rm drain}) \tag{16}$$

where $\eta = 0.5$ and $B_{\rm target} = B_{\rm s}/2$ is the target buffer level. $R'_{T,n}$ is further adjusted by the PI compensator ξ_n to the target bit rate $R_{T,n}$

$$R_{T,n} = R'_{T,n} + \xi_n \times R'_{T,n}. \tag{17}$$

In traditional rate control algorithms, the target bit rate is used to solve the rate quantizer function to get the target quanitizer for encoding purpose in general. However, it ignores the impact on the picture quality varying in the time domain. In our buffer feedback control, the target bit rate calculated in the (17) is utilized as the upper bound for actual target bit rate as we employ the PI controller which tolerates larger buffer deviation in a short term.

B. Constrained R-D Tradeoff

From above analysis, we get the bit rate constraint which can be regarded as an upper bound to avoid buffer overflow. The R-D tradeoff mechanism is designed to smooth the video quality within the bit rate constraint. Meanwhile, to control the buffer deviation in a safe margin, a buffer fluctuation function is incorporated in the R-D tradeoff mechanism. This buffer fluctuation function is to dominate the cost function when the buffer deviation turns to be large. So the cost function to be minimized is

$$\mathbf{C}_n(Q) = V_n(Q) + \alpha F_{{\rm buffer},n}(Q) \quad \text{s.t.} \quad R_n(Q) \le R_{T,n} \tag{18}$$

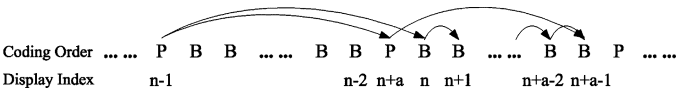


Fig. 5. Illustration on distortion variation reference selection.

where $V_n(Q)$ is the quality variation function and $F_{\text{buffer},n}(Q)$ is the buffer fluctuation function and α is a weighting factor with typical value of 0.5 to balance the two different measurements of $V(Q)$ and $F(Q)$.

Motion compensated video compression implies that the source rate and distortion of current frame depend not only on the current quantizer Q_n, where n denotes the current frame index, but also those in the reference frames, i.e., neighboring quantizers $Q_{n-u},\ldots,Q_{n+v}$ where u and v are non-negative integers. So the R-D pairs can be represented as $R_n(Q_{n-u},\ldots,Q_{n+v})$ and $D_n(Q_{n-u},\ldots,Q_{n+v})$, respectively. So the global distortion variation criterion is defined as [13]

$$V_n(Q_{n-u-1},\ldots,Q_{n+v}) = |D_n(Q_{n-u},\ldots,Q_{n+v}) - D_{n-1}(Q_{n-u-1},\ldots,Q_{n+v-1})| \quad (19)$$

where $n-u-1$ and $n+v$ satisfy the boundary conditions, N presents the number of frames in the sequence. In real-time applications, the video coding is straight forward such that the $Q_{n-u-1},\ldots,Q_{n-1}$ have been determined before encoding current frame and the $Q_{n+1},\ldots,Q_{n+v}$ are not applicable. So the distortion variation criterion can be simplified as

$$V_n(Q_n) = |D_n(Q_n) - D_{n-1}(Q_{n-1})|. \quad (20)$$

In this work, we try to smooth the video quality by minimizing the quality variations. Since in our R-D tradeoff framework, we will consider buffer deviation and integrate the buffer deviation and quality variation into a unified framework, we define $V_n(Q)$ by the normalized quality variation

$$V_n(Q) = |D_{\text{YUV},n}(Q) - D_{\text{YUV},n-1}| / D_{\text{YUV},n-1}. \quad (21)$$

It is noted that when B-frames are included, the quality variation reference selection problem is more complicated since the coding order is different to the display order. We present Fig. 5 as an example. In Fig. 5, we assume there are a B-frames between two successive P-frames. In the display order, the P-frame $n-1$ is the nearest frame to the B-frame n. So when we calculate the distortion variation $V_n(Q)$ [(21)] for encoding the B-frame n, the P-frame $n-1$ should be chosen as the distortion variation reference. For the P-frame $n+a$, the nearest coded frame in the display order is the P-frame $n-1$. For the B-frame without successive B-frames, i.e., B-frame $n+a-1$, there are two nearest frames available in display order before encoding, i.e., B-frame $n+a-2$ and P-frame $n+a$. So we propose to use the average distortion of frames $n+a-2$ and $n+a$, i.e., $(D_{\text{YUV},n+a-2}+D_{\text{YUV},n+a})/2$, as the distortion variation reference $D_{\text{YUV},n-1}$ in (21).

As far as we know, buffer control is very important since underflow and overflow should be avoided as overflow will result in frameskip and cause jerky motion and underflow will waste bandwidth. However, certain buffer fluctuation should be allowed since it is helpful to guarantee smoothed coding quality. If the current buffer deviation is not large, we can let the quality variation function dominate the cost function such that minimizing the cost function (18) is to provide minimum quality variation. When the buffer deviation is large, the buffer fluctuation function should be able to dominate the cost function such that minimizing the buffer deviation can get higher priority. In this case, quality stability is sacrificed to avoid buffer overflow or underflow. So the buffer fluctuation function, $F_{\text{buffer},n}(Q)$, is defined in an exponential expression

$$F_{\text{buffer},n}(Q) = \exp\left(\beta \frac{|R_n(Q) + B_{n-1} - R_{\text{drain}} - B_s/2|}{B_s/2}\right) \quad (22)$$

where

$$R_n(Q) = R_{\text{header},n} + R_{\text{YUV},n}(Q) \quad (23)$$

is the estimated bit rate and $R_{\text{header},n}$ is the number of bits for header information. Here we use the number of bits of previous frame as prediction. β is the weighting factor to control the tradeoff between the quality smoothness and buffer fluctuation. We set $\beta=1$ to achieve smooth video quality and allow larger buffer fluctuation. The use of larger β in the exponential function in (22) results in smaller buffer fluctuation but larger quality variation. As the bit rate is bounded by the buffer feedback control mechanism, no buffer overflow will occur. In H.264, the quantization parameter Q is allowed to vary within a certain range, i.e. ± 2, of previous frame quantizer [41]. So the quantizer of current frame is

$$Q_n^* = \arg\min_{Q_n} \mathbf{C}(Q_n) \quad \text{s.t.} \quad R_n(Q_n) \le R_{T,n} \quad (24)$$

where Q_n in the admissible quantizer candidates and Q_n^* is the optimal one. By using the R-D information corresponding to the quantization parameter candidates, Q_n, which can be obtained from the color video R-D models (6), (7), we get the optimal solution for (24).

C. Summary of the R-D Tradeoff Algorithm

The R-D tradeoff algorithm is as follows.

1) *Initialization*: Initialize the encoding parameters and model parameters.
2) *Buffer Feedback Control*: The target bit rate R_T for incoming frame is determined by the buffer feedback control as described in Section III-A.
3) *R-D Tradeoff*: The cost function (18) is calculated based on the quantization parameter candidates and the corresponding color video R-D models. The optimal solution for (24) is used as the quantization parameter for frame encoding.
4) *Encoding*: Encode frame and record the corresponding bit rate and distortion information of luminance, $R_{\text{texture},1}$, $D_{\text{texture},1}$ and chrominance components, $R_{\text{texture},2}$, $D_{\text{texture},2}$, together with the header information R_{header}.

TABLE II
SIMULATION RESULTS OF VIDEOCONFERENCING APPLICATIONS

Sequence	Format	Bit Rate (kbps)	Frame Rate (fps)	Average PSNR Variation (dB)			Maximum PSNR Variation (dB)			Average PSNR (dB)		Execute Time (Second/100 frames)	
				JM	RDT	Gain	JM	RDT	Gain	JM	RDT	JM	RDT
Mobile	QCIF	64	15	0.14	0.07	50.0%	1.35	0.68	49.6%	27.2	27.2	73.6	73.6
Foreman	QCIF	32	15	0.45	0.25	44.4%	1.90	1.10	42.1%	32.1	32.1	71.0	71.7
Akiyo	QCIF	32	15	0.29	0.14	51.7%	2.04	1.41	30.9%	42.1	41.8	71.0	71.3
Bus	CIF	256	15	0.39	0.19	51.3%	0.98	0.70	28.6%	31.3	31.3	281.3	281.3
Flower	CIF	1024	30	0.44	0.25	43.2%	1.61	1.16	28.0%	33.2	33.3	310.8	311.4
Football	CIF	1024	30	0.47	0.41	12.8%	3.52	2.43	31.0%	36.3	36.1	285.2	285.8
Tempete	CIF	256	15	0.32	0.09	71.9%	1.27	1.01	20.5%	31.7	31.7	285.0	285.4

TABLE III
SIMULATION RESULTS OF VIDEO STREAMING APPLICATIONS

Sequence	Format	Bit Rate (kbps)	Frame Rate (fps)	Average PSNR Variation (dB)			Maximum PSNR Variation (dB)			Average PSNR (dB)		Execute Time (Second/100 frames)	
				JM	RDT	Gain	JM	RDT	Gain	JM	RDT	JM	RDT
Akiyo	QCIF	64	15	0.44	0.34	22.7%	1.52	0.62	59.2%	43.3	43.7	351.7	353.3
Foreman	QCIF	64	15	0.72	0.48	33.3%	1.98	1.25	36.9%	38.4	38.2	356.7	358.7
Mobile	QCIF	64	15	1.03	0.28	72.8%	3.14	1.14	63.7%	30.8	30.4	360.0	360.3
News	QCIF	32	15	0.48	0.43	10.4%	1.61	1.26	21.7%	38.0	38.4	347.3	349.0
Bus	CIF	1024	30	0.78	0.69	11.5%	2.56	1.97	23.0%	36.3	36.2	1431.3	1434.7
Flower	CIF	1024	30	1.06	0.56	47.2%	3.86	1.81	53.1%	35.6	35.7	1446.8	1448.4
Football	CIF	1024	30	1.35	0.70	48.1%	3.85	2.18	43.4%	36.9	36.8	1442.0	1452.8
Tempete	CIF	1024	30	1.24	0.31	75.0%	3.09	1.04	66.3%	36.6	36.4	1441.5	1443.5

5) *Update the Model Parameters*: The model parameters for both the luminance and chrominance R-D models are updated based on the encoding results of the current frame as well as the past frames within the slide window.
6) *Update Buffer*: After encoding a frame, the buffer level is updated

$$B(t) = B(t-1) + R_c - R_{\text{drain}} \tag{25}$$

where R_c denotes the actual encoding bit rate. To avoid the buffer overflow, if the current buffer level $B(t)$ is higher than 80% of the buffer size B_s then the incoming frame will be skipped. After the frame is skipped, the buffer level is updated again. Then goto step 2 until the end of coding.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed R-D tradeoff approach, namely RDT, we employ the H.264 video codec and compare it with the H.264 joint model (JM) [41]. The test sequences and simulation conditions are listed in Tables II and III for the video conferencing and video streaming applications [49], respectively. All the sequences are in 4:2:0 QCIF or CIF format. For color video, the final quality measurement, PSNR_{YUV} for 4:2:0 color video sequence is described as

$$\text{PSNR}_{\text{YUV}} = 20 \log \frac{255}{D_{\text{YUV}}} \tag{26}$$

where D_{YUV} is the distortion between the original color video sequence and reconstructed one as described in (7). To measure the smoothness of the video quality, we use the following measurement [13]:

$$V_{\text{average}} = \frac{1}{N-1} \sum_{n=2}^{N} V_n \tag{27}$$

where N is the length of the coded video frames and V is the PSNR variation. We also list the maximum quality variation in Tables II and III to evaluate the robustness of these two algorithms.

We first consider the videoconferencing applications which require small delay and real-time encoding. In these tests, only the first frame is intracoded. All the other frames are coded as P-frames. The target buffer level is set as half of the buffer size where the buffer size is defined as half of the bit rate. We use the configuration with one reference frame full search motion estimation (search range 16), and R-D optimized mode decision. As shown in the simulation results in Table II, the average and maximum quality variations are improved up to 70% and 50% which demonstrate that the proposed algorithm can achieve smoother video quality. In Figs. 6 and 7, we present the PSNR curves of test sequences *Mobile* and *Foreman* for comparisons. We can observe that the the curve from the proposed approach is smoother. Other simulations also present similar performance. The buffer level curves in Figs. 6 and 7 as well as in other simulations report no buffer overflow or underflow. It is also noted that the overall quality, average PSNR, of two algorithms are very close in our tests.

As we have discussed, larger β in (22) results in smaller buffer fluctuation but larger quality variation. We present our test results of the proposed algorithm at $\beta = 4$ for the *Mobile* sequence as shown in Fig. 8. As we can observe, the buffer is more stable than the results of the JM algorithm as we have set a large β.

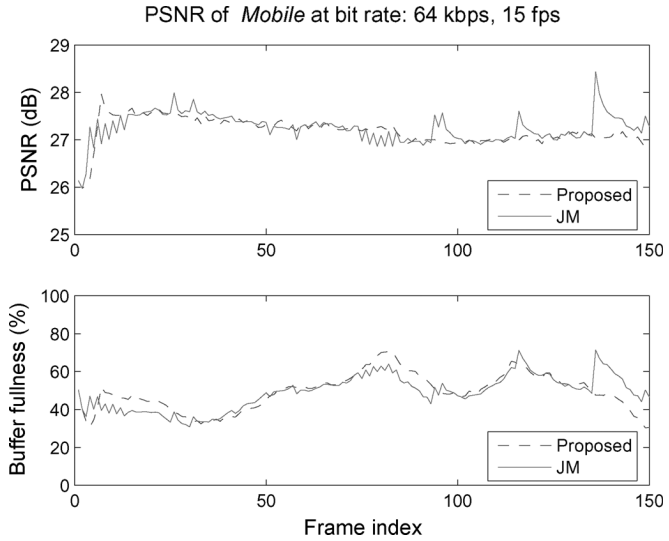


Fig. 6. PSNR curves and buffer levels of the proposed algorithm and JM algorithm for the test sequence *Mobile* (QCIF) at frame rate of 15 fps and bit rate of 64 kbps.

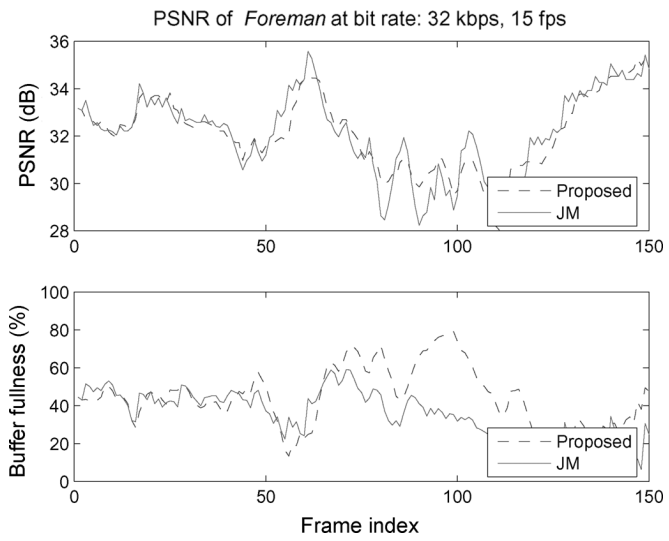


Fig. 7. PSNR curves and buffer levels of the proposed algorithm and JM algorithm for the test sequence *Foreman* (QCIF) at frame rate of 15 fps and bit rate of 32 kbps.

The corresponding PSNR of the proposed algorithm is 27.4 dB whilst the PSNR of the JM algorithm is 27.2 dB. This means that we do not sacrifice the global quality. However, the decoded video quality is not as smooth as a result of the smaller β as shown in Fig. 6.

We present some reconstructed pictures to illustrate the subjective quality improvement. Figs. 9 and 10 show the reconstructed successive frames of QCIF test sequence *Mobile* and the corresponding distortion pictures from JM and the proposed algorithms, respectively. We can tell from the comparisons, where the difference of picture quality (or the difference of the distortions) in JM algorithm is larger than the proposed algorithm. In the results from JM, the PSNR difference is 1.35 dB while the PSNR difference is 0.1 dB in our proposed

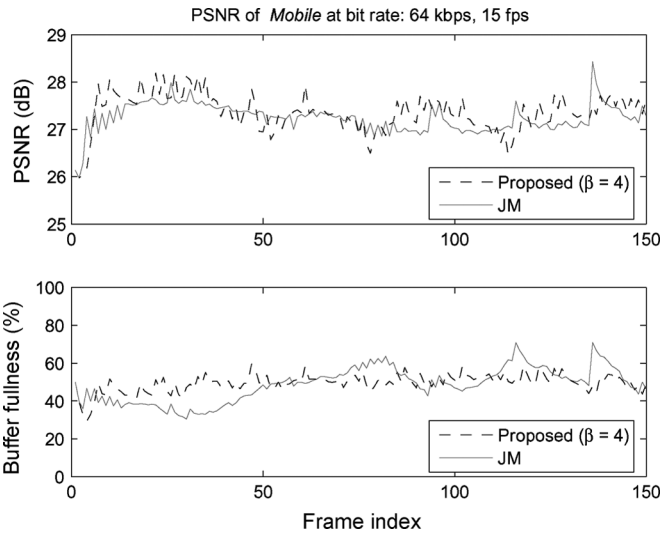


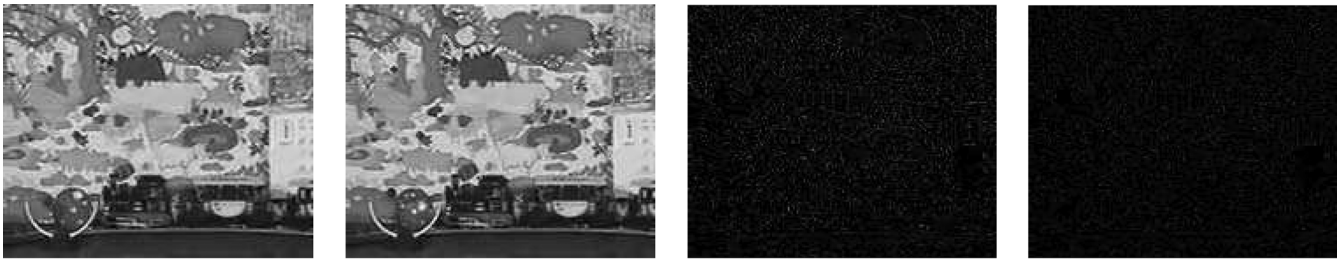
Fig. 8. PSNR curves and buffer levels of the proposed algorithm ($\beta = 4$) and JM algorithm for the test sequence *Mobile* (QCIF) at frame rate of 15 fps and bit rate of 64 kbps.

algorithm. The distortion maps also provide a subjective comparison where we can tell the difference of distortion maps from JM algorithms. Small distortion difference is expected since it results in better subjective visual quality by providing smoother video.

We also consider the video streaming applications which can enable advanced features of H.264 coder. In these tests, we modified the configuration by using five reference frames, two B-pictures between each two successive P-pictures. The simulation results are listed in Table III. Both the average and maximum quality variation achieve improvements. This shows that the RDT algorithm can achieve smoother reconstructed video quality. Similar to our first set of tests, the average PSNR of the decoded results are close.

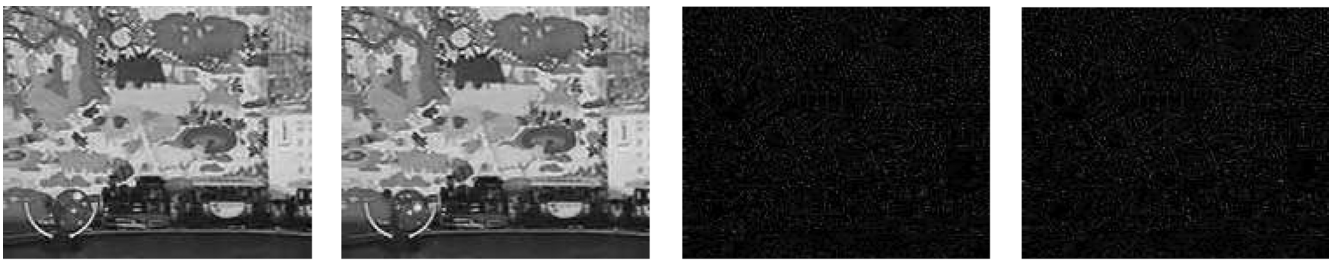
To compare with the state-of-the-art work in [50] which also reduces the quality fluctuation for H.264 coding, we implemented the RDT approach on H.264 JM8.4 reference encoder software. In the simulation tests, we used IPPP... GOP structure with the GOP size of 12. The first 120 frames of each sequence were used. We configured the encoder with two reference frames in full search motion estimation with the search range of 16, 1/4-pel motion vector resolution, and CABAC for entropy coding. The results are listed in Table IV. To evaluate the performance, we only use the PSNR values of the luminance (Y) components. As we can observe from the results, the RDT approach outperforms the frame bit allocation algorithm of [50] in most cases in terms of reducing quality fluctuation. Meanwhile, the average PSNR performances are comparable. Although the RDT approach suffered from larger buffer fluctuation, no buffer overflow or underflow occurred.

Since we need to model the luminance and chrominance components separately, the R-D modeling complexity is doubled. However, as we have shown in the Tables II and III, the execution times for these JM- and RDT-based encoders are very close since the computational complexity of the rate control part is insignificant compared to that of the whole video coding system.



(a) (b) (c) (d)

Fig. 9. Reconstructed successive frames 135 and 136 of QCIF test sequence *Mobile* using JM algorithm, PSNR difference: 1.35 dB. (a) Reconstructed frame 135 (PSNR: 27.06 dB). (b) Reconstructed frame 136 (PSNR: 28.41 dB). (c) Distortion of reconstructed frame 135. (d) Distortion of reconstructed frame 136.



(a) (b) (c) (d)

Fig. 10. Reconstructed successive frames 135 and 136 of QCIF test sequence *Mobile* using the proposed algorithm, PSNR difference: 0.10 dB. (a) Reconstructed frame 135 (PSNR: 27.10 dB). (b) Reconstructed frame 136 (PSNR: 27.00 dB). (c) Distortion of reconstructed frame 135. (d) Distortion of reconstructed frame 136.

TABLE IV
SIMULATION RESULTS COMPARED WITH [50]

Sequence	Format	Bit Rate (kbps)	Average PSNR Variation (dB)		Average PSNR (dB)	
			[50]	RDT	[50]	RDT
Akiyo	QCIF	60	0.76	0.37	38.4	38.1
News	QCIF	100	0.86	0.46	35.9	36.2
Foreman	QCIF	100	0.72	0.49	33.9	34.3
Mobile	CIF	1500	0.43	0.58	32.6	32.7
Tempete	CIF	1500	0.68	0.58	35.8	35.7

It is noted the RDT algorithm is also more complex than the algorithm presented in [50] which is faster than the JM algorithm. However, this price is reasonable in exchange for the improved performance and neglectable additional complexity in the encoder.

V. CONCLUSION

In this paper, we investigated the color video R-D modeling problem and proposed the separable R-D models. Based on a control-theoretic adaptation approach, we further proposed a novel R-D tradeoff framework to smooth the video quality which shows a good performance as demonstrated in our simulation results. The overall quality is guaranteed and the the bandwidth is efficiently utilized without buffer overflow or underflow. From the results, we assert that a better subjective visual quality is achieved by the proposed algorithm since a smoother visual quality is more pleasing to look at than a fluctuating one.

REFERENCES

[1] *Video Codec for Audiovisual Services at $p \times 64$ kbit/s*, ITU-T Rec. H.261 ver. 2, 1990.

[2] *Video Coding for Low Bitrate Communication*, ITU-T Rec. H.263 ver. 2, 1998.

[3] *Coding of Moving Pictures and Associated Audio Fordigital Storage Media at Up to About 1.5 Mbps*, ISO/IEC 11172-2:1991, 1991.

[4] *Information Technology—Generic Coding of Moving Pictures and Associated Audio*, ISO/IEC 13818-2:1994, 1994.

[5] *Information Technology—Coding of Audio/Visual Objects*, ISO/IEC 14496-2:1999, 1999.

[6] *Draft ITU-T Rec. Final Draft Int. Standard of Joint Video Specification (ITU-T Rec. H.264 |ISO/IEC 14496-10 AVC)*, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, 2003.

[7] T. V. Lakshman, A. Ortega, and A. R. Reibman, "VBR video: Tradeoffs and potentials," *Proc. IEEE*, vol. 86, no. 5, pp. 952–973, May 1998.

[8] A. Ortega, "Variable bit-rate video coding," in *Compressed Video Over Networks*, M.-T. Sun and A. R. Reibman, Eds. New York: Marcel Dekker, 2000, pp. 343–382.

[9] N. Mohsenian, R. Rajagopalan, and C. A. Gonzales, "Single-pass constant- and viriable-bit-rate MPEG-2 video compression," *IBM J. Res. Develop.*, vol. 43, pp. 489–509, Jul. 1999.

[10] P. H. Westerink, R. Rajagopalan, and C. A. Gonzales, "Two-pass MPEG-2 variable-bit-rate encoding," *IBM J. Res. Develop.*, vol. 43, pp. 471–488, Jul. 1999.

[11] Y. Yu, J. Zhou, Y. Wang, and C. W. Chen, "A novel two-pass VBR coding algorithm for fixed-size storage application," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 345–356, Mar. 2001.

[12] G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression*. Norwell, MA: Kluwer, 1997.

[13] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, pp. 446–459, Aug. 1998.

[14] M. F. Ringenburg, R. E. Ladner, and E. A. Riskin, "Global MINMAX interframe bit allocation for embedded video coding," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2004, pp. 222–231.

[15] G. Shavit, R. E. Ladner, and E. A. Riskin, "MINMAX bit allocation for quantization-based video coders," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2005, pp. 299–308.

[16] Y. Sermadevi and S. S. Hemami, "Efficient bit allocation for dependent video coding," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2003, pp. 232–241.

[17] ——, "MINMAX frame rate control using a rate-distortion optimized wavelet coder," in *Proc. Int. Conf. Image Process.*, Kobe, Japan, Oct. 1999, pp. 551–555.

[18] Y. Yang and S. S. Hemami, "Rate control for VBR video over ATM: simplification and implementation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 9, pp. 1045–1058, Sep. 2001.
[19] A. Ortega and K. Ramchandran, "Rate-distrotion methods for image and video compression," *IEEE Signal Process. Mag.*, pp. 23–50, Nov. 1998.
[20] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 1, pp. 12–20, Feb. 1996.
[21] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
[22] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172–185, Feb. 1999.
[23] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 186–199, Feb. 1999.
[24] H. Song and C.-C. J. Kuo, "Rate control for low-bit-rate video via variable-encoding frame rates," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 4, pp. 512–521, Apr. 2001.
[25] Z. He, Y. K. Kim, and S. K. Mitra, "Low-delay rate control for DCT video coding via ρ domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.
[26] K. N. Ngan, T. Meier, and Z. Chen, "Improved single video object rate control for MPEG-4," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 5, pp. 385–393, May 2003.
[27] Z. Chen and K. N. Ngan, "Rate-constrained arbitrarily shaped video object coding with object-based rate control," *Vis., Image Signal Process.*, vol. 151, pp. 250–256, Aug. 2004.
[28] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
[29] H.-M. Hang and J. J. Chen, "Source model for transform video coder and its application—Part I: Fundamental theory," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 287–298, Apr. 1997.
[30] A. J. Viterbi and J. K. Omura, *Principle of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
[31] B. Tao, B. W. Dickinson, and H. A. Peterson, "Adaptive model-driven bit allocation for MPEG video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 1, pp. 147–157, Feb.. 2000.
[32] H.-J. Lee, T. Chiang, and Y.-Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 878–894, Sep. 2000.
[33] J.-B. Cheng and H.-M. Hang, "Adaptive piecewise linear bits estimation model for MPEG based video coding," *J. Vis. Commun. Image Represent.*, vol. 8, pp. 51–67, Mar. 1997.
[34] Z. He, W. Zeng, and C. W. Chen, "Low-pass filtering of rate-distortion functions for quality smoothing in real-time video communication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 973–981, Aug. 2005.
[35] B. Xie and W. Zeng, "A sequence-based rate control framework for consistent quality real-time video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 1, pp. 56–71, Jan. 2006.
[36] H. S. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 598–603, Jul. 2003.
[37] F. Bossen, ABT cleanup and complexity reduction, JVT-E087 Oct. 2002.
[38] R. Reininger and J. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM–31, no. 6, pp. 835–839, Jun. 1983.
[39] S. R. Smoot and L. A. Rowe, "Study of DCT coefficient distributionss," in *Proc. Human Vision and Electronic Imaging*, San Jose, CA, Jan. 1996, pp. 403–411.
[40] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Process.*, vol. 9, pp. 1661–1666, Oct. 2000.
[41] *Joint Model 9.8 (JM9.8)*. [Online]. Available: http://bs.hhi.de/suehring/tml/download/jm98.zip
[42] Z. Chen and K. N. Ngan, "Joint texture-shape optimization for MPEG-4 multiple video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 9, pp. 1170–1174, Sep. 2005.
[43] N. Brady, "MPEG-4 standardized methods for the compression of arbitrarily shaped video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 8, pp. 1170–1189, Dec. 1999.
[44] *MPEG-4 Video Verification Model v18.0*, ISO/IEC JTC1/SC29/WG11, Jan. 2001.
[45] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control Systems*. Reading, MA: Addison-Wesley, 1990.
[46] F. Blanchini, R. L. Cigno, and R. Tempo, "Robust rate control for integrated services packet networks," *IEEE/ACM Trans. Netw.*, vol. 10, pp. 644–652, Oct. 2002.
[47] Y. Sun and I. Ahmad, "A robust and adaptive rate control algorithm for object-based video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 10, pp. 1167–1182, Oct. 2004.
[48] J. L. Hellerstein, Y. Diao, S. Parekh, and D. M. Tibury, *Feedback Control of Computing Systems*. Hoboken, NJ: Wiley, 2004.
[49] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.
[50] N. Kamaci, Y. Altunbasak, and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 994–1006, Aug. 2005.
[51] S. C. Liew and D. C.-Y. Tse, "A control-theoretic approach to adapting VBR compressed video for transport over a CBR communications channel," *IEEE/ACM Trans. Netw.*, vol. 6, no. 1, pp. 42–55, Feb 1998.

Zhenzhong Chen (S'02) received the B.E. degree from Huazhong University of Science and Technology, Wuhan, China. He is currently pursuing the Ph.D. degree at the Department of Electronic Engineering, the Chinese University of Hong Kong, Hong Kong.

From 2001 to 2003, he conducted graduate research at Nanyang Technological University, Singapore. He was a recipient of Microsoft fellowship. His current research interests are visual signal processing and communications.

Mr. Chen is a student member of SPIE and IEE, U.K.

King Ngi Ngan (M'79–SM'01–F'00) received the Ph.D. degree in electrical engineering from Loughborough University of Technology, Loughborough, U.K.

He is currently a Chair Professor in the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, and was previously a Full Professor at the Nanyang Technological University, Nanyang, Singapore, and at the University of Western Australia, Australia. He has chaired a number of prestigious international conferences on video signal processing and communications and served on the advisory and technical committees of numerous professional organizations. He has published extensively including three authored books, five edited volumes, and over 200 refereed technical papers in the areas of image/video coding and communications.

Prof. Ngan is an Associate Editor of the *Journal on Visual Communications and Image Representation*, as well as an Area Editor of EURASIP *Journal of Signal Processing: Image Communication*, and served as an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and *Journal of Applied Signal Processing*. He is a Fellow of IEE (U.K.), and of the IEAust (Australia).