Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding

Zhihai He, *Member, IEEE*, Jianfei Cai, *Student Member, IEEE*, and Chang Wen Chen, *Senior Member, IEEE*

***Abstract*—In this paper, we first develop a rate-distortion (R-D) model for DCT-based video coding incorporating the macroblock (MB) intra refreshing rate. For any given bit rate and intra refreshing rate, this model is capable of estimating the corresponding coding distortion even before a video frame is coded. We then present a theoretical analysis of the picture distortion caused by channel errors and the subsequent inter-frame propagation. Based on this analysis, we develop a statistical model to estimate such channel errors induced distortion for different channel conditions and encoder settings. The proposed analytic model mathematically describes the complex behavior of channel errors in a video coding and transmission system. Unlike other experimental approaches for distortion estimation reported in the literature, this analytic model has very low computational complexity and implementation cost, which are highly desirable in wireless video applications. Simulation results show that this model is able to accurately estimate the channel errors induced distortion with a minimum delay in processing. Based on the proposed source coding R-D model and the analytic channel-distortion estimation, we derive an analytic solution for adaptive intra mode selection and joint source-channel rate control under time-varying wireless channel conditions. Extensive experimental results demonstrate that this scheme significantly improves the end-to-end video quality in wireless video coding and transmission.**

***Index Terms*—End-to-end distortion, error propagation, joint source-channel coding, rate-distortion analysis, wireless video.**

I. Introduction

WITH the increasing bandwidth in the next-generation mobile network and rapidly growing demand for visual communication, wireless video transmission has become possible and received much attention during the last few years. Due to the limited bandwidth of the wireless channels, video signals have to be highly compressed by efficient coding algorithms, such as H.263 [1] and MPEG-4 [2]. On the other hand, under the error-prone wireless environments, highly compressed video data becomes extremely vulnerable [3], [4]. A single bit error may cause severe degradation in video quality. Therefore, it is necessary for the video encoder to provide adequate error resilience features to protect the video data from the channel errors. Two effective approaches for error resilience and protection are: 1) error control and 2) intra update of macroblocks (MBs) [3]–[6]. Error control, such as forward error correction (FEC) and automatic repeat request (ARQ), are often used to correct the bit errors in the compressed video data by adding controlled redundancy information. Due to stringent delay constraint for real-time video transmission, it is often considered more beneficial to use FEC than to apply ARQ. The second approach of MB intra update, also called intra refreshing, is a fairly efficient way to stop error propagation, because the decoding of an intra MB does not need the information from its previous frames which may have already been "corrupted" by channel errors. In contrast, for an intercoded MB, even if its bit stream has been correctly received and decoded, the channel errors introduced in previous frames may still propagate to the current frame along the motion-compensation path [4], [5], [7].

A. Problem Formulation

Increased error resilience often comes at the cost of higher bandwidth consumption. For example, intra coding of a MB or a frame often requires much more bits than inter coding. This is because motion compensation in the inter coding mode can largely remove the temporal redundancy between two neighboring video frames. However, the inter coding of MBs, although having much better R-D performance than the intra mode, enables channel error propagation along the motion-compensation path, which has significant impact on the video quality. Therefore, a tradeoff needs to be made when selecting the MB coding mode. Let β be the intra refreshing rate, the percentage of MBs coded with intra mode. The tradeoff problem can then be formalized as follows: given the transmission channel conditions, such as bandwidth and bit error rate (BER) p_e, how to determine the optimal β such that the overall picture quality at the receiver end is maximized.

In video coding and transmission over noisy channels, Reed–Solomon (RS) code is one of the widely used FEC schemes [8]. An (N, K) RS code with length N and dimension K encodes K information bits with N bits. Clearly, the $(N-K)$ bits are the FEC overhead that consume a portion of the total bandwidth. However, with this overhead redundancy, the RS decoder can correct certain amount of channel errors, which in turn would significantly reduce the picture distortion induced by channel errors. It is true that if we assign more bits to the RS code, the RS decoder can correct more bit errors. However, because of the limited overall channel bit rate, fewer bits will be assigned to the source encoder which results in

Manuscript received August 15, 2001; revised April 15, 2002.

Z. He and C. W. Chen are with the Interactive Media Group, Sarnoff Corporation, Princeton, NJ 08543 USA (e-mail: zhe@sarnoff.com; cchen@sarnoff.com).

J. Cai is with the Interactive Media Group, Sarnoff Corporation, Princeton, NJ 08543 USA and also with the Department of Electrical Engineering, University of Missouri–Columbia, Columbia, MO 65201 USA (e-mail: cai@ee.missouri.edu).

Publisher Item Identifier 10.1109/TCSVT.2002.800313.

1051-8215/02$17.00 © 2002 IEEE

increased source coding distortion. Therefore, in a FEC-based video coding and transmission system, one of the key problems is the joint source-channel bit allocation and rate control. To be more specific, given the transmission channel conditions, such as bandwidth and BER, the encoder needs to find the optimal bit allocation between source coding and channel coding such that the overall picture distortion at the receiving end is minimized. This is the second tradeoff problem to be examined by this work.

Let R_s and R_c be the source and channel coding bit rates. Let D be the overall picture distortion at the receiver end, defined as the mean square error (MSE) between the decoded video frame and the original one. In an end-to-end wireless video coding and transmission system, there are two major types of picture distortion. The first one is the quantization distortion introduced in source encoding. The second type of distortion is caused by channel errors. For the convenience of our presentation, in this work, we call these two types of distortion as "source distortion" and "channel distortion," denoted by D_s and D_c, respectively. More precisely, source distortion refers to the MSE between the reconstructed video frame at the encoder (used for motion estimation and compensation of the next frame) and the original one. Channel distortion refers to the MSE between the decoded video frame at the receiver and the reconstructed frame at the encoder. Note that if no bit errors occur, the two reconstruction frames at the encoder and decoder should be exactly the same. Note that, D_s is a function of R_s and β, denoted by $D_s(R_s, \beta)$, and D_c is a function of code rate $r = R_s/R$ and p_e, denoted by $D_c(r, p_e)$. Obviously, in order to perform optimal bit allocation between source and channel coding and to select the optimal intra refreshing rate β, we need first to analyze the R-D behaviors of video encoder and decoder and estimate the functions $D_s(R_s, \beta)$ and $D_c(r, p_e)$.

B. Analysis of Source-Coding Distortion

For a given source-coding bit rate R_s, to estimate the corresponding source distortion D_s, we need to model and analyze the R-D behavior of the video encoder. Due to the varying characteristics of the input videos and the sophisticated data representation scheme employed by the coding algorithm, accurate analytic estimation of the R-D behavior of the video encoder remains a challenging problem [12]. Because of this, operative R-D estimation is often adopted, in which the R-D functions are assumed to follow some mathematical model. The coding algorithm is then run over the input video several times to generate several R-D measurements, which are used to estimate the model parameters [11], [12]. To reduce the computational complexity, MPEG-4 TM7 [17] and H.263 TMN8 [18] rate control algorithms use the coding statistics of previous frames or MBs to estimate the model parameters for the current frame or MB.

With adaptive intra refreshing, it is even more difficult to analyze the R-D behavior of the video encoder. This is because the input data to the quantizer and the entropy encoder also changes as different coding modes are applied to the MBs. Very little research has been done in the literature to investigate the impact of the intra refreshing rate on the R-D performance of the video encoder. In our previous work [9], [10], we have developed accurate and robust R-D model for DCT-based video coding by introducing the so-called "ρ-domain" R-D analysis methodology. In this work, we extend this model to incorporate the intra refreshing rate β. Our experimental results show that the extended R-D model can accurately estimate the source coding distortion function $D_s(R_s, \beta)$.

C. Analysis of Channel Distortion

Standard video coding schemes, such as H.263 and MPEG-4, employ a motion-compensation based discrete cosine transform (MC-DCT) coding scheme. As indicated earlier, while motion compensation significantly improves the coding efficiency, it also causes inter-frame propagation of channel errors, and significantly degrades the picture quality at the receiver end. For this reason, the complex error propagation in the video decoding loop has to be accurately modeled in channel-distortion analysis. Obviously, the modeling process needs to consider the specific source/channel encoding and decoding schemes, packetization method, patterns of the channel errors, error concealment, and so on. Several approaches for channel-distortion estimation have been proposed in the literature [3], [4]. To analyze the video transmission over lossy channels, a heuristic approach is introduced in [3], where the channel-distortion formula is derived through a leaking filter model. This distortion formula has several control parameters. To estimate these parameters, one needs to run the codec over the video a few times to generate some measurement points and match the model to the experimental data. Obviously, this type of estimation scheme is not desired in real time video coding and communication. A statistical simulation of the video decoder is employed in [4] to estimate the channel distortion with error concealment at the decoder. Using this decoder simulation, the encoder understands how much the picture at decoder is "corrupted" by the random channel errors. Such estimation scheme involves potentially high computational complexity and implementation cost. In addition, this type of simulation approach does not allow further analysis for global optimization.

In this paper, based on the statistical analysis of the error propagation, error concealment, and channel decoding, we develop a theoretical framework to estimate the channel distortion. Our extensive experimental results demonstrate that the proposed statistical model can estimate the channel distortion very accurately and robustly. Coupled with the R-D model for source coding, an adaptive mode selection and rate control algorithm is proposed for wireless video coding and transmission. This end-to-end R-D analysis framework can be applied to any standard video coding system and any video sequence. Our simulations show that the optimal mode selection and joint source-channel rate control can achieve up to a 2–3-dB PSNR gain in picture quality, comparing with the other methods reported in the literature [20], [21].

The rest of the paper is organized as follows. Section II presents the extension of the source coding R-D model and rate control algorithm developed in [9], [10] by incorporating the MB intra refreshing rate. The extended R-D model is able to estimate the source distortion function $D_s(R_s, \beta)$. The channel-distortion model and the corresponding estimation scheme are described in Section III. Theoretical analysis of the asymptotic behavior of channel distortion is also given in

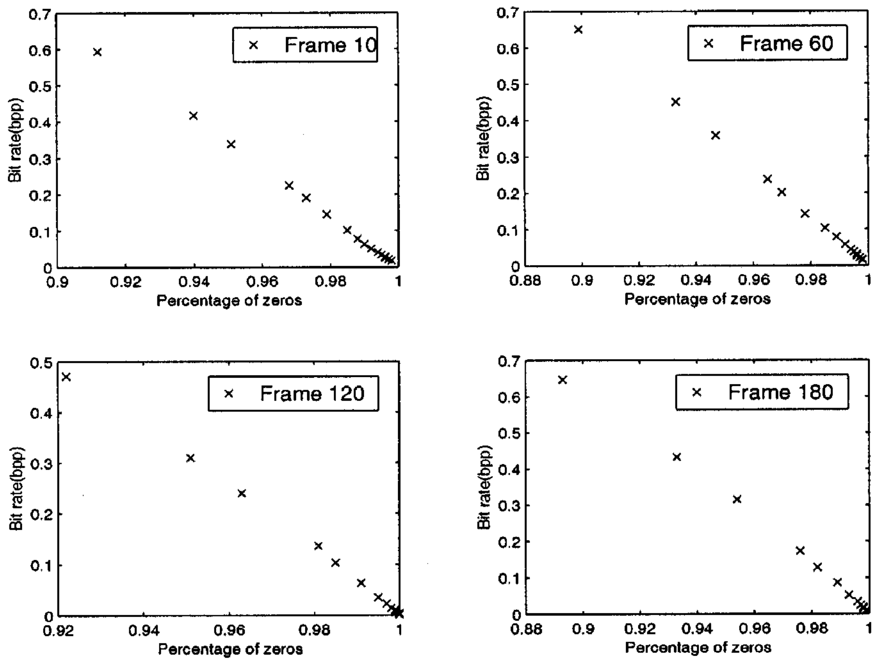


Fig. 1. Linear relationship between the source coding bit rate R_s and the percentage of zeros ρ. The test frames are from the "Foreman" QCIF video.

Section III. In Section IV, based on the source and channel R-D models, we propose an adaptive mode selection scheme and show the corresponding simulation results. Section V presents the optimal joint source-channel bit allocation and rate control algorithm and the corresponding simulation results. Concluding remarks are given in Section VI.

II. R-D Analysis for Source Coding

In this section, we first review the ρ-domain R-D model developed in [9], [10]. By incorporating the MB intra refreshing rate β, we extend this R-D model for DCT video coding with adaptive mode selection.

A. ρ-Domain R-D Model

In the previous work [9], [10], a robust and accurate R-D model is developed for DCT-based video coding. Specifically, in this model, we consider the source coding bit rate R_s and distortion D_s as functions of ρ, which is the percentage of zeros among the quantized DCT coefficients. This consideration is based on the following observation. In the classical R-D analysis [13]–[15], R_s and D_s are treated as functions of the quantization parameter (or step size) q. Notice that in standard video coding, such as H.263 and MPEG-4, ρ monotonically increases with q. This implies that there is a one-to-one mapping between them. Therefore, mathematically, R_s and D_s are also functions of ρ. A study of R_s and D_s as functions of ρ is termed *ρ-domain analysis*. We observe that, in the ρ domain, the R-D functions $R_s(\rho)$ and $D_s(\rho)$ have unique behaviors. Specifically, R_s has a linear relationship with ρ; i.e.,

$$R_s(\rho) = \theta \cdot (1-\rho) + C_h \tag{1}$$

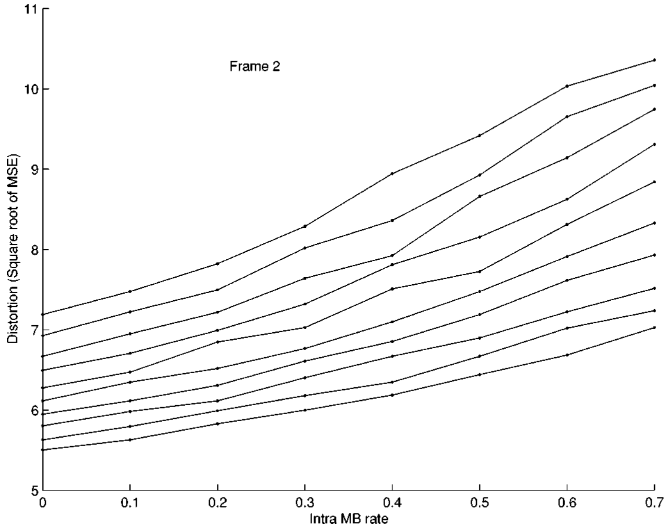
where θ is a constant and C_h refers to the number of bits for header information and motion vectors. Note that C_h does not depend on the quantization. To better understand this linear rate model and for the integrity of our presentation, we reproduce the simulation results in [10]. With the MPEG-4 video codec [16], we encode the test video sequence at a series of quantization step sizes. In Fig. 1, we plot $R_s(\rho)$ for several frames from the "Foreman" QCIF video sequence. It can be seen that there is a clear linear relationship between R_s and ρ. We have performed this test over a wide range of video sequences and with different coding algorithms, this linear rate model has been found to hold [9], [10]. Within the ρ-domain, we also have developed the following distortion model for source coding:

$$D_s(\rho) = \sigma^2 e^{-\alpha(1-\rho)} \tag{2}$$

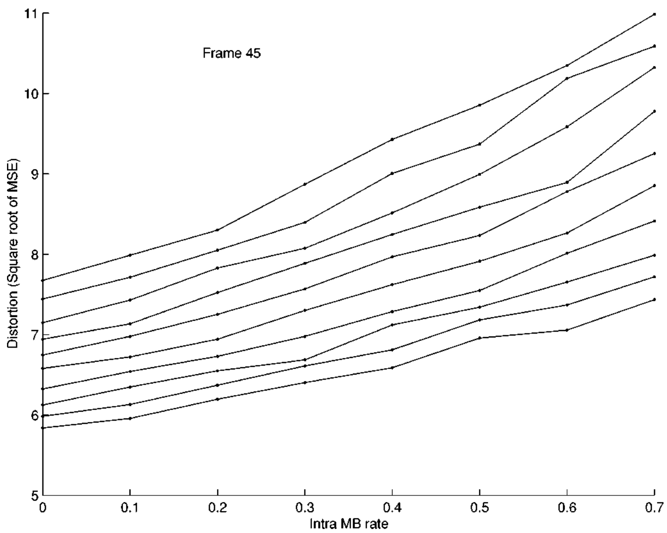
where σ^2 is the variance of the source data and α is a constant. The extensive experimental results in [9] have shown that the above R-D model is very accurate. Based on the rate model (1), a linear rate control algorithm has also been developed, with which we can control the video encoder to achieve the target bit rate accurately and robustly. A detailed treatment of the above R-D model and rate control algorithm can be found in [9], [10].

B. R-D Functions With MB Intra Refreshing

For a given input picture, using the R-D model presented in Section II-A, we can estimate its R-D function before quantization and coding. In video coding with adaptive intra mode selection, if we change the coding mode of each MB or the intra refreshing rate β of the video frame, the input to the video coding algorithm is also changed. For example, if a MB is intra coded, the input video data is just a MB in the original video frame. If it is inter coded, the input video data is then the motion-compensation difference. Therefore, in the R-D analysis for video coding with adaptive mode selection, we also need to consider the impact of β on the R-D behavior of the video encoder. In general, as β increases and more MBs are forced to be intra coded, the average coding bit rate becomes higher. In



(a)



(b)

Fig. 2. Source coding distortion as a function of β for (a) Frame 2 and (b) Frame 45 of "Carphone" QCIF video at different bit rates. In each plot, different curves correspond to different coding bit rates R_s.

other words, for a given coding bit rate R_s, the source distortion D_s also increases with β. In Fig. 2, we plot the $D_s(R_s, \beta)$ for Frames 2 and 45 of "Carphone" QCIF video at different values of R_s. Our extensive simulations over various additional video sequences show similar behavior of $D_s(R_s, \beta)$. Two extreme cases here are $\beta = 0$ and 1 when all the MBs are inter and intra coded, respectively. The corresponding distortion values are denoted by $D_s(R_s, 1)$ and $D_s(R_s, 0)$. From Fig. 2, we can see that as β increases from 0 to 1, $D_s(R_s, \beta)$ increases from $D_s(R_s, 0)$ to $D_s(R_s, 1)$. During our simulations, we find out that the following quadratic approximation is sufficiently accurate:

$$D_s(R_s, \beta) = D_s(R_s, 0) + \beta(1 - \lambda + \lambda\beta) \cdot [D_s(R_s, 1) - D_s(R_s, 0)] \quad (3)$$

where λ is a constant which depends on the specific characteristics of the video sequence. Therefore, to estimate $D_s(R_s, \beta)$, we only need to estimate the R-D functions of the current frame at two extreme modes: all intra mode and all inter mode. In the case of all intra mode, the input data is exactly the original video frame. While the case of all inter mode and no MB is forced to be intra coded, the input data is the original motion-compensation difference picture. Note that both types of input data are available at the encoder before R-D modeling. Therefore, for each mode, the R-D estimation can be carried out with (1) and (2). The distortion function for adaptive intra refreshing is then obtained by (3). It should be mentioned that the model parameters θ and α are determined from the coding statistics of previous frame, as explained in [9].

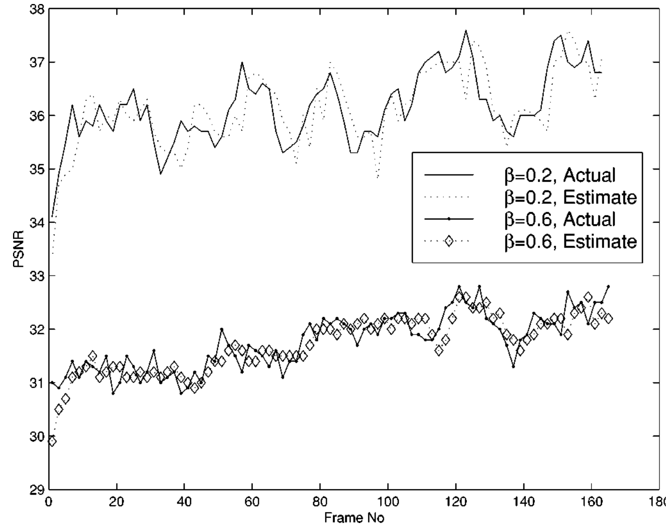


Fig. 3. Source coding distortion estimation results for "Carphone" QCIF video coded with MPEG-4 at 96 kbps.

C. Experimental Results

To test the accuracy of the proposed R-D model for source coding with adaptive intra refreshing, we run the MPEG-4 codec [16] on "Carphone" and "Flowergarden" QCIF videos at different bit rates and different values of β, and estimate the distortion function $D_s(R_s, \beta)$ before quantization and coding. In Fig. 3, we plot the actual distortion and the estimation for "Carphone" coded at 96 kbps with $\beta = 0.2$ and 0.6. The estimation result for "Flowergarden" at 256 kbps is shown in Fig. 4. In this case, the test values of β are 0.2 and 0.8. It can be seen that the proposed R-D model gives a very accurate and robust estimation of the source coding distortion. Our tests over other video sequences and encoder settings yield similar results.

III. Analysis of Channel Distortion

In wireless video coding and transmission, channel coding such as RS code is often used to correct bit errors in the coded video data stream. Due to the limited error correction capacity of the channel decoder, residual bit errors often still exist after error correction. When a corrupted codeword in the bit stream cannot be properly decoded, the video encoder will jump to the next packet starting with a resynchronization mark and skip all the intermediate bits. This introduces visible picture distortion at the receiver end. Note that at the decoder the current reconstruction frame "corrupted" by bit errors will still be used as the motion-compensation reference for the next frame. In this way,

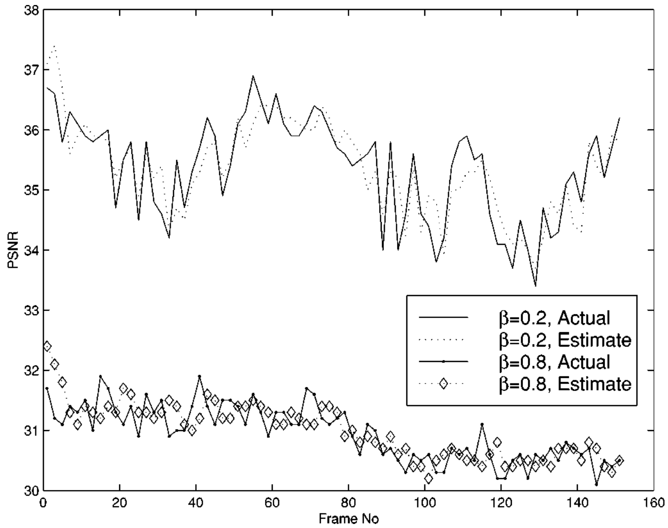


Fig. 4. Source-coding distortion estimation results for "Flowergarden" QCIF video coded with MPEG-4 at 256 kbps.

the channel distortion propagates along the motion-compensation path, which often severely degrades the video presentation quality. Therefore, the channel-distortion model needs to consider the complex error propagation in video decoding [3]–[5]. Furthermore, the channel-distortion model also needs to consider the varying characteristics of the input video data, specific channel conditions such as channel bandwidth and BER, complex data representation and coding scheme employed by the video encoder, sophisticated error resilience and concealment methods, as well as the operating mechanism of the video decoder. Most important of all, the channel-distortion model has to deal with the random nature of the bit errors. Therefore, accurate and robust modeling of the channel distortion remains a challenging problem.

In an end-to-end video coding and communication system, the channel-distortion analysis and optimization can operate at different levels. The decoder simulation approach in [4] estimates the average channel distortion at the pixel level, and performs bit allocation and control at the MB level. In [6], an empirical error-sensitivity metric is used to select the coding mode for each MB. Despite its simplicity, the maximization of the overall picture quality is not guaranteed. The channel-distortion analysis in [3] operates at the video sequence level. In general, channel-distortion analysis at low levels, such as the pixel and MB levels, is able to capture the local behavior of bit errors, especially when the channel BER is very small and very few MBs are corrupted. Special treatment of these MBs may be beneficial [6]. It should be noted that this approach often needs an immediate transmission feedback from the decoder. Otherwise, after a relatively large feedback delay, the random bit errors have already propagated to many other MBs and may spread over the whole video frame. In this way, the channel distortion exhibits a frame-level statistical behavior. In this case, there is no need to perform the MB-specific analysis and optimization, which often involves potentially high computational complexity. In addition, accurate modeling and estimation of the R-D behavior of one MB is often very difficult. This is because the R-D analysis of source coding is a statistical process that needs sufficiently large amount of data to achieve reasonable estimation accuracy. Using an inaccurate R-D model for source coding, the MB-level optimization actually can not achieve truly optimized and robust picture quality. In this work, we try to develop a statistical model to describe the overall behavior of the channel distortion. This analytic statistical model allows global optimization of picture quality through joint source-channel coding and adaptive selection of error-resilience parameters, such as intra refreshing rate, synchronization frequency, etc. [5]. Such analysis also provides very useful information for resource allocation and Quality of Service (QoS) control in network transmission [19].

A. End-to-End Distortion

We denote the packet-loss ratio as p. If we assume each packet contains the same number of MBs (or pixels), then the loss ratio of a pixel is also p [4]. Let $F(n, i)$ be the original value of pixel i in the nth video frame, and $\hat{F}(n, i)$ be the corresponding reconstruction value in the feedback loop at the encoder. We denote the reconstruction value at the receiver end as $\tilde{F}(n, i)$. For inter coded MBs, let $e(n, i)$ be the motion-compensation difference at the encoder. Let $\hat{e}(n, i)$ and $\tilde{e}(n, i)$ be the corresponding reconstruction values at the encoder and decoder, respectively. Due to the randomness of bit errors, $\tilde{F}(n, i)$ and $\tilde{e}(n, i)$ are actually random variables. Therefore, we can only model and analyze the expected picture distortion at the receiver end which is given by

$$D(n) = E\{[F(n, i) - \tilde{F}(n, i)]^2\}. \tag{4}$$

It should be noted that $E\{x(n, i)\}$ here represents the average (over all pixels) expected value of the random variable $x(n, i)$. According to their definitions, the source-coding distortion $D_s(n, i)$ and channel distortion $D_c(n)$ are given by

$$D_s(n) = E\{[F(n, i) - \hat{F}(n, i)]^2\}, \tag{5}$$

$$D_c(n) = E\{[\hat{F}(n, i) - \tilde{F}(n, i)]^2\} \tag{6}$$

respectively. In this paper, we assume the $D_s(n)$ and $D_c(n)$ are uncorrelated with each other. That is to say

$$D(n) = D_s(n) + D_c(n). \tag{7}$$

To justify this assumption, we code the "Foreman" QCIF video at 96 kbps and 15 fps with MPEG-4 and simulate the transmission with random packet loss at a loss ratio of 2%. In Fig. 5, we plot the $D(n)$ and $D_s(n) + D_c(n)$ for each frame. It can be seen that $D(n)$ is approximately equal to $D_s(n) + D_c(n)$. Similar test has been performed over other video sequences and encoding settings. The average relative difference e_D between the $D(n)$ and $D_s(n) + D_c(n)$, defined by

$$e_D = \frac{1}{T} \sum_{n=1}^{T} \frac{|[D_s(n) + D_c(n)] - D(n)|}{D(n)} \times 100\% \tag{8}$$

for each test is listed in Table I. Here, T is the total number of video frames. It can be seen that e_D is very small. This implies that it is quite reasonable to assume that source distortion

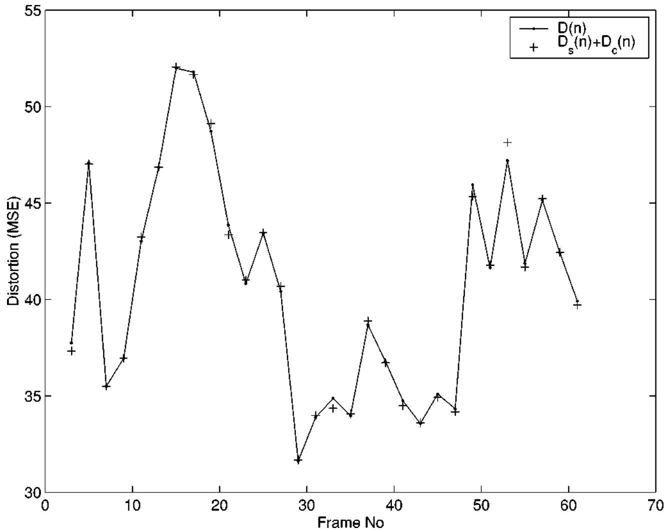


Fig. 5. Comparison between the overall distortion $D(n)$ and $D_s(n)+D_c(n)$ for "Foreman" QCIF video coded at 96 kbps and 15 fps and a loss rate of 2%.

TABLE I
RELATIVE DIFFERENCE e_D BETWEEN THE END-TO-END DISTORTION $D(n)$ AND THE SUM OF SOURCE DISTORTION $D_s(n)$ AND CHANNEL DISTORTION $D_c(n)$

Video sequence	Coding bit rate	Packet loss ratio	Relative difference e_D
Foreman	150 kbps	2%	0.9%
Foreman	96 kbps	5%	1.2%
News	96 kbps	2%	1.2%
News	96 kbps	5%	1.5%
Salesman	96 kbps	2%	1.2%
Salesman	96 kbps	5%	1.3%
Carphone	128 kbps	2%	1.1%
Carphone	128 kbps	5%	1.0%

and channel distortion are uncorrelated with each other and that $D(n) = D_s(n) + D_c(n)$. Using the R-D model presented in Section II, we can accurately estimate $D_s(n)$. Therefore, to estimate the end-to-end distortion $D(n)$, the only thing left is to estimate $D_c(n)$, which is explained in the following.

B. Statistical Analysis of Channel Distortion

At the decoder side, we employ the following error concealment scheme: if a MB is skipped by the decoder, both the motion vectors and the texture information are supposed to be lost if the data partition syntax option is turned off [2]. In this case, the decoder simply copies the MB at the same location from the previous decoded frame. With this simple and efficient error concealment scheme, we develop a statistical analysis of the channel distortion. In standard video coding, such as H.263 and MPEG-4, there are two basic types of MBs: intra and inter. For a pixel in intra MBs, in case of no channel errors, its reconstruction value is $\hat{F}(n, i)$. If the MB is lost, the reconstruction value of pixel i is $\tilde{F}(n-1, i)$, which is copied from the previous decoded frame. Therefore, the expected channel distortion is

$$\begin{aligned} D_c^I(n) &= E\{[\hat{F}(n, i) - \tilde{F}(n, i)]^2\} \\ &= p \cdot E\{[\hat{F}(n, i) - \tilde{F}(n-1, i)]^2\} \\ &= p \cdot E\{[\hat{F}(n, i) - \hat{F}(n-1, i) + \hat{F}(n-1, i) \\ &\qquad - \tilde{F}(n-1, i)]^2\} \\ &= p \cdot E\{[\hat{F}(n, i) - \hat{F}(n-1, i)]^2\} \\ &\quad + p \cdot E\{[\hat{F}(n-1, i) - \tilde{F}(n-1, i)]^2\} \\ &= p \cdot RFD(n, n-1) + p \cdot D_c(n-1) \end{aligned} \tag{9}$$

where $RFD(n, n-1)$ represents the mean square error (MSE) between the reconstructed frames n and $n-1$. It should be noted that the fourth identity in (9) is based on assumption that the frame difference and the channel distortion are uncorrelated with each other. Note that the joint source-channel bit allocation and intra refreshing rate selection operate before quantization and coding of the current frame. At this stage, $\hat{F}(n, i)$ is not available. However, we do know the MSE between the original frames n and $n-1$, defined as

$$\mathcal{F}_d(n, n-1) = E\{[F(n, i) - F(n-1, i)]^2\}. \tag{10}$$

If we assume

$$RFD(n, n-1) = a \cdot \mathcal{F}_d(n, n-1) \tag{11}$$

where a is a constant, (9) then becomes

$$D_c^I(n) = ap \cdot \mathcal{F}_d(n, n-1) + p \cdot D_c(n-1). \tag{12}$$

If we regard the video encoder as low-pass filter [3], then the reconstruction frame is the filter output of the original frame. Note that a low-pass filter removes the energy in the original signal. From this point of view, the constant a can be regarded as the energy loss ratio of the encoder filter. It mainly depends how much information is discarded by the coding algorithm. In other words, it depends on the video quality level of the current wireless video communication session. More precisely, it is related to the average quantization step size. In this work, it is estimated using the statistics from previous frames.

For a pixel in inter MBs, in case of no channel errors, its reconstruction value is $\hat{e}(n, i) + \tilde{F}(n-1, j)$ where pixel j is the motion prediction of pixel i. (If the half-pel motion estimation is used, j could point to a half-pel position.) If the MB is lost, the reconstruction value of pixel i is $\tilde{F}(n-1, i)$, which is copied from the previous decoded frame. Therefore, the expected channel distortion is

$$\begin{aligned} D_c^P(n) &= E\{[\hat{F}(n, i) - \tilde{F}(n, i)]^2\} \\ &= (1-p) \cdot E\{[\hat{F}(n, i) - \hat{e}(n, i) - \tilde{F}(n-1, j)]^2\} \\ &\quad + p \cdot E\{[\hat{F}(n, i) - \tilde{F}(n-1, i)]^2\} \\ &= (1-p) \cdot E\{[\hat{F}(n-1, j) - \tilde{F}(n-1, j)]^2\} \\ &\quad + p \cdot RFD(n, n-1) + p \cdot D_c(n-1). \end{aligned} \tag{13}$$

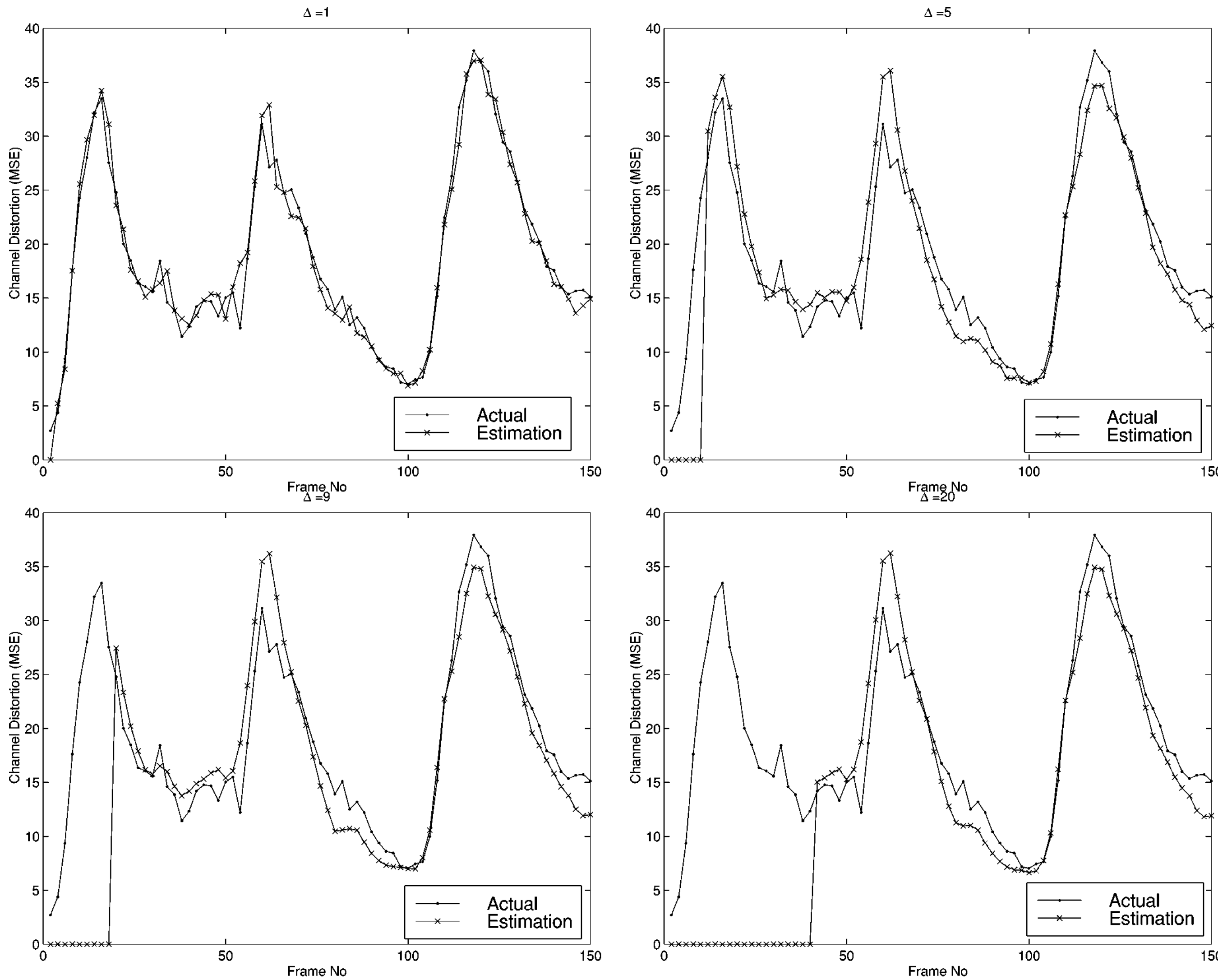


Fig. 6. Channel-distortion estimation results for "Salesman" QCIF video coded at 64 kbps.

Note that $\{\hat{F}(n-1, j)\}$ is the motion-compensation reference frame. If we assume

$$E\{[\hat{F}(n-1, j) - \tilde{F}(n-1, j)]^2\} = b \cdot D_c(n-1) \quad (14)$$

where b is a constant, we have

$$D_c^P(n) = [(1-p)b+p] \cdot D_c(n-1) + pa \cdot \mathcal{F}_d(n, n-1). \quad (15)$$

Note that b is a constant describing the motion randomness of the video scene. In frame n, let M be the total number of MBs and L be the number of intracoded MBs. $\beta = L/M$ is then the intra refreshing rate. The overall channel distortion is then given by

$$\begin{aligned} D_c(n) &= \beta D_c^I(n) + (1-\beta) D_c^P(n) \\ &= [(1-\beta)(1-p)b + p] \cdot D_c(n-1) \\ &\quad + pa \cdot \mathcal{F}_d(n, n-1) \\ &= \Gamma_1 \cdot D_c(n-1) + \Gamma_2 \cdot \mathcal{F}_d(n, n-1) \end{aligned}$$

where

$$\begin{aligned} \Gamma_1 &= (1-\beta)(1-p)b + p \\ \Gamma_2 &= pa. \end{aligned} \quad (16)$$

We can see that this model reveals the inherent relationship between the channel distortion and the characteristics of the input video data.

C. Asymptotic Behavior of Channel Distortion

Let T be the total number of coded video frames. The average channel distortion of all frames, denoted by $\overline{\mathcal{D}}_c(T)$, is given by

$$\begin{aligned} \overline{\mathcal{D}}_c(T) &= \frac{1}{T} \sum_{n=1}^{T} D_c(n) \\ &= \frac{1}{T} \sum_{n=1}^{T} \Gamma_1^n \cdot D_c(0) + \frac{\Gamma_2}{T} \sum_{n=1}^{T} \sum_{i=1}^{n} \Gamma_1^i \cdot \mathcal{F}_d(i, i-1) \\ &= \frac{1}{T} D_c(0) \frac{1}{1-\Gamma_1} + \frac{\Gamma_2}{1-\Gamma_1} \frac{1}{T} \\ &\quad \cdot \sum_{i=1}^{T} \mathcal{F}_d(i, i-1)(1-\Gamma_1^{T-i}). \end{aligned} \quad (17)$$

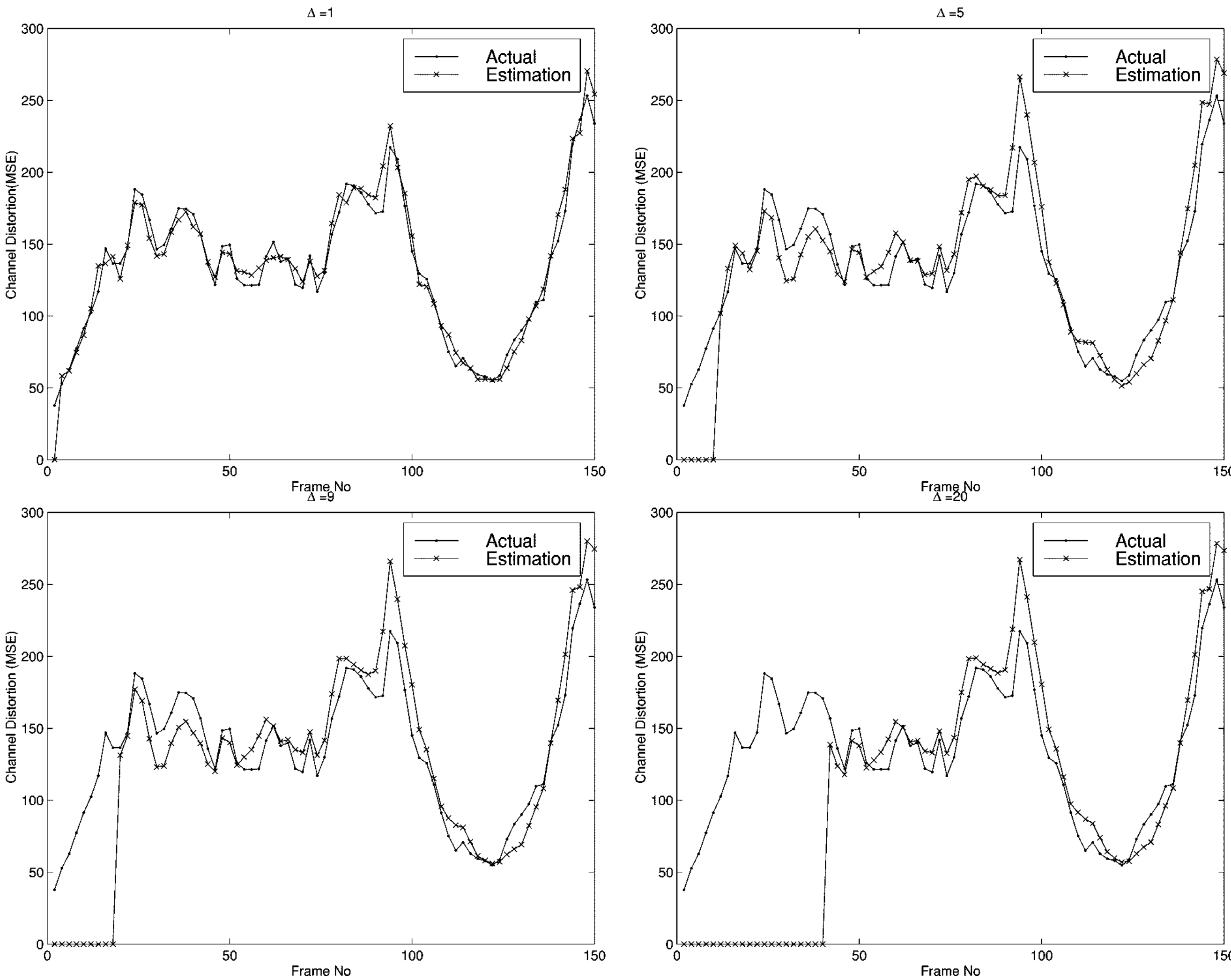


Fig. 7. Channel-distortion estimation results for "Foreman" QCIF video coded at 96 kbps and 15 fps.

It can be seen from (10) that $\mathcal{F}_d(n, n-1)$ is upper bounded, at least by $C = 255 \times 255$. Therefore

$$\sum_{i=1}^{T} \mathcal{F}_d(i, i-1)(1-\Gamma_1^{T-i}) \leq C\frac{1}{1-\Gamma_1}. \quad (18)$$

We then have

$$\overline{\mathcal{D}}_c = \lim_{T\to\infty} \overline{\mathcal{D}}_c(T) = \frac{\Gamma_2}{1-\Gamma_1} E[\mathcal{F}_d(n, n-1)] \quad (19)$$

$$= \frac{a}{(1-b+b\beta)}\,\frac{p}{1-p}\,E[\mathcal{F}_d(n, n-1)] \quad (20)$$

where $E[\mathcal{F}_d(n, n-1)]$ is the average value of the frame difference $\mathcal{F}_d(n, n-1)$ over the whole video scene. From (20), we observe that, asymptotically, the average channel distortion caused by packet loss is proportional to the mean frame difference.

D. Fast Channel-Distortion Estimation

In wireless video communication over noisy channels, with the feedback information on the channel condition and transmission status, the encoder can determine the decoded picture quality of frame $n-\Delta$ and its previous frames, where Δ is the feedback delay (in the unit of frame interval) and n is the current frame number. In other words, $\{D_c(n-\Delta-m)|m \geq 0\}$ are available at the encoder through channel feedback. With $\{D_c(n-\Delta-m)\}$, we can apply the channel model in (16) recursively to compute the channel distortion $D_c(n)$ for the current frame n as follows:

$$D_c(n) = \Gamma_1^{\Delta} D_c(n-\Delta) + \Gamma_2 \sum_{l=0}^{\Delta-1} \Gamma_1^{l} \cdot \mathcal{F}_d(n-l, n-l-1). \quad (21)$$

E. Experimental Results

To test the performance of the proposed fast channel-distortion estimation scheme, we simulate packet loss in MPEG-4 video coding and use this scheme to estimate the channel distortion for different videos at different encoding settings and channel conditions. The configuration of each test is shown in Table II. The packet size is 96 bytes. In each test, the video sequence is simulated 20 times and the average channel distortion is computed. The estimation results for "Salesman" with channel feedback delay $\Delta = 1$, 5, 9, and 20 frames are shown in Fig. 6. The estimation results for "Foreman" and "Carphone" are shown in Figs. 7 and 8, respectively. From these extensive

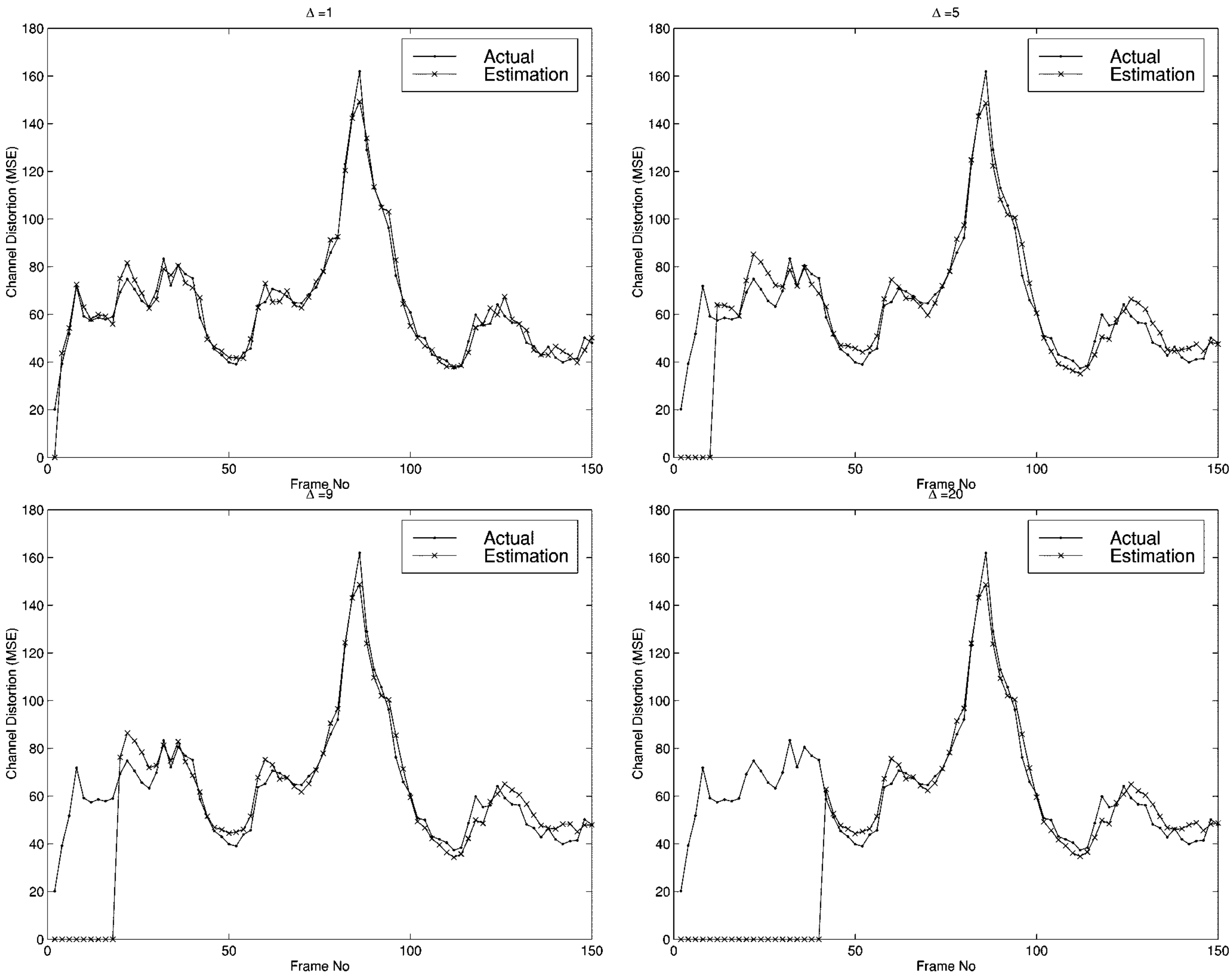


Fig. 8. Channel-distortion estimation results for "Carphone" QCIF video coded at 96 kbps and 15 fps.

TABLE II
EXPERIMENT CONFIGURATION FOR PERFORMANCE TEST OF THE FAST CHANNEL DISTORTION ESTIMATION SCHEME

Video Name	Source Coding Rate	Frame Rate	p	β
Salesman	64 kbps	15 fps	0.15	0.1
Foreman	96 kbps	15 fps	0.15	0.1
Carphone	96 kbps	15 fps	0.05	0.02

simulation results we can see that the channel-distortion model is very accurate. More importantly, when the channel feedback delay is significantly increased, the model estimation accuracy is largely maintained. This implies that the model is also very robust.

Another way to understand the channel-distortion model and the above experimental results is as follows. If we know the channel distortion $D_c(n)$ of the current frame, we can accurately predict the channel distortion of many frames ahead (as many as Δ frames). Note that during the prediction process, the model only needs the frame differences of the original sequence, which is easily available at the encoder. Based on the predicted channel-distortion behavior, we can achieve better resource allocation and video quality at the receiver end. The accuracy and robustness of experimental results also suggest that the proposed channel-distortion model reveals the close relationship between the channel distortion and the characteristics of the input video data (the frame difference information). This statistical model has provided helpful insight on the behavior of channel distortion, as well as its impact on the end-to-end video quality.

F. Estimate Packet-Loss Ratio

Note that in the proposed channel-distortion model, p refers to the probability of packet loss. In wireless video communication over a noisy channel, the parameter BER, denoted as p_e, is often used to describe the channel error condition. Therefore, we need to find the relationship between p_e and p. In this paper, we consider a random binary symmetric channel (BSC) model. In many cases, bursty errors can be converted into random errors with pre-interleaving [22]. The (N, K) RS block code with 8 bits per symbol is used for channel coding. The code rate

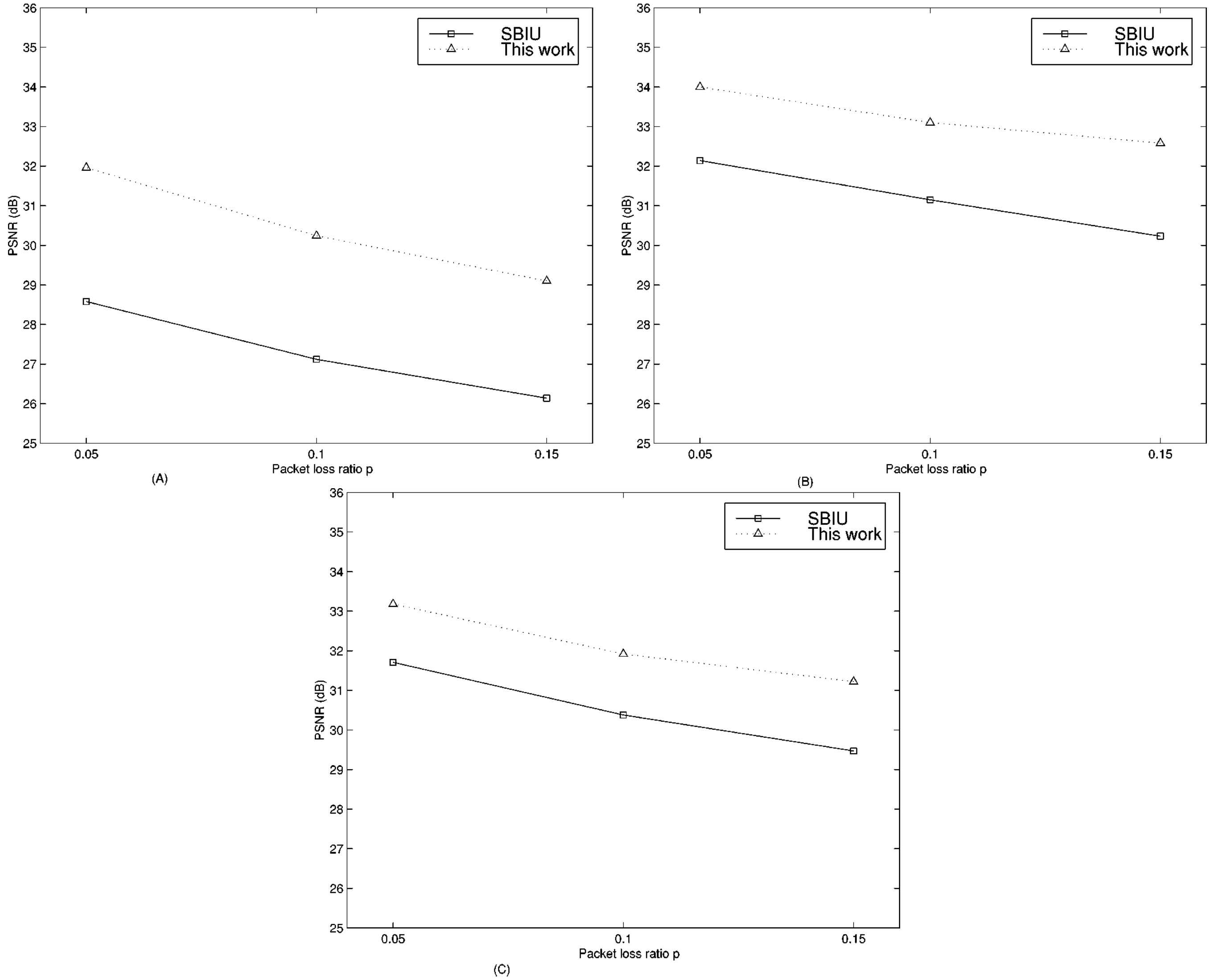


Fig. 9. PSNR performance of the proposed AIR algorithm for three test videos: (a) "Foreman" at 250 kbps; (b) "News" at 150 kbps; and (c) "Carphone" at 150 kbps.

$r = K/N$ is determined by the joint source-channel bit allocation scheme proposed in Section V. Note that in the RS code, any error pattern with no more than

$$T = \left\lfloor \frac{N-K}{2} \right\rfloor \tag{22}$$

symbol errors can be corrected. We denote the symbol error rate (SER) as $\mathcal{E}$. The decoded symbol error rate is then given by

$$\mathcal{E}_d = 1 - \sum_{i=0}^{K} \sum_{j=0}^{N-K} \binom{K}{i} \binom{N-K}{j} \mathcal{E}^{i+j} (1-\mathcal{E})^{N-i-j} \cdot \eta(i, j) \tag{23}$$

where

$$\eta(i, j) = \begin{cases} 1, & \text{if } i + j \le T \\ (K-i)/K, & \text{otherwise.} \end{cases} \tag{24}$$

Once a symbol cannot be corrected by the RS decoder, it will be detected by the video encoder. In this case, the decoder will jump to the next packet starting with a re-synchronization mark and skips all the intermediate symbols. Suppose the packet has l symbols. The packet-loss ratio is determined by

$$p = 1 - (1 - \mathcal{E}_d)^l \tag{25}$$

which is used by the channel-distortion model (16). Typical plots of p as a function of the channel BER p_e can be found in [5].

IV. Adaptive Intra Mode Selection

Intra MB refreshing can significantly improve the error resilience capability of the coded video data. This can be observed from (9), (13), and (16). Note that $D_c^P(n)$ is always larger than $D_c^I(n)$. Their difference is $(1-p)b \cdot D_c(n-1)$ which is exactly the portion of channel distortion in the previous frames propagated to the current frame through motion compensation. If the intra refreshing rate β is increased from β_0 to β_1, the corresponding channel distortion will be reduced by the amount of $(\beta_1 - \beta_0)\ (1-p)b \cdot D_c(n-1)$. However, from (3), we can see that the source distortion increases. This leads to a tradeoff in selecting the value of β. Clearly, the optimal tradeoff point corresponds to the value of β which minimizes the overall distortion $D(n) = D_s(n) + D_c(n)$.

A. Adaptive Intra Refreshing Algorithm

With source and channel-distortion models developed in Sections II and III, we can estimate the distortion $D(n)$ for any given R_s and β. We then find the value of β which minimizes the distortion function $D(n)$. Once the optimal intra refreshing rate β is determined, we need to select $L = \beta \cdot M$ MBs to be intra coded. The simplest way is just to use random selection. Statistically, this random selection scheme will achieve the minimum average channel distortion. In some intra mode selection algorithms reported in the literature [4], [5], [20], the intra/inter mode decision is made individually for each MB. However, we notice that at the decoder the previous frame is a random variable due to the randomness of channel errors. All the R-D computations and optimization should be handled in a statistical sense. This statistical procedure considers only the overall R-D behavior of the whole frame/video, instead of the specific characteristics of each individual MB. Based on this observation, we believe that on average, it is sufficient to randomly select the intra-coded MBs, since the overall optimization is already guaranteed at the frame level. This will be demonstrated by our experimental results in the following section. Another advantage of the random selection is its significantly reduced computational complexity and implementation cost. For convenience, we refer to the proposed Adaptive Intra Refreshing algorithm as "AIR."

B. Experimental Results

We implement the AIR algorithm in the MPEG-4 codec and compare its performance with the Scattered-Block Intra Update (SBIU) algorithm [20]. The three test QCIF video sequences are "Foreman" at 250 kbps, "News" at 150 kbps, and "Carphone" at 150 kbps. The frame rate is 15 fps. The packet size is set to be 96 bytes. It should be noted that, for fair comparison, in this simulation no RS channel coding is applied, and all the available bit rate is assigned for source coding. Using the AIR algorithm, the encoder can adjust intra refreshing rate β according to the channel conditions and the characteristics of the input video data. The average PSNR results for different packet-loss ratios p are depicted in Fig. 9. The average PSNR is computed over 20 different channel realizations. It can be seen that the proposed AIR algorithm significantly improves the video quality, especially for high motion video sequences. For example, for the "Foreman" video sequence, the quality improvement is about 3.2 dB. The PSNR improvement for each coded frame from the "Carphone" video is depicted in Fig. 10. From the experimental results reported in [4], we can see that the proposed frame-level AIR algorithm has almost similar performance to the MB-level adaptive intra mode selection scheme. However, the channel-distortion estimation in [4] is based on a statistical simulation of the decoder at the encoder side, which involves potentially high computational complexity and implementation cost, and therefore is not suitable for real-time video coding and communication over wireless devices. Next, we test the performance of the AIR algorithm

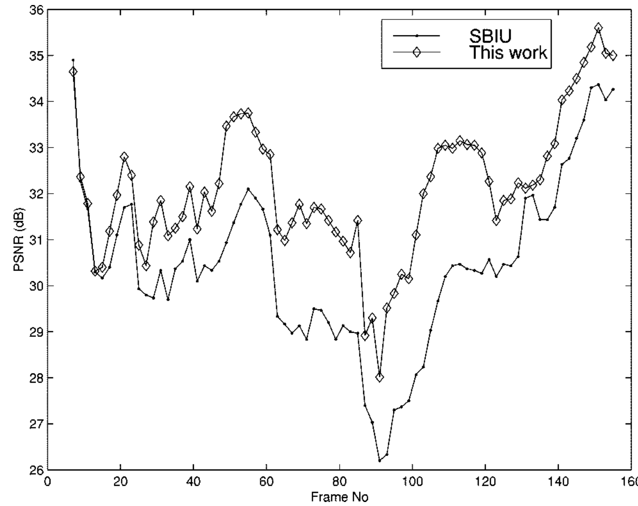


Fig. 10. Intra-mode selection results for "Carphone."

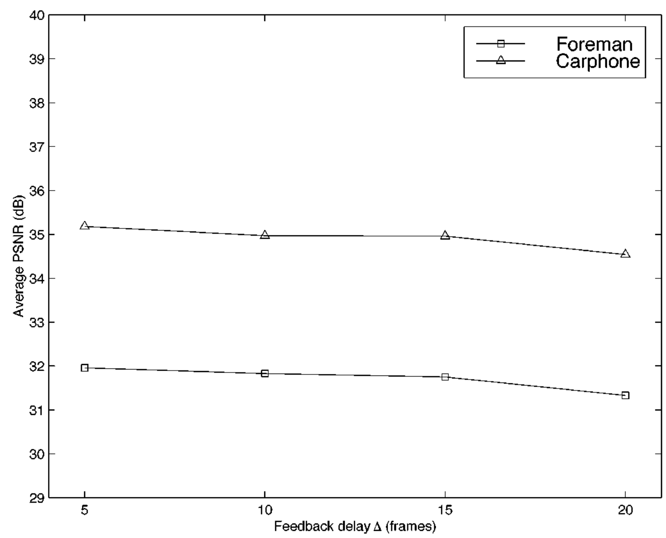


Fig. 11. PSNR performance of the AIR algorithm with different channel feedback delay. The videos are coded at 250 kbps and 15 fps. The packet-loss ratio is $p = 0.05$.

at different channel feedback delays. In Fig. 11 we plot the average PSNR for "Foreman" and "Carphone" both coded at 250 kbps and 15 fps. The packet-loss ratio is $p = 0.05$. We can see that the performance of the AIR algorithm degrades little as the feedback delay increases. Experiments over other video sequences and coding settings yield similar results.

V. JOINT SOURCE CHANNEL RATE CONTROL

A. Algorithm Description

Intra MB refreshing is an effective way to stop the channel-distortion propagation at the decoder. To reduce the amount of channel distortion introduced to the video data stream, we need to resort to channel coding to correct the bit errors using channel coding. In this case, the overall bit rate is divided into two parts: source coding bit rate R_s and channel coding bit rate R_c. One important parameter of the RS coder is its code rate $r = R_s/R$. To minimize the overall picture distortion at the

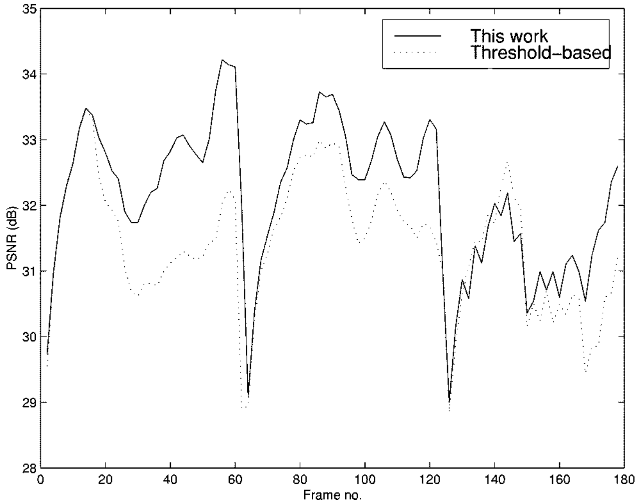


Fig. 12. Decoded picture quality comparison when the proposed joint source-channel bit allocation scheme and the threshold-base scheme are applied. The video sequence is "Foreman" coded at 96 kbps and 15 fps. The threshold SER is 0.1%.

receiver end, we need to adjust the parameters of the video encoder and the RS channel codes for different input video sequences and different channel conditions. Specifically, given a target bit rate R which is determined by the available channel bandwidth, we need to perform optimal bit allocation between source and channel coding. In addition, we need to adjust the encoder settings to achieve the allocated bit budget. From (16), we can see that the channel distortion D_c is a function of p, which is in turn a function of the channel coding rate r. Therefore, the optimal joint source-channel bit allocation can be formulated as follows:

$$\min_{0 \le r, \beta \le 1} D(r, \beta) = D_c([1-r] \cdot R, \beta) + D_s(r \cdot R, \beta). \quad (26)$$

In Sections II and III, we have explicitly estimated the source and channel-distortion functions. Based on the estimation, we only need to find the values of (r, β) which minimize the objective function in (26). Once r is determined, we can use the linear rate control algorithm proposed in our previous work [10] to select the quantization parameters to achieve the target bit rate $R_s = r \cdot R$ at the video encoder.

B. Experimental Results

We implement the above joint source-channel rate control (JSCRC) algorithm with optimal MB intra refreshing in the MPEG-4 codec [16]. The compressed video data stream is further coded by RS codes. At the receiver end, the corrupted bit stream is first RS decoded to correct the bit errors. For different input video sequences and channel conditions, we test the proposed joint source-channel bit allocation and rate control algorithm, and compare its performance with the conventional threshold-based bit allocation scheme [21]. In the threshold-based scheme, no joint source-channel bit allocation is performed. A fixed portion of the channel bit rate is assigned to the RS coder such that its error correction capacity is above some given threshold. In this way, the video encoder and RS codes do not consider the varying characteristics of the input video data and the transmission channel.

TABLE III
PICTURE QUALITY COMPARISON FOR DIFFERENT VIDEO SEQUENCES AND SER THRESHOLDS

Video Name	Channel Bit rate	PSNR This work	PSNR Threshold-Base 0.05%	0.1%	0.2%
Foreman	96	31.40	30.23	30.71	29.81
Flowergarden	256	30.20	29.02	29.26	27.60
Salesman	96	34.40	34.21	33.28	33.52
Akiyo	64	36.98	36.75	36.24	36.56

In Fig. 12, we plot the decoded picture quality for "Foreman" QCIF video coded at 96 kbps and 15 fps when the JSCRC algorithm and the conventional threshold-based bit allocation algorithm are applied. In this case, for the threshold-based scheme, the SER threshold is set to 0.1%. In Table III, we summarize the PSNR results for different input video sequences and different SER threshold settings. It can be seen that with adaptive rate allocation and control, the video encoder always chooses appropriate source/channel coding bit rates and encoder settings, which yields significantly improved picture quality at the receiver end.

VI. CONCLUSION

There are three major contributions in this work. First, we have developed an R-D model for DCT-based video coding which incorporates the MB intra refreshing rate. This model can accurately estimate the R-D function for any given intra refreshing rate before a video frame is coded. Second, we have developed a theoretical analysis of the channel distortion caused by channel errors and inter-frame propagation. Our statistical channel-distortion model reveals the inherent relationship between the channel distortion and input video characteristics. This model has very low computational complexity and implementation cost and is therefore suitable for wireless applications. Experimental results show that it is able to estimate the channel distortion accurately and robustly with a minimum delay in processing. Finally, based on the proposed source and channel-distortion models, we have developed a scheme for adaptive intra mode selection and joint source-channel rate control. Extensive experimental results have demonstrated that this scheme significantly improves the end-to-end video quality for wireless video coding and transmission.

REFERENCES

[1] ITU-T, "Video coding for low bit rate communications," ITU-T Recommendation H.263, version 1, version 2, Jan. 1998.
[2] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 19–31, Feb. 1997.
[3] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1012–1032, June 2000.
[4] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 966–976, June 2000.
[5] G. Cote, S. Shirani, and F. Kossentini, "Optimal mode selection and synchronization for robust video communications over error-prone networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 952–965, June 2000.

[6] J. Y. Liao and J. Villasenor, "Adaptive intra block update for robust transmission of H.263," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 30–35, Feb. 2000.
[7] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by ATM cell loss," in *Proc. Int. Conf. Acoust., Speech, and Signal Processing*, vol. 3, San Francisco, CA, Mar. 1992, pp. 545–548.
[8] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Indust. Appl. Math.*, June 1960.
[9] Z. He, Y. Kim, and S. K. Mitra, "Object-level bit allocation and scalable rate control for MPEG-4 video coding," in *Proc. Workshop and Exhibition on MPEG-4*, San Jose, CA, June 2001.
[10] ——, "A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, May 2001.
[11] 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, pp. 12–20, Feb. 1996.
[12] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 38, pp. 82–93, Jan. 1990.
[13] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sept. 1968.
[14] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[15] 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, pp. 287–298, Apr. 1997.
[16] MoMuSys codec, "MPEG4 verification model version 7.0," ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Associated Audio MPEG97, Bristol, U.K., Mar. 1997.
[17] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 246–250, Feb. 1997.
[18] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172–185, Feb. 1999.
[19] Q. Zhang, W. Zhu, and Y. Zhang, "Resource allocation with adaptive QoS for multimedia streaming," in *IEEE Wireless Conf. Communications and Networking Conference (WCNC)*, Chicago, IL, Sept. 2000.
[20] G. Cote and F. Kossentini, "Optimal intra coding of blocks for robust video communication over the Internet," *Image Commun.*, pp. 25–34, 1999.
[21] V. Parthasarathy, J. W. Modestino, and K. S. Vastola, "Reliable transmission of high-quality video over ATM networks," *IEEE Trans. Image Processing*, vol. 8, pp. 361–374, Mar. 1999.
[22] B. Girod and N. Farber, "Wireless video," in *Compressed Video Over Networks*, M.-T. Sun and A. R. Reibman, Eds. New York, NY: Marcel Dekker, Inc., 2000.

Zhihai He (M'01) received the B.S. degree from Beijing Normal University, Beijing, China, and the M.S. degree from the Institute of Computational Mathematics, Chinese Academy of Sciences, Beijing, both in mathematics, in 1994 and 1997, respectively. He received the Ph.D. degree in electrical engineering from the University of California, Santa Barbara, in 2001.

In 2001, he joined Sarnoff Corporation, Princeton, NJ, as a Member of Technical Staff. His current research interests include image compression, video coding, network transmission, wireless communication, and embedded system design.

Dr. He received the IEEE Circuits and Systems Society TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Best Paper Award in 2002.

Jianfei Cai (S'00) received the B.S. degree in electrical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1994, the M.S. degree in electrical engineering from Southeast University in 1997, and the Ph.D. degree in electrical engineering from the University of Missouri-Columbia in 2002.

From 1997 to 2002, he was a Research Assistant in the Department of Electrical Engineering, University of Missouri-Columbia. He was with Phillips Research Center, Briarcliff Manor, NY, in the summer of 1998, with Microsoft Research Asia, Beijing, China, in the summer of 2000, and with Sarnoff Corporation, Princeton, NJ, in 2001. His research interests include multimedia communications, image/video coding and streaming, wireless communications, digital signal processing, image processing, and pattern recognition. He has authored over 20 refereed technical journal and conference publications.

Dr. Cai is a member of Sigma Xi.

Chang Wen Chen (S'86–M'90–SM'97) received the B.S. degree from University of Science and Technology of China in 1983, the M.S.E.E. degree from University of Southern California at Los Angeles in 1986, and the Ph.D. degree from University of Illinois at Urbana-Champaign in 1992.

He is currently Head of the Interactive Media Group at Sarnoff Corporation, Princeton, NJ. From August 1992 to September 1996, he was on the faculty of the Department of Electrical Engineering, University of Rochester, Rochester, NY. Since September 1996, he has been Assistant Professor and then Associate Professor of Electrical Engineering at the University of Missouri-Columbia. He has consulted with Kodak, NASA Goddard Space Flight Center, and Microsoft Research. His current research interests include multimedia processing and compression, image and video coding, MPEG-4, MPEG-7, mobile wireless multimedia communication, Internet multimedia streaming, telemedicine, biomedical image processing, and visualization.

Dr. Chen is currently serving as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He is also on the Editorial Board of the *Journal of Visual Communication and Image Representation* and the *Journal of Computerized Medical Imaging and Graphics*. He has been a Guest Editor for the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATION Special issue on Error-Resilient Image and Video Transmission, and a Guest Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Special Issue on Wireless Video. He has received research awards from NSF, NASA, the Whitaker Foundation, and industry.