Rate Control for Robust Video Transmission over Burst-Error Wireless Channels

Chi-Yuan Hsu, Antonio Ortega, and Masoud Khansari

Abstract— **We study the problem of rate control for transmission of video over burst-error wireless channels, i.e., channels such that errors tend to occur in clusters during fading periods. In particular we consider a scenario consisting of packet based transmission with Automatic Repeat reQuest (ARQ) error control and a back channel. We start by showing how the delay constraints in real time video transmission can be translated into rate constraints at the encoder, where the applicable rate constraints at a given time depend on future channel rates. With the acknowledgments received through the back channel we have an estimate of the current channel state. This information, combined with an *a priori* model of the channel allow us to statistically model the future channel rates. Thus the rate constraints at the encoder can be expressed in terms of the expected channel behavior. We can then formalize a rate distortion optimization problem, namely, that of assigning quantizers to each of the video blocks stored in the encoder buffer such that the quality of the received video is maximized. This requires that the rate constraints be included in the optimization, since violating a rate constraint is equivalent to violating a delay constraint and thus results in losing a video block. We formalize two possible approaches. The first one seeks to minimize the distortion for the *expected rate constraints* given the channel model and current observation. The second approach seeks to allocate bits so as to *minimize the expected distortion* for the given model. We use both dynamic programming and Lagrangian optimization approaches to solve these problems. Our simulation results demonstrate that both the video distortion at the decoder and packet loss rate can be significantly reduced when incorporating the channel information provided by the feedback channel and the *a priori* model into the rate control algorithm.**

Keywords— **ARQ, channel model, channel feedback, packet video, delay constraint, VBR video.**

I. Introduction

WIRELESS channels are increasingly being considered as a transport medium for various types of multimedia information. While the appeal of tetherless mobility is great, numerous issues need to be resolved in order for wireless transport of real time multimedia data to become a reality (including communications issues, low power implementation, etc.) We consider a scenario where, due to the user's mobility, the channel behavior will be inherently time-varying, with periods of correct transmission alternating with periods of high error rates. In this paper, we concentrate on how a real time video application can be supported over such a time varying burst-error channel, rather than on the specifics of the physical layer of the channel. We will only assume that the channel behavior can be characterized by simple burst-error models and will provide experimental results for two such models [1], [2]. However, the proposed techniques are applicable to more general cases. In particular, our goal is not to validate specific channel models but rather to show how, given such channel models, optimized rate control strategies can be implemented. Thus, our techniques could be applied with alternative channel models for the same wireless environments considered, or they could be used, with appropriate models, for other time-varying channel scenarios. Therefore, the performance of the described techniques will depend in part on how well the channel model matches the actual channel behavior.

Transmission of real time video is challenging because of the delay constraints involved (i.e., information which arrives too late at the decoder is considered lost), and because of the negative impact of channel losses on the perceptual quality of video at the decoder. Thus to achieve high fidelity video quality at the decoder requires a robust transmission scheme [3]. Indeed, uncorrected channel errors may result in significant quality degradation at the decoder. This is particularly evident in standard coders, such as those based on MPEG or H.263, where variable length coding is used (the variable length decoder is likely to lose synchronization) or where compression involves a predictive coding scheme, such as motion compensation (error can therefore propagate through several frames.) While numerous approaches for error concealment have been described in the literature [4], it is in general preferable to ensure as error-free a transmission as possible.

To provide the required protection one can use error control techniques, which can be roughly categorized into open-loop (e.g., forward error correction, FEC) and closed-loop (e.g., automatic repeat request, ARQ). Obviously error correction comes at the cost of reduced bandwidth available for transmission: this is due to the error correction overhead in the FEC case and to the need to retransmit data in the ARQ case. While FEC is often used for wireless mobile channels [5], [6], in a two-way communication system the available feedback channel can be used for error resilience by allowing the receiver to request the retransmission of erroneous packets using ARQ [7]. Using ARQ error control for the mobile radio channels has been recently proposed as an alternative to a purely FEC based

Chi-Yuan Hsu is with the Sony Semiconductor Company of America, San Jose, California 95134-1940

Antonio Ortega is with the Integrated Media System Center, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564

Chi-Yuan Hsu and Antonio Ortega were supported in part by the National Science Foundation under grant MIP-9502227 (CAREER) and by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, with additional support from the Annenberg Center for Communication at the University of Southern California, the California Trade and Commerce Agency

Masoud Khansari is with the Hewlett-Packard Laboratories, Palo Alto, California 94304-1126

approach [2], [1]. In [8], [9], ARQ feedback is also used for error concealment of the transmitted video. ARQ information can also be used within the recently finalized H.263+ standard to ensure that motion compensated prediction is only performed with respect to frames that have already been acknowledged [10].

ARQ approaches, assuming the existence of a back channel and sufficiently long end-to-end delays, are appealing in that retransmission is only required during periods of poor channel conditions. Thus ARQ schemes are inherently variable rate. However, to take full advantage of the error control capabilities of an ARQ scheme, we propose to *combine the ARQ feedback mechanism with the rate control mechanism at the video encoder.* By combining the ARQ feedback with a rate control algorithm at the encoder one can achieve an intuitively appealing result: the rate for the encoded video is reduced during the periods of poor channel conditions.

We concentrate here on selective-repeat ARQ where packets are continuously transmitted without waiting to receive acknowledgments of previously transmitted packets. We consider a general framework for the real time transmission of video data over a channel subject to random bursty losses. We first show how the end-to-end delay constraints on the video delivery can be translated into rate constraints at the encoder. These rate constraints depend on the channel behavior, which is assumed to be random but can be estimated based on a priori models and the channel state feedback.

In order to achieve higher compression ratios, video encoders typically utilize lossy compression techniques, that is, the decoded video is not an exact reproduction of the original. Thus, one can reduce the source rate at the cost of reduced quality (distortion) at the decoder. This leads to Rate-Distortion (R-D) techniques being natural frameworks to evaluate the performance of video transmission systems. Our goal will then be to design systems which minimize the distortion at the decoder for the given set of rate constraints, as derived from the end-to-end delay constraints. While perceptual quality does not always correlate exactly with the measured objective distortion (e.g., mean square error, MSE, which will be used in this paper), objective distortion measures are still a useful tool to evaluate received video quality.

Most commonly used video compression standards, such as MPEG, H.261 or H.263, (see [11], [12] for example) share a similar structure based on motion compensated frame prediction and block-based Discrete Cosine Transform (DCT) coding. In such DCT-based video source encoders, the resulting encoding rate and the incurred distortion are determined by how coarsely the DCT coefficients are quantized. In our formulation we allow the encoder to select one quantizer for each group of blocks (GOB), chosen out of a predefined discrete set of quantizers.

R-D based approaches have been used to solve numerous problems in image and video compression, from simple bit allocation in image coding [13] to various aspects of video coding including motion estimation [14], rate control [15], [16], [17], [18] and shape coding [19]. A description of some of the basic techniques for R-D optimization and a survey of recent work can be found in [20]. A more detailed description of these techniques can be found in [19]. Recent work has focused on the R-D optimization of video transmission over variable rate channels [21], [22], [17], [18], i.e., those where the number of bits transmitted changes for each time interval. However, the work in [21], [22], [17], [18] concentrated on situations (specifically transmission over an Asynchronous Transfer Mode, ATM, network) where the transmitter was allowed to select the channel rate and thus channel rates were known deterministically. Instead here we consider a scenario where the effective channel rate varies randomly. The main novelty of this work, as compared to earlier work on R-D driven rate control is that we introduce a probabilistic component in the rate control: the constraints on the rate that can be allocated to each video block depend on the future channel rates, for which we only have a statistical characterization. The work presented here extends the one we initially reported in [23], [24].

Also note that recent work [25], [26] has addressed similar scenarios where rate control at the video encoder is used to provide enhanced transmission robustness. However, to the best of our knowledge, no other work has considered this problem in an R-D optimization framework as that presented here.

Our proposed formulation is very general and it allows us to show how one can modulate the rate of the source to minimize the expected distortion at the receiver, given estimates of the channel state. The experimental results presented in this paper focus on a point-to-point wireless link for video transmission, where a feedback channel is available to the encoder, but the same algorithms may be applicable to other environments where transmission is subject to errors, e.g., video transmission over the Internet.

The paper is organized as follows: in Section II we define the delay constraints arising in a real-time video transmission system. From the delay constraints, a set of rate constraints on encoded video data are derived, such that meeting the rate constraints guarantees that the delay constraints are not violated. In Sections III and IV, we briefly describe the structure of our wireless video transmission system, and propose rate control algorithms for video transmission under such burst-error channels. We assume a wireless environment with a feedback channel and assume that an *a priori* probabilistic model of the channel behavior is available. The burst-error wireless channel is modeled as a Markov chain, and an ARQ scheme is used for error control. At the source encoder, we formulate an optimal rate control problem in an R-D framework, where the channel state observation and an a priori model of the channel are given. We formalize two possible approaches. The first one seeks to minimize the distortion for the *expected rate constraints* given the channel model and current observation. The second approach seeks to allocate bits so as to *minimize the expected distortion* for the given model. In Section V we use both dynamic programming and Lagrangian optimization approaches to solve these problems. Our simulation results, presented in Section VI, demonstrate that both the video distortion at the decoder and packet loss rate can be significantly reduced when incorporating the channel information provided by the feedback channel and the *a priori* model is incorporated into the rate control algorithm. We also evaluate how inaccurate channel models and delay in the channel state feedback affect our proposed rate control approaches.

II. Equivalence between Delay and Rate Constraints

In this paper, we define a *real-time* video transmission system as a system in which each video frame is captured, encoded, transmitted, decoded and displayed in real-time within some acceptable delay interval. Such delay interval is referred to as the "end-to-end" delay of video transmission. We consider here the case where the video frame rate (number of frames per second) is constant, and the same, at both encoder and decoder (i.e., frames cannot be "dropped"). Under these conditions, the end-to-end delay per frame, ΔT, will have to be constant. Thus, each frame captured at the transmitter at time t will have to be decoded and available for display before time $t + \Delta T$. Video data that arrives at the decoder too late to be decoded by its scheduled display time is useless and is considered lost.

Clearly, frame skipping at the decoder results in quality loss, especially when motion compensated video coding is used, while skipping frames at the encoder can be done without as heavy a quality penalty. However in this latter case there will also be end-to-end delay constraints for those frames that are transmitted. Different video applications have different requirements in terms of delay. In interactive video communications (e.g., video conferencing) low delay is required, while in one-way video transmission, (e.g., broadcast or video on demand) the end-to-end delay is only noticeable to the user as an initial latency, i.e., the time interval between the start of the video transmission session and the time the first video frame is displayed. While our formulation is generic it is clear that it will be more suited for scenarios with longer end-to-end delay.

Typical compressed video bitstreams are Variable Bit Rate (VBR), i.e., each frame is compressed using a different number of bits. When transmitting over a Constant Bit Rate (CBR) channel, buffers at the encoder and decoder are required to smooth out the variations in the encoding rates. Buffering data requires extra memory at both encoder and decoder and introduces additional delay to data transmission. As the encoder is allowed to produce a more variable rate (more bits for difficult frames, fewer bits for easy frames, for example) the overall quality will be better. Thus larger buffers, or equivalently, increased end-to-end delay, will tend to result in higher video quality [27], [15]. Traditionally, rate control has been studied from the point of view of memory, i.e., rate control was required to avoid overflowing the available buffers at encoder and decoder. However, in what follows we will tackle the problem as in [27], [21], i.e., we will assume that sufficient physical memory is available and formulate the problem from the point of view of end-to-end delay.

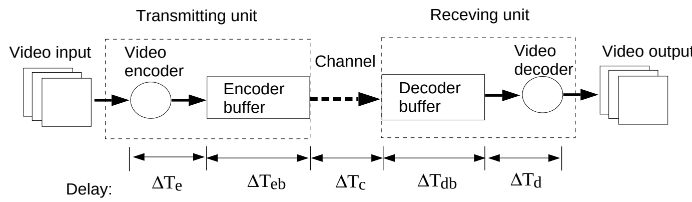

Fig. 1. Delay components of a communication system.

A. Delay Components

Following the MPEG, H.263 standard structures we will assume that the basic encoding/decoding units are so-called macroblocks which comprise several 8×8 pixel blocks[1]. These macroblocks can then be grouped as follows: A *processing group of blocks* (P-GOB) contains a set of macroblocks that are processed together by encoder or decoder. That is, in order for correct encoding (resp. decoding) to take place all the macroblocks in a P-GOB have to be available *simultaneously* at the encoder (resp. decoder) input. The size of the P-GOB depends on the exact architecture of the video encoder and decoder. A *quantization group of blocks* (Q-GOB) contains a set of macroblocks which get assigned the same quantization scale. In other words, the Q-GOB size defines the minimum granularity of the quantization assignment. The Q-GOB size is a source coding parameter and typically allowing quantizers to be chosen for smaller units (Q-GOB size small) results in better quality video. We will assume that the P-GOB and Q-GOB sizes are constant. Typical sizes for both can range from 1 macroblock to a complete frame.

We will now discuss the delay constraints affecting each P-GOB. We will also assume without loss of generality that the P-GOB and Q-GOB sizes are the same[2]. In the remainder of the paper we will thus refer to P-GOBs and Q-GOBs indistinctly as GOBs or blocks.

In typical video communications systems, the end-to-end delay each GOB experiences (from the time it is obtained from the input video buffer to the time it is placed in the display video buffer) consists of several delay components. For example the total constant delay ΔT experienced by each frame (see also Fig. 1) can be broken up into:

$$\begin{aligned} \Delta T &= \Delta T_e \quad \text{(Encoder delay)} \\ &+ \Delta T_{eb} \quad \text{(Encoder buffer delay)} \\ &+ \Delta T_c \quad \text{(Channel delay)} \\ &+ \Delta T_{db} \quad \text{(Decoder buffer delay)} \\ &+ \Delta T_d \quad \text{(Decoder delay).} \end{aligned} \tag{1}$$

[1]Several alternative macroblock sizes are possible, for example, an MPEG-2 macroblock could include 16x16 luminance pixels and 2 sets of 8x8 chrominance pixels [11]

[2]The delay constraints depend only on the size of the P-GOBs. Once this has been fixed one can find solutions for the resulting constrained optimization problems: the size of the Q-GOBs only constraints the set of available quantization settings.

In a more general case, assume that there are G GOBs in each frame, then it is still true that the overall delay for the frame can at most be ΔT. However the delay may be different for specific GOBs. For example, for the i-th GOB in one particular frame the following must hold[3],

$$\Delta T-(i\times\Delta T'_e)-((G-i+1)\times\Delta T'_d)=\Delta T'_{eb}+\Delta T'_c+\Delta T'_{db}, \tag{2}$$

where each $\Delta T'$ has the same meaning as the ΔT's in (1) but represents GOB-level, rather than frame-level, delays. Eq. (2) simply states that the i-th GOB within the frame is placed in the encoder buffer after it and the previous $i-1$ GOBs have been processed $(i\times\Delta T'_e)$ and has to be available at the decoder in time to allow its own decoding and that of that of the following $G-i$ GOBs ($(G-i+1)\times\Delta T'_d$) before the corresponding frame has to be displayed. If we use the frame as the basic unit, i.e., a GOB contains one frame (G=1), then (2) reduces to (1). Thus, in general, the delay for each GOB, $\Delta T(GOB)$, does not have to be constant. To simplify things let $\Delta T'_{e/d}=(i\times\Delta T'_e)-((G-i+1)\times\Delta T'_d)$ be constant, i.e., we let the encoding and decoding times per GOB be identical. Then each GOB will have the same delay constraint and will have to be available at the decoder within $\Delta T-\Delta T'_{e/d}$ time units of the time it was placed in the encoder buffer. While assuming a different delay constraint for each GOB within a frame can be easily accommodated within our framework, it does not add any substantial intuition to the development and results. Thus we simplify the ensuing discussion (and notation) by assuming a fixed delay requirement per GOB.

The channel delay $\Delta T'_c$ in (2) may be variable in general. For example, in transmission over shared networks, significant delay variations may occur due to queuing in the network routers. However, in this work we consider a point-to-point wireless channel connecting the base station to each of the mobile units, and thus the variation of the channel delay is comparably small. Thus we assume $\Delta T'_c$ to be constant so that, from (1), the sum of delays introduced by encoder and decoder buffers on a particular GOB $\Delta T'_{eb}+\Delta T'_{db}$ will be constant,

$$\Delta T'_{eb}+\Delta T'_{db}=\Delta T-\Delta T'_{e/d}-\Delta T'_c \tag{3}$$

Let us denote T_f the duration, in seconds, of a frame interval and let T_g be the duration of a GOB, so that $T_f = G\times T_g$, where G is the number of GOBs per frame. Then, ΔN, the total number of GOBs stored in either the encoder and the decoder buffer will be *constant* and given by,

$$\Delta N=\frac{\Delta T'_{eb}+\Delta T'_{db}}{T_g}, \tag{4}$$

where $\Delta T'_{eb}+\Delta T'_{db}$ are given by (3).

B. Rate Constraints

We will now show how these delay constraints can be translated into rate constraints that the encoder has to meet to guarantee that there are no losses due to excessive delay. The encoded video data is packetized into constant-size packets before transmission. We make the assumption that each packet interval has a duration T_p seconds. In order to facilitate our discussion, we use T_p as the basic time unit, i.e., we discretize the time by T_p. Therefore a GOB spans over F packet intervals where $F=\frac{T_g}{T_p}$. For convenience we assume that F is integer. The n-th GOB is encoded at time $n\times F$, or, equivalently, at time t (t-th packet interval) block $n=\lfloor\frac{t}{F}\rfloor$ is the last GOB that is encoded and released into the encoder buffer. Due to the delay constraint of ΔN GOBs, the n-th block has to be received by the decoder no later than $(n+\Delta N)\times F$.

Assume that at time t, block m is the GOB currently being transmitted by the channel. Define $R(i)$ as the number of bits used for encoding GOB i, and $R'(m)$ as the number of bits of GOB m that are still in the encoder buffer and waiting for transmission at time t. Therefore at time t, the content of the encoder buffer consists of $R'(m), R(m+1),\ldots,R(n)$ bits of data from GOBs $m, m+1,\ldots,n$, respectively (refer to Fig. 2).

In addition, we are modeling an ARQ based system and thus we will assume that BL (backlog) bits in the buffer are used to store packets that have been transmitted but not yet acknowledged. Assuming that the delay in receiving acknowledgments is constant and equal to b packet intervals, we will need $BL=b\times\bar{C}$ bits to store the b packets that are waiting to be acknowledged ($\bar{C}$ is the payload per packet, in bits). Since in the worst case all b packets will have to be retransmitted, we will take into account this separate ARQ buffer of size BL in deriving our rate constraints [4]. The delay in receiving acknowledgments will be $b=\left\lceil\frac{T_b}{T_p}\right\rceil$ packets, where T_b is the feedback delay in seconds.

Denote $C(k)$ as the number of bits transmitted by the channel at time k. The condition for the i-th GOB, $i\in\{m+1,\ldots,n\}$, to arrive at the decoder in time for decoding is that all the data corresponding to i, as well as to all the previous blocks in the encoder buffer, has to be transmitted by the due time $(i+\Delta N)\times F$, thus:

$$BL+R'(m)+\sum_{j=m+1}^{i}R(j)\leq\sum_{k=t+1}^{(i+\Delta N)\times F}C(k) \tag{5}$$

We assume that if part of block m has been transmitted, then the encoding rate for block m can no longer be changed. Therefore we are interested in the rate constraints for the remaining blocks (block $m+1$ to block n) in the encoder buffer. Analogous constraints apply to each of the blocks currently in the buffer. Thus, we can write for all $i=m+1,\ldots,n$ the set of constraints (6) (see also Fig. 2).

Selecting the encoding rates $R(m+1),\ldots,R(n)$ such that all the rate constraints in (6) are met guarantees that the end-to-end delay constraints will also be complied with.

[3]Note that here we are referring to the i-th GOB within a frame, while in the later discussion we will consider the overall index, i.e., the i-th GOB in the video sequence.

[4]Note that this results in a somewhat conservative estimate since in most cases not all the b packets will have to be retransmitted.

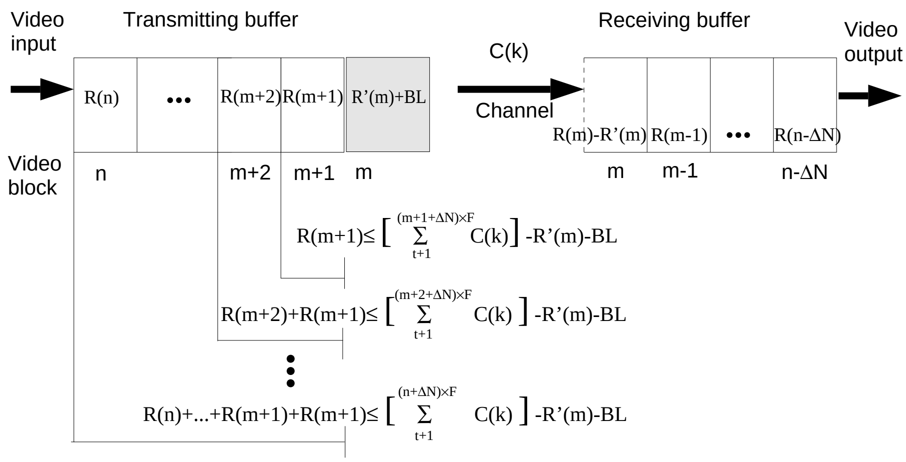

Fig. 2. Constraint on encoding rate for each buffered video block at time t.

$$
\begin{aligned}
R(m+1) &\le \left[\sum_{k=t+1}^{(m+1+\Delta N)\times F} C(k)\right] - R'(m) - BL \\
R(m+2) + R(m+1) &\le \left[\sum_{k=t+1}^{(m+2+\Delta N)\times F} C(k)\right] - R'(m) - BL \\
&\vdots \\
R(n) + \ldots + R(m+2) + R(m+1) &\le \left[\sum_{k=t+1}^{(n+\Delta N)\times F} C(k)\right] - R'(m) - BL
\end{aligned}
\tag{6}
$$

Q-GOB:	Quantization group of blocks,
P-GOB:	Processing group of blocks,
ΔT (sec):	End-to-end delay for each frame,
$\Delta T'_{e/d}$ (sec):	Encoder/decoder delay for each GOB,
T_f (sec):	Frame interval,
T_g (sec):	GOB interval,
T_p (sec):	Packet interval,
ΔN (GOB):	Number of GOBs in the encoder and decoder buffers,
G (GOB):	Number of GOBs/frame,
F (packets):	Number of packets/GOB.

TABLE I
Summary of notations

III. Formulation of the Rate Control Problem

From (6) we can observe that the general delay constraint (each GOB has to arrive at the decoder within $\Delta N \times T_g$ seconds) can be translated into rate constraints which depend on the future channel transmission rates $C(k)$, $k > t$. If the $C(k)$ were known, there would be many choices of encoding rates $R(m+1), \ldots, R(n)$ which would meet the constraints of (6). Thus we adopt a rate-distortion approach where our goal is to obtain the encoder rate allocation which produces *the minimum distortion at the decoder for the given rate constraints.*

In (6) GOBs $m+1$ to n are buffered in the encoder buffer at time t. We will assume that the quantizer assignment of a given block can be modified while the block is in the buffer, but before it starts to be transmitted (for example at time t we cannot modify the allocation for GOB m, but the rates for blocks $m+1$ to n can be adjusted.) In a DCT-based video compression scheme, a possible implementation of the system (see Fig. 3) would consist of having several parallel buffers, each storing every GOB quantized with one particular quantizer. Then the data for the block currently being transmitted (m in our example) will be drawn from the appropriate buffer. An alternative, and more elegant, approach would be to have each block encoded as an embedded bitstream (i.e., the bits corresponding to a low rate version of the GOB are embedded in all the higher rate versions).

As indicated earlier, we choose to use identical P-GOBs and Q-GOBs sizes. Thus we assume that each GOB is encoded with a quantizer from a finite quantizer set $\mathcal{Q}$. Denote $x(i) \in \mathcal{Q}$ as the choice of quantizer for block i, and $R_{x(i)}(i)$ and $D_{x(i)}(i)$ as the associated encoding rate and distortion. Consider the example depicted in Fig. 2 in which the quantizers for blocks $m+1$ to block n are dynamically selected under the rate constraints of (6). The choices

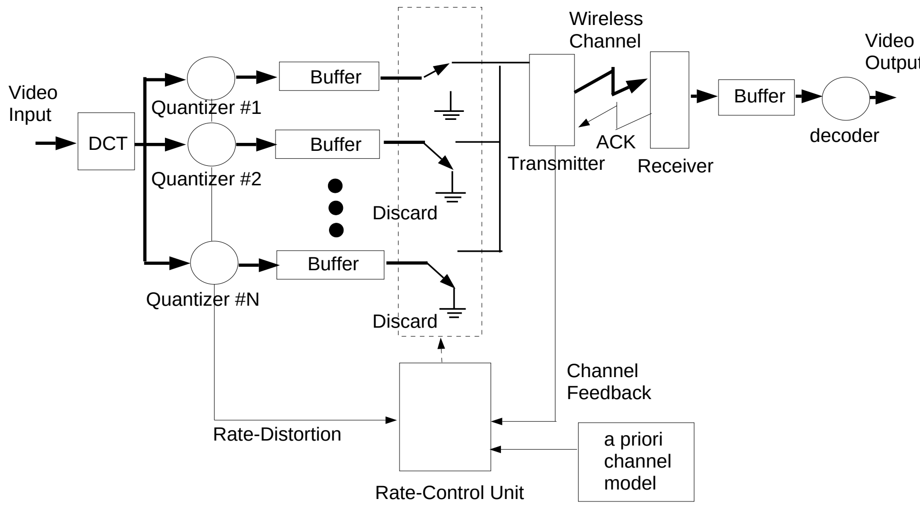

Fig. 3. System block diagram.

of quantizers $\boldsymbol{x}(m+1,n) = \{x(m+1), x(m+2), \ldots, x(n)\}$ which can result in the minimal video distortion for the video segment in the encoder buffer are those solving the following problem:

Formulation 1: Optimal quantizer selection for known channel rates
Find the optimal quantizer choices $\boldsymbol{x}^*(m+1,n)$ *at time* t *such that,*

$$\boldsymbol{x}^*(m+1,n) = \arg \min_{\boldsymbol{x}(m+1,n)} \sum_{j=m+1}^{n} D_{x(j)}(j), \tag{7}$$
$$\textit{where} \quad n = \left\lfloor \frac{t}{F} \right\rfloor,$$

subject to the constraint set:

$$\sum_{j=m+1}^{i} R_{x(j)}(j) \leq \left(\sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \right) - R'(m) - BL, \tag{8}$$
$$\forall i = m+1, \ldots, n.$$

Note that in the above formulation, the optimization is based on the R-D data of video frames currently stored in the encoder buffer, since under our assumption of real-time encoding we do not have access to the R-D data of future frames. Thus, successively finding the optimal solution to Formulation 1 for each n is equivalent to using a "sliding window" mechanism to compute the quantizer assignment for the complete sequence. Thus the quantizer for block $m+1$ is selected based on blocks $m+1$ through n and the corresponding channel rates. This is a greedy approach which cannot guarantee overall optimality, but it is the best approach given that only those blocks are known at the encoder.

We are considering a lossy transmission channel, where packets are retransmitted if they are received with too many errors to be decoded correctly. Thus, in our system $C(k)$ can be either the nominal packet payload $\bar{C}$ (if the packet is received correctly) or zero (if an error occurred). Therefore we cannot guarantee that the rate constraints of (6) will not be violated, and thus that no losses will occur, because this would require knowledge of the future channel transmission rates $C(k)$ for packets k $(k > t)$. In this paper we show how to make use of a probabilistic model of the channel and observations of the current channel state in the context of this rate control problem.

We propose two alternative formulations, which both assume that, given the observation and the *a priori* model, estimates of future channel behavior can be obtained. Our first approach consists of modifying Formulation 1 so as to use *expected rate constraints*, given the current state of the buffer. In the second approach, we instead minimize the *expected distortion*, where the distortion of a given block depends not only on the choice of quantizer but on the probability that the block is lost. In what follows we will denote $S(t-b)$ the latest *observation* of the channel state at time t, where b is the delay (in number of packets) with which we obtain feedback information. For example, if $b = 0$ the decoder would know immediately whether transmission in the prior time slot was successful. Typically we will have $b > 0$ since the encoder has to wait for acknowledgments from the decoder to be received in order to determine whether transmission was successful.

In the first approach, assuming that the encoder can estimate the expected value of the future channel rates (as will be discussed in Section IV-C), we can replace the rate constraints in (8) by their expected values:

$$\sum_{j=m+1}^{i} R(j)$$
$$\leq E\left[\sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \;\middle|\; S(t-b) \right] - R'(m) - BL, \tag{9}$$
$$\forall i = m+1, \ldots, n,$$

so that our problem can be formulated as:

Formulation 2: Rate control under estimated rate constraints

Find the optimal quantizer choices $\boldsymbol{x}^*(m+1,n)$ *at time* t *such that,*

$$\boldsymbol{x}^*(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \sum_{j=m+1}^{n} D_{x(j)}(j), \qquad (10)$$

$$\textit{where} \quad n = \left\lfloor \frac{t}{F} \right\rfloor,$$

subject to the expected rate constraints (9).

In the above formulation data loss caused by exceeding the delay constraints may still happen even if the encoding rates $R(m+1), \ldots, R(n)$ meet all the expected constraints of (9), because the actual channel rates may be lower than our predicted values. Because data loss may result in significant distortion in the decoded video, it may be better to replace the average expected rate in (9) by, say, the future rates which are guaranteed with probability 90%. This will obviously result in a more conservative rate allocation at the encoder and hence higher distortion at the decoder. This would be a form of effectively trading off the source rate distortion (sending fewer bits) for the distortion due to losses (if fewer bits are sent they are more likely to be received correctly). This trade-off can be made explicit if we assume that the distortion incurred by data loss can be estimated.

Thus, our second rate control approach will seek to minimize the "expected" distortion, which combines the distortion caused by encoding and that caused by data loss. More specifically, denote $D_0(i)$ as the incurred distortion on video block i, when the block is lost. Let $p_{loss}(i)$ be the probability that video block i does not arrive at the decoder in time. This will happen if the rate constraints corresponding to block i are violated and thus we can write

$$p_{loss}(i) = \Pr\Bigg[\sum_{j=m+1}^{i} R(j) > \left(\sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \right) - R'(m) - BL \,\Bigg|\, S(t-b) \Bigg]. \qquad (11)$$

Then the expected distortion of block i can be defined as:

$$E\left[D(i)|S(t-b)\right] = (1 - p_{loss}(i)) \times D(i) + p_{loss}(i) \times D_0(i), \qquad (12)$$

and the problem can be reformulated as:

Formulation 3: Rate Control for Minimum Expected Distortion
Find the optimal quantizer choices $\boldsymbol{x}^*(m+1,n)$ *at time* t *such that,*

$$\boldsymbol{x}^*(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \sum_{j=m+1}^{n} E\left[D_{x(j)}(j)|S(t-b)\right] \quad (13)$$

IV. Probabilistic Modeling of Channel Behavior

The formulations we propose are very general and do not rely on any specific characteristics of the statistical channel behavior. However, the available solutions may differ substantially depending on the specific channel characteristics. Indeed, for channels with random (rather than bursty) errors, the proposed real time feedback approach may not provide any gains in performance, as compared to a open loop FEC approach. We now present specific parameters and models for the burst error channels that will be used in our experiments. While the optimization techniques to be presented later have general applicability, we focus our discussion on the case of burst-error channels. As pointed out in the introduction, our goal is not to advocate the use of specific models, but instead to show how channel models can be incorporated into a rate control mechanism: for the same application, similar results could be achieved with alternative models.

A. Physical Layer and Error Control

The specific channel under consideration is a wireless CDMA spread spectrum system [28] for a mobile transmission environment [6], where channel errors tend to occur in burst during channel fading periods. The wireless channel consists of two radio links, namely uplink (mobile-to-base) and downlink (base-to-mobile). The encoded video bitstream is packetized into constant-size packets for transmission. Note that one could use an interleaved FEC scheme, such as that defined in ITU-T recommendation I.363 [29]. However the robustness in such an FEC scheme would come at the price of additional delay, since the degree of interleaving may have to be significant. Instead, here we choose to use shorter interleaving periods (see transceiver description in [30], [1]) combined with an ARQ approach. We then use models for the resulting error probability for the data packets (i.e., after processing).

In our Selective Repeat (SR) ARQ scheme, the reception of a packet is acknowledged by the receiver by sending either an acknowledgment (ACK) or a negative acknowledgment (NAK) to the transmitter. Only the erroneous packets are retransmitted. A time-out mechanism is used so that, if the feedback information is corrupted, data is retransmitted anyway. Packets that have been sent are stored in the ARQ buffer until they are acknowledged. Packets awaiting transmission are stored in the encoder buffer (see Fig.1) and the decoder buffer can be used to rearrange the received packets, which may be out-of-order due to retransmission. Refer to Fig. 4 for a diagram of the various buffers in the system. Because video transmission is subject to a delay constraint as discussed in Section II, the retransmission of any packet is attempted only while its due time has not been exceeded. Data losses occur during the channel fading intervals whenever the data cannot be retransmitted before its due time.

B. Channel models

Previous studies [31], [32] show that a first-order Markov chain, such as a two-state Markov model [33], [34] or a finite-state model [35], [30], [1] provide a good approximation in modeling the error process at the packet level in fading channels. Here we use a *two-state Markov model* and an *N-state Markov model* to emulate the process of packet errors. Note that the transition probabilities of the two

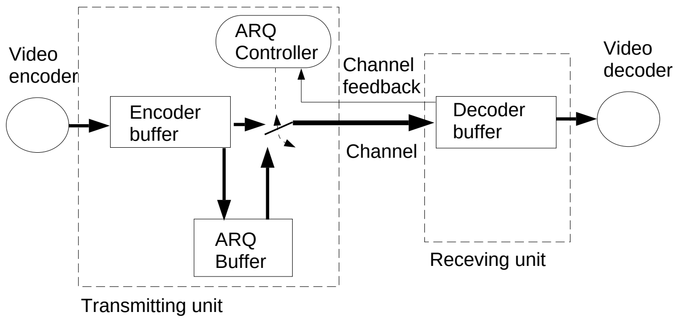

Fig. 4. Diagram of buffers in the system

models are chosen such as to have the same overall probability of error, although the average burst lengths will be different.

B.1 Two-state Markov model:

In this model, the channel switches between a "good state" and a "bad state", s_0 and s_1, respectively: packets are transmitted correctly when the channel is in state s_0, and errors occur when the channel is in state s_1 [5]. p_{ij} for $i,j \in \{0,1\}$ are the transition probabilities (see Fig. 5). The transition probability matrix for this channel model

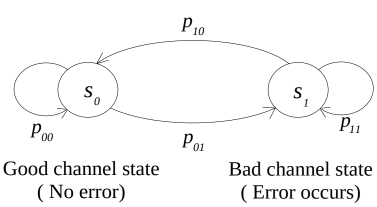

Fig. 5. Two state Markov channel model.

then can be set up as:

$$\boldsymbol{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \tag{14}$$

B.2 N-state Markov model:

We use a simplified version of the more general finite-state Markov model described in [35]. In this N-state model, introduced in [30], [1] (see Fig. 6), the channel states are defined as s_n, $n = 0, \ldots, N-1$ in which s_0 represents the "good state" and all other states represent the "bad states". When the channel is in state s_n, $n \in \{0, \ldots, N-2\}$, the transition of the channel state is either to the next higher state or back to state s_0 based on the status of the currently received data packet. If the channel is in state s_{N-1}, it will always return to state s_0. With this model, it is only possible to generate burst errors of at most length $N-1$ (see Fig. 6 and Table II).

Define $p_n = Prob(s_{n+1}|s_n)$ as the transitional probability from state s_n to s_{n+1}. The transition probability matrix for this N-state Markov chain model can be set up as:

$$\boldsymbol{P} = \begin{bmatrix} 1-p_0 & p_0 & 0 & 0 & \cdots & 0 \\ 1-p_1 & 0 & p_1 & 0 & \cdots & 0 \\ 1-p_2 & 0 & 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 1-p_{N-2} & 0 & 0 & 0 & \cdots & p_{N-2} \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{15}$$

The state transition probability for the uplink and downlink channels at BER $= 10^{-3}$ are shown in the following table, where N was found to be 15 and 6 (equivalent to maximum burst error lengths of 70 msec and 25 msec) for the downlink and the uplink channel, respectively. These values are found by matching the parameters of the Markov chains to simulations of the transceivers [30], [1].

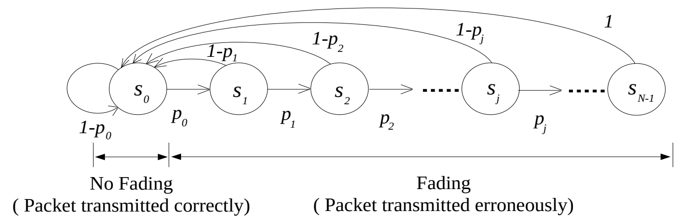

Fig. 6. N-state Markov model

	Downlink	Uplink
p_0	0.001469	0.064292
p_1	0.516068	0.100324
p_2	0.778388	0.164083
p_3	0.854118	0.149606
p_4	0.936639	0.526316
p_5	0.873529	0.000000
p_6	0.905724	
p_7	0.881041	
p_8	0.831224	
p_9	0.893401	
p_{10}	0.863636	
p_{11}	0.717105	
p_{12}	0.853211	
p_{13}	0.763441	
p_{14}	0.000000	

TABLE II
Transitional probability of the downlink and the uplink channels for a 15-state Markov model

C. Channel Rate Estimation

Assuming that at time t the channel state at time $t-b$, $S(t-b)$, is known, the average channel rates in Formulation 2 can be derived from the channel models. In this section we focus on the estimation for the N-state model, but similar approaches can be applied to the two-state model as well (refer to [36] for a more detailed derivation for the two-state model case).

[5] More general classes of two-state Markov models can also be used, where for example each state in the model has associated a different probability of error.

In this N-state Markov channel model with transition probabilities (15), define the state probability $\pi_n(k \mid S(t-b))$ as the probability that the channel is in state s_n at time k given the channel state observation $S(t-b)$. A vector of state probabilities can be written as:

$$\boldsymbol{\pi}(k \mid S(t-b)) = [\pi_0(k|S(t-b)), \pi_1(k|S(t-b)), ..., \pi_{N-1}(k|S(t-b))].$$

The initial state probability $\boldsymbol{\pi}(t-b \mid S(t-b))$ at time $t-b$ can be set up as:

$$\forall\, n \in \{0, \dots, N-1\}, \quad \pi_n(t-b \mid S(t-b)) = \begin{cases} 1, & \text{when } S(t-b) = s_n; \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

In the Markov model, the state probabilities $\boldsymbol{\pi}(k \mid S(t-b))$ at time k can be derived from the state probabilities $\boldsymbol{\pi}(k-1 \mid S(t-b))$ at the previous time slot and the transition probability matrix $\boldsymbol{P}$ as:

$$\boldsymbol{\pi}(k \mid S(t-b)) = \boldsymbol{\pi}(k-1 \mid S(t-b)) \cdot \boldsymbol{P} \tag{17}$$

By recursively using (17), channel state probabilities at time k, where $k > t-b$, can then be calculated from $\boldsymbol{\pi}(t \mid S(t-b))$ and $\boldsymbol{P}$ as:

$$\boldsymbol{\pi}(k \mid S(t-b)) = \boldsymbol{\pi}(t-b \mid S(t-b)) \cdot \boldsymbol{P}^{k-t+b} \tag{18}$$

In our channel model, packets are transmitted correctly ($\bar{C}$ bits are transmitted) when the channel is in state s_0, while errors occur (0 bits are transmitted) when the channel is in any other state s_i, $i \in \{1, \dots, N-1\}$. Therefore $\pi_0(k)$ is the probability of correct transmission at time k. The expected channel rate $E[C(k)|S(t-b)]$ given the observation of channel state $S(t-b)$ can be calculated as:

$$E[C(k)|S(t-b)] = \bar{C} \times \pi_0(k \mid S(t-b)) \tag{19}$$

and thus the sum of expected channel rates in (9) can be written as:

$$E\left[\sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \,\middle|\, S(t-b)\right] = \sum_{k=t+1}^{(i+\Delta N)\times F} E[C(k)|S(t-b)] = \bar{C} \cdot \sum_{k=t+1}^{(i+\Delta N)\times F} \pi_0(k|S(t-b))$$

D. Expected Distortion

From (11), the probability of losing the i-th GOB, $p_{loss}(i)$, depends on the accumulated rate in the encoder buffer $B = \sum_{j=m+1}^{i} R(j)$ and on the future channel rates. Therefore, given the channel state observation $S(t-b)$ at time $t-b$, we define a probability distribution function $\Phi_{i,t}(B \mid S(t-b))$, with the accumulated encoding rate B as variable, as:

$$\Phi_{i,t}(B \mid S(t-b)) = \Pr\left[B > \left(\sum_{k=t+1}^{(i+\Delta N)\times F} C(k)\right) - R'(m) - BL \,\middle|\, S(t-b)\right] \tag{20}$$

Given that the quantizer choices are $x(m+1), \dots, x(i)$, the encoder can estimate the distortion of i-th GOB given $B = R_{x(m+1)}(m+1) + \dots R_{x(i)}(i)$ as:

$$\begin{aligned} E[D(i)] &= [1 - \Phi_{i,t}(B \mid S(t-b))] \times D_{x(i)}(i) \\ &\quad + \Phi_{i,t}(B \mid S(t-b)) \times D_0(i), \end{aligned} \tag{21}$$

where $D_{x(i)}(i)$ is the distortion for the GOB when quantizer $x(i)$ is used, and $D_0(i)$ is the distortion incurred when the whole GOB is lost (we assume $D_0(i)$ is the distortion incurred when replacing each block in the GOB by a block with constant intensity equal to the average intensity of the original block.)

The probability distribution function $\Phi_{i,t}(B \mid S(t-b))$ at time t can also be derived from the channel model given the channel observation $S(t-b)$. For any given value B, define η as the number of packets needed for transmitting those B bits of encoded data as:

$$\eta = \left\lceil \frac{B}{\bar{C}} \right\rceil \tag{22}$$

where $\bar{C}$ is the packet size. Then $\Phi_{i,t}(B \mid S(t-b))$ is equivalent to the probability that less than η packets are successfully transmitted during the time interval $[t+1, \dots, (i+\Delta N) \times F]$.

To calculate $\Phi_{i,t}(B \mid S(t-b))$ for an N-state channel model, define $q_{n,r}(t,k)$, where $n \in \{0, \dots, N-1\}$ and $r \le k-t$, as the probability that, initialized from time t, the channel visits state s_0 (successful packet transmission) r times and arrives at state s_n at time k. Given the observed channel state $S(t-b)$ at time $t-b$, the initial state probabilities when $k = t$ (i.e., $q_{n,r}(t,t)$) can be set up as:

$$\forall n \in \{0, \dots, N-1\}, \quad q_{n,r}(t,t) = \begin{cases} \pi_n(t \mid S(t-b)), & \text{when } r = 0; \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The value of $q_{n,r}(t,k)$ can be obtained recursively from the Markov chain model as:

$$q_{n,r}(t,k) = \begin{cases} \sum_{n=0}^{N-1} (1-p_n) \cdot q_{n,r-1}(t,k-1), & \text{when } n = 0; \\ p_{n-1} \cdot q_{n-1,r}(t,k-1), & \text{when } n = 1, \dots N-1. \end{cases}$$

Therefore the value of the probability distribution function $\Phi_{i,t}(B \mid S(t-b))$ for a given value B is

$$\Phi_{i,t}(B \mid S(t-b)) = \sum_{r=0}^{\eta-1} \sum_{n=0}^{N-1} q_{n,r}\left(t, (i+\Delta N) \times F\right), \quad \text{where } \eta = \lceil \tfrac{B}{\bar{C}} \rceil \tag{24}$$

Note that since the number of channel states is discrete and the transition probabilities are known a priori, it is possible to use tables to generate the relevant constraints and probabilities from the channel observations.

E. *Effectiveness of Rate Control Based on Channel Model and Feedback*

In the two proposed rate control approaches, i.e., rate control under estimated rate constraints (Formulation 2) and rate control for maximum expected distortion (Formulation 3), the estimation of the expected rate constraints (19) and expected distortion (21) relays on the *a priori* probabilistic model of the channel. However, if there is a mismatch between the channel model used in the rate control and the actual channel behavior, the performance of the rate control will obviously suffer accordingly.

Another factor that may affect the performance of rate control is the feedback delay b. We can expect performance to be good if feedback is available at the encoder immediately. Conversely, as b grows the information available at the encoder becomes increasingly inaccurate (we are using an "old" observation of the channel state), and therefore the efficacy of the rate control will be reduced.

In Section VI we will provide some experiments to assess the effect of channel modeling mismatches and feedback in the performance of our algorithms.

V. Rate Control Algorithms

A. *Encoding Rate Selection under Estimated Rate Constraints*

The problem of Formulation 2 can be solved at every time instant using the estimated rate constraints from the selected channel model. Here we propose solutions based on dynamic programming and Lagrangian optimization.

A.1 Encoding Rate Selection by Dynamic Programming

We consider a scheme which encodes all video frames of a video sequence in intra-frame mode. In this case the resulting encoding rate and distortion for each block depends only on the quantizer selection for the block and not on quantizer selections for other blocks. Because the quantizers are selected from a finite set $\mathcal{Q}$, the optimal choices of quantizers that can achieve minimum distortion can be searched using the dynamic programming technique. A formulation for video rate control based on dynamic programming was introduced in [15], while dependent quantization was considered in [16]. Even though most video coding schemes are dependent we focus here in the independent case. In lossy environments, intra-frame schemes may actually be preferable (since there will be no error propagation in case of errors). In addition, it is possible to develop approximate solutions for the dependent cases with simple modifications of the proposed algorithms (by making them iterative) as was done in [21].

Consider Fig. 7. Our goal in order to solve Formulation 2 is to find the best quantizer choice for blocks $m+1$ through n (all those currently in the buffer) so that none of the constraints of (9) are violated. The y-axis in Fig. 7 represents the accumulated rate (or state) and the x-axis represents the GOB considered (or stage). The initial state (stage $i = m$) represents the initial contents of the buffer, $R'(m) + BL$. Each branch in the trellis represents a choice of quantizer, thus a branch linking stages m and $m+1$ represents a choice of quantizer for block $m+1$, and the y-coordinate of the branch's end represents the accumulated rate. Each branch at a given stage has an associated distortion which is the distortion of the block at that stage when using the quantizer corresponding to the branch. A path consists of a sequence of branches and has associated a total path cost which is equal to the sum of the distortions of the branches in the path.

More formally, given that the quantizer choices $\boldsymbol{x}(m+1,i) = \{x(m+1),\ldots,x(i)\}$ are used for encoding blocks $m+1$ to block i, we define $B_{\boldsymbol{x}(m+1,i)}$ as the accumulated encoding rates for these blocks as:

$$B_{\boldsymbol{x}(m+1,i)} = \sum_{j=m+1}^{i} R_{x(i)}(i) \tag{25}$$

At each stage of the trellis, each state represents a possible level of accumulated rate given that a particular selection of quantizers $\boldsymbol{x}(m+1,i) = \{x(m+1),\ldots,x(i)\}$ is used. We define $\mathcal{S}_i(B)$ as the state in stage i when an accumulated encoding rate B is used. State $\mathcal{S}_i(B)$ is associated with the accumulated distortion $\sum_{j=m+1}^{i} D_{x(j)}(j)$ as the cost for that state. Because of the rate constraints (9), only the states with state variable B that meet the rate constraints (9) are valid. That is, state $\mathcal{S}_i(B)$ is valid if:

$$B \le E\left[\sum_{k=t+1}^{i+\Delta N} C(k)\right] - R'(m) - BL \tag{26}$$

Suppose a set of quantizer choices $\boldsymbol{x}(m+1,i) = \{x(m+1),\ldots,x(i)\}$ results in an accumulated encoding rate $B_{\boldsymbol{x}(m+1,i)}$. Given that a choice of quantizer $x(i+1)$ is used to encode block $i+1$ and results in the encoding rate $R_{x(i+1)}(i+1)$ and distortion $D_{x(i+1)}(i+1)$, the resulting accumulated rate $B_{\boldsymbol{x}(m+1,i+1)}$ is then:

$$B_{\boldsymbol{x}(m+1,i+1)} = B_{\boldsymbol{x}(m+1,i)} + R_{x(i+1)}(i+1) \tag{27}$$

and arrives at the state $\mathcal{S}_{i+1}(B_{\boldsymbol{x}(m+1,i+1)})$. Such choice of quantizer is represented by a branch that connects the node of state $\mathcal{S}_i(B_{\boldsymbol{x}(m+1,i)})$ at stage i to the node of state $\mathcal{S}_{i+1}(B_{\boldsymbol{x}(m+1,i+1)})$ with cost $\sum_{j=m+1}^{i+1} D_{x(j)}(j)$ at stage $i+1$. Therefore in the trellis representation, a path that consists of branches that are connected from stage $m+1$ to stage $i+1$ represents a set of quantizer choices $\boldsymbol{x}(m+1,i+1) = \{x(m+1),\ldots,x(i+1)\}$.

With (26) we test whether the quantizer choice $x(i+1)$ may cause the violation of the rate constraint or not, and any branches that violate the rate constraint are pruned out. If two or more sets of quantizer choices result in the same accumulated encoding rate B and arrive at the same state $S_{i+1}(B)$ at stage $i+1$, only the path that results in the minimum accumulated distortion at the given node is kept and all the other sub-optimal paths are pruned out. This is based on Bellman's optimality principle [37] and it is easy to see that (because rate and distortion are decoupled for each block) it also applies in this case. Therefore the cost

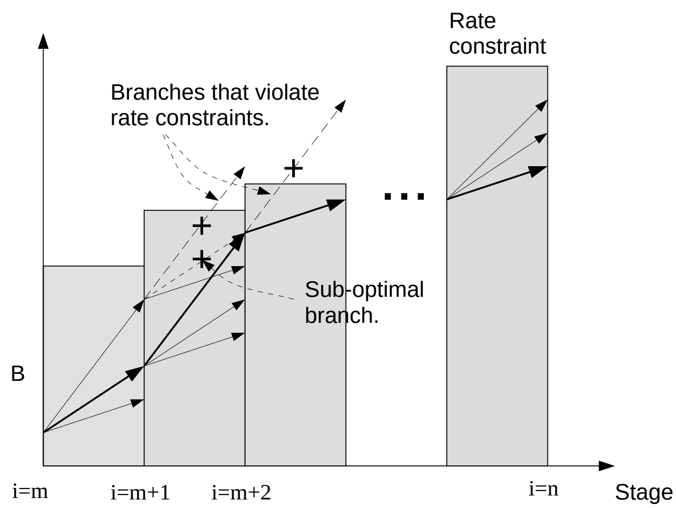

Fig. 7. Trellis tree in dynamic programming for searching the optimal encoding rate allocation.

associated with state $S_{i+1}(B_{\boldsymbol{x}(m+1,i+1)})$ is the minimum among those resulting in the same accumulated rate.

By pruning sub-optimal solutions at every intermediate stage, the combination of quantizer choices that can achieve minimum distortion can thus be found without trying all the possible combinations. However the complexity can still be fairly high depending on the number of stages and the number of states per stage. This prompts us to consider a faster optimization approach.

A.2 Encoding Rate Selection by Lagrangian Optimization

Using Lagrangian optimization for rate control under multiple rate constraints was previously studied in [38], [22]. In the Lagrangian optimization approach, the constrained optimization problem in Formulation 2 is equivalent to the unconstrained problem derived by introducing a non-negative Lagrange multiplier λ_i associated with each rate constraint in (9). The optimization formulation then becomes:

Formulation 4: Find the quantizer choice $\boldsymbol{x}^*(m+1,n)$ *at time t such that*

$$\boldsymbol{x}^*(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \sum_{j=m+1}^{n} D_{x(j)}(j) + \sum_{i=m+1}^{n} \lambda_i \cdot \left(\sum_{j=m+1}^{i} R_{x(j)}(j) \right) \quad (28)$$

where we introduce $n-m$ Lagrange multipliers to replace the $n-m$ constraints (9). Then the problem that remains is to find the appropriate multipliers $\lambda_{m+1},\ldots,\lambda_n$ such that no constraint is violated. Define λ'_j as:

$$\lambda'_j = \sum_{i=j}^{n} \lambda_i, \quad \forall j \in \{m+1,\ldots,n\}. \quad (29)$$

then (28) can be rearranged as:

$$\boldsymbol{x}^*(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \sum_{j=m+1}^{n} \left(D_{x(j)}(j) + \lambda'_j \cdot R_{x(j)}(j) \right) \quad (30)$$

Since $\lambda_{m+1},\ldots,\lambda_n$ are all non-negative values, from (29) we have

$$\lambda'_{m+1} \geq \lambda'_{m+2} \geq \ldots \geq \lambda'_n. \quad (31)$$

Because the mapping $\{\lambda_{m+1},\ldots,\lambda_n\} \rightarrow \{\lambda'_{m+1}\ldots,\lambda'_n\}$ is one-to-one, it is equivalent to finding the appropriate non-negative values of $\{\lambda'_{m+1},\ldots,\lambda'_n\}$ such that no constraint is violated. Define $J_i(\lambda'_i, x(i))$, the cost for block i, as:

$$J_i(\lambda'_i, x(i)) = D_{x(i)}(i) + \lambda'_i \cdot R_{x(i)}(i), \quad \forall i \in \{m+1,\ldots,n\}. \quad (32)$$

If we use intra-frame mode, the quantizer for each video block can be independently chosen by minimizing the cost for each block $J_i(\lambda'_i, x(i))$ as:

$$x^*(i) = \arg\min_{x(i)\in\mathcal{Q}} J_i(\lambda'_i, x(i)), \quad \forall i \in \{m+1,\ldots,n\}. \quad (33)$$

Then the problem remains of how to determine a set of Lagrange multipliers $\{\lambda'_{m+1},\ldots,\lambda'_n\}$ such that the rate constraints are met. In [38] a similar problem is solved by iteratively increasing the lower bounds on the multipliers, defined as $\{\Lambda'_{m+1},\ldots,\Lambda'_n\}$, such that the violation of rate constraints can be prevented, and adjusting the values of $\{\lambda'_{m+1},\ldots,\lambda'_n\}$ until an optimal bit allocation, where none of the constraints is violated, is found.

Initially the quantizer choices $\hat{\boldsymbol{x}}(m+1,n)$ are selected by Lagrangian optimization subject to only a constraint on the total rate for all blocks in the buffer: $\sum_{j=m+1}^{n} R(j) \leq E\left[\sum_{k=t+1}^{(n+\Delta N)\times F} C(k)|S(t-b)\right] - R'(m) - BL$. Only one multiplier λ_n is associated with the constraint as:

$$\hat{\boldsymbol{x}}(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \left\{ \sum_{j=m+1}^{n} D_{x(j)}(j) + \lambda_n \cdot \left(\sum_{j=m+1}^{n} R_{x(j)}(j) \right) \right\} \quad (34)$$

From (29), this is equivalent to setting $\lambda'_{m+1} = \lambda'_{m+2} = \ldots = \lambda'_n = \lambda_n$, and the optimization can be solved by minimizing the cost for each block individually as in (33) using a single Lagrange multiplier λ'_n in every cost function as:

$$\hat{x}(i) = \arg\min_{x(i)\in\mathcal{Q}} J_i(\lambda'_n, x(i)), \quad \forall\, i \in \{m+1,\ldots,n\}. \quad (35)$$

The optimal quantizer choices and the appropriate value of λ'_n can be found simultaneously by the bisection search technique. If the quantizer choices $\hat{\boldsymbol{x}}(m+1,m)$ that are selected do not cause violation of the other rate constraints, then the quantizer choices $\hat{\boldsymbol{x}}(m+1,n)$ are the solution to the optimization problem of Formulation 4. Otherwise, if the resulting encoding rates violate any other rate constraints, then the quantizer choices $\hat{\boldsymbol{x}}(m+1,n)$ are not the desired solution and other rate constraints also have to be taken into account in the optimization process by including more Lagrange multipliers.

Assuming that block v, where $v < N$, is the "last" block which violates the rate constraints given that the quantizer choices $\hat{\boldsymbol{x}}(m+1,n)$ are used, i.e.,

$$\sum_{j=m+1}^{v} R_{x(j)}(j) > E\left[\sum_{k=t+1}^{(v+\Delta N)\times F} C(k)\right] - R'(m) - BL, \quad (36)$$

and there is no other rate constraint violation for the video segment from block $v+1$ to block n, then the encoding rates for the video segments from block $m+1$ to block v have to be reduced. Another Lagrange multiplier λ_v has to be included in the optimization process to take into account the constraint as:

$$\hat{\boldsymbol{x}}(m+1,n) = \arg\min_{\boldsymbol{x}(m+1,n)} \left\{ \sum_{j=m+1}^{n} D_{x(j)}(j) + \lambda_v \cdot \left(\sum_{j=m+1}^{v} R_{x(j)}(j)\right) + \lambda_n \cdot \left(\sum_{j=m+1}^{n} R_{x(j)}(j)\right)\right\} \quad (37)$$

Again from (29) it is equivalent to setting λ'_i as:

$$\lambda'_i = \begin{cases} \lambda'_v = \lambda_v + \lambda_n, & \text{when } m+1 \le i \le v; \\ \lambda'_n = \lambda_n, & \text{when } v+1 \le i \le n. \end{cases}$$

and the optimization problem (37) can be rewritten as:

$$\hat{x}(i) = \begin{cases} \arg\min_{x(i)\in\mathcal{Q}} J_i(\lambda'_v, x(i)), & \text{when } m+1 \le i \le v; \\ \arg\min_{x(i)\in\mathcal{Q}} J_i(\lambda'_n, x(i)), & \text{when } v+1 \le i \le n; \end{cases} \quad (38)$$
$$\text{where } \lambda'_v = \lambda_v + \lambda_n, \quad \lambda'_n = \lambda_n.$$

The overall encoding rates for the video segments from block $m+1$ to block v can be reduced because the multipliers λ'_i, $i \in \{m+1,\ldots,v\}$, in the cost functions of block $m+1$ to block v are increased by a non-negative value λ_v. Define Λ'_v as the value of λ'_v when the rate constraint (36) can just be avoided. Then Λ'_v is the lower bound of λ'_v and all the applicable values of λ'_v should be greater than Λ'_v. The lower bound Λ'_v can also be derived using a bisection search technique on the video segments from block $m+1$ to block v as:

$$\Lambda'_v = \arg\min_{\boldsymbol{x}(m+1,v)} \left\{ \sum_{j=m+1}^{v} D_{x(j)}(j) + \lambda'_v R_{x(j)}(j) \right\} \quad (39)$$

given the rate constraint:

$$\sum_{j=m+1}^{v} R_{x(j)}(j) \le E\left[\sum_{k=t+1}^{(v+\Delta N)\times F} C(k)\right] - R'(m) - BL. \quad (40)$$

The choice of quantizer is found again where the multipliers $\lambda'_{m+1},\ldots,\lambda'_v$ are lower-bounded as:

$$\hat{x}(i) = \arg\min_{\substack{x(i)\in\mathcal{Q} \\ \lambda'_i \ge \Lambda'_i}} J_i(\lambda'_i, x(i)), \quad \forall i \in \{m+1,\ldots,n\} \quad (41)$$

The search for the optimal quantizers $\boldsymbol{x}(m+1,n)$ and the appropriate multipliers $\{\lambda'_{m+1},\ldots,\lambda'_n\}$ is repeated until a choice of quantizers that does not violate any rate constraints is found. The algorithm can be summarized as follows:

Step 0 Initially the quantizer choices $\hat{\boldsymbol{x}}(m+1,n)$ are obtained by using a single Lagrange multiplier λ'_n for all blocks in (33), subject to only one constraint: $\sum_{j=m+1}^{n} R(j) \le E\left[\sum_{k=t+1}^{(n+\Delta N)\times F} C(k)\right] - R'(m) - BL$.

Step 1 If $\hat{\boldsymbol{x}}(m+1,n)$ is such that all rate constraints in (9) are met, then $\hat{\boldsymbol{x}}(m+1,n)$ is the optimal solution $\boldsymbol{x}^*(m+1,n)$ for Formulation 4. Otherwise, assume that frame v is the "last" frame which violates the rate constraint, i.e., $v < n$ and no other frame between frame $v+1$ and frame n violates the rate constraint. Find the minimum value of Lagrange multiplier Λ'_v for the video segment from frame $m+1$ to frame v which just prevents violation of the rate constraint as in (39).

Step 2 Find the quantizer choices $\hat{\boldsymbol{x}}(m+1,n) = \{\hat{x}_{m+1},\ldots,\hat{x}_n\}$ as in Step 0 except that the Lagrangian multiplier for the video segment from frame $m+1$ to frame v is lower-bounded by Λ'_v as $\lambda'_v \longleftarrow \max(\Lambda'_v, \lambda'_v)$.

Step 3 Go to Step 1. Repeat until all the rate constraints in (9) are met.

Refer [38] for a detailed description of the algorithm and the proof of optimality.

B. Encoding Rate Selection for Minimum Expected Distortion

The dynamic programming approach discussed in Section V-A.1 can be used to find the optimal quantizer choices formulated in Formulation 3. To use a dynamic programming framework to minimize the expected distortion, the cost associated with state $S_i(B)$ is the sum of the expected distortions $\sum_{j=m+1}^{i} E\left[D_{x(j)}(j) \mid S(t-b)\right]$ along the path. From (20) we can observe that the loss probability $\Phi_{i,t}(B \mid S(t-b))$ only depends on the accumulated encoding rate B. Since each state $S_i(B)$ is uniquely defined by its accumulated encoding rate B, we can associate a unique loss probability $\Phi_{i,t}(B \mid S(t-b))$ to each state, and this independently of future quantization choices. Therefore all the paths that arrive at the same state $S_i(B)$ will have the same loss probability no matter what were their previous states. Thus the optimality principle also applies in this case and paths that are sub-optimal (higher expected distortion) up to a given state are also guaranteed to be suboptimal overall. We can solve the problem using dynamic programming as described before, with the only modification being that the branch cost is now the expected distortion, rather than the deterministic distortion due to coding as in Section V-A.1.

However, the Lagrangian optimization approach can not be used because the choice of quantizers for other video blocks can affect the value of expected of distortion $E\left[D_{x(j)}(j)\right]$. To be more specific, if the problem is formulated as that of finding the quantizer choice $x(i)$

to minimize the cost function $J_i(\lambda_i', x(i)) = E[D_{x(i)}(i)] + \lambda_i' R_{x(i)}(i)$, the choice of previous encoding rate $R(m+1), \ldots, R(i-1)$ may affect the value of $E[D_{x(i)}(i)]$ (since it determines B) and thus the optimization can not be achieved independently for each block as in (28).

VI. Experimental Results and Conclusions

A. Performance of the Proposed Algorithms

In order to assess the effectiveness of the proposed rate-control algorithms, we implement them with simulated channel behavior based on the models we described for downlink and uplink channels in Section IV-B. In each experiment we use channel models to generate error patterns, and the results we provide are averaged over several realizations of the channel error patterns. In each case the encoder has knowledge of the statistical model of the channel behavior and makes use of it in the rate control algorithm. However in some of our experiments we will consider that there exists a mismatch between the actual behavior and that assumed by the rate control algorithm. Table III summarizes the characteristics of the various models used in our simulations. Note that average losses for the downlink channels are smaller but the corresponding burst durations are also longer. The uplink channel behavior we simulate with the 2-state Markov model is very close to being a channel with uniformly distributed losses, since the average burst length is close to 1.

The video test sequence "Susie" (first 100 frames) is used in our experiments. The input sequence is in QCIF format (176 × 144 pixels for each frame), and is encoded using an H.261 encoder [39], [40] with the quantization step size chosen from four values: 12, 14, 20 and 30. The H.261 encoder is used in the intra mode, which allows us to allocate quantizers independently to each frame. In the QCIF format, each frame is subdivided into macroblocks (MB) with size 16 × 16 pixels. Therefore each frame consists of 99 (11 × 9) macroblocks. In our simulation, we select the frame rate such that the duration of 3 MB's equals to one packet transmitting interval. Every three MB's are grouped together as a Q-GOB with a single quantizer being assigned to each video block. A packet is transmitted by the channel every 5 msec with 41 bytes payload, thus video data is transmitted at the rate about 6 frames/sec on average. In our simulations, except in those that consider explicitly the effect of feedback delay (Figs. 12 and 13), we assume that $b = 2$, that is, the state of the channel is known with a delay of two packet intervals.

Our results are summarized in Figs. 8, 9 for the N-state Markov model and Figs. 10, 11 for the two state Markov model. We provide results of both PSNR[6] and packet loss rates. The packet losses we plot are the losses due to video information not arriving in time at the decoder. That is, every time a packet cannot be decoded due to excessive bit errors we have the chance to retransmit it (or retransmit the same video information coded at a lower rate); if the retransmission is successful, then we do not count the packet loss. Thus we only count the packet losses observed by the decoder, i.e., the video information missing at the time it has to be decoded even after retransmission has been attempted. Note that because we are subject to a delay constraint we cannot guarantee that losses will not occur: for example if a packet has to be retransmitted and it arrives to the decoder too late to meet its delay constraint it will be considered as a lost packet (even though physically the packet did arrive to the decoder; it just arrived too late) and thus our results will show non-zero packet loss rates. Also note that our algorithm does not reduce the "raw" channel losses, but it helps reduce the resulting "video" losses by reducing the number of bits used for each video block and thus increasing the probability of timely retransmission. We plot our average distortion and loss results for different end-to-end delay values.

We provide results for the following rate control algorithms:

- The Dynamic Programming based algorithm for the expected average rate constraint case of Formulation 2.
- The Lagrangian optimization solution to the same problem introduced in Section V-A.2.
- The Dynamic Programming based algorithm for the expected distortion case of Formulation 3.
- The Dynamic Programming with average constraints is also used in the case when no knowledge of the channel is available; in this case the video encoder assumes the average rate is available, i.e., $\bar{C} \times P_e$, where P_e is the probability of packet loss.
- Finally, we also consider the unrealistic scenario where the encoder has advance knowledge of the future channel rates. This gives us an indication of the loss in performance due to imperfect channel model.

Based on our experimental results it is easy to see that the performance, as is to be expected, improves as we increase the information available about the channel state. Thus, performance when no feedback is given is worse than in the case where real time feedback and a channel model are available, which in turn has worse performance than the case where future rates are known.

It can also be seen that the approach based on expected distortion generally outperforms the expected rate approaches. In general the distortion due to packet losses will be much higher than that due to using a coarse quantizer so, even in the system based on expected distortion, the rate control algorithms will tend to minimize the packet losses. We can observe that in all cases (except in the case where no feedback is available, obviously) the distortion and packet loss is reduced when the end-to-end delay in the system increases. Note in particular that the losses in some cases can be made very close to zero.

Finally, it is worth pointing out the difference between uplink and downlink channels. The former are nearly random and error bursts tend to be very short (of the order of magnitude of the feedback delay) thus the difference in performance between having and not having exact knowledge

[6]The Peak Signal to Noise Ratio for a sequence is defined as $PNSR = 10 * log_{10}(255^2/MSE)$, where MSE is the average Mean Squared Error for the whole sequence

	Downlink Channel		Uplink Channel	
	Two-state	N-state	Two-state	N-state
Pr(Good state)	0.9940	0.9940	0.9328	0.9328
Pr(Good $\rightarrow$ Bad)	0.001035	0.001469	0.03382	0.06429
Pr(Bad $\rightarrow$ Good)	0.1720	0.2442	0.46945	0.8924
Avg. burst length (packets)	5.8136	4.0950	2.1302	1.1205

TABLE III
Summary of the characteristics of the channels used in our experiments

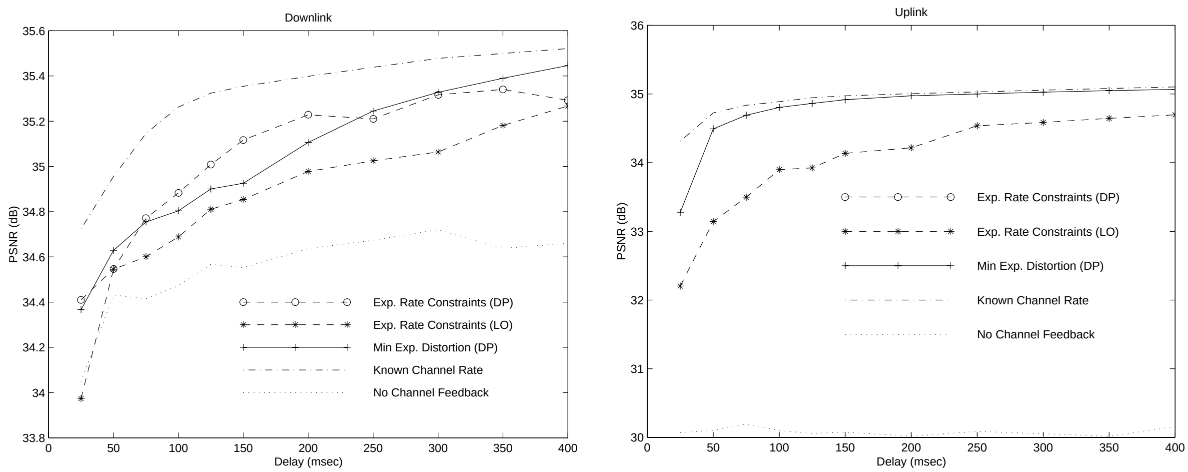

Fig. 8. N**-state Markov channel model:** Resulting PSNR of the decoded video under end-to-end delay constraint from 50 msec to 400 msec. Dynamic Programming (DP) and Lagrangian Optimization (LO) are used for selecting encoding rates.

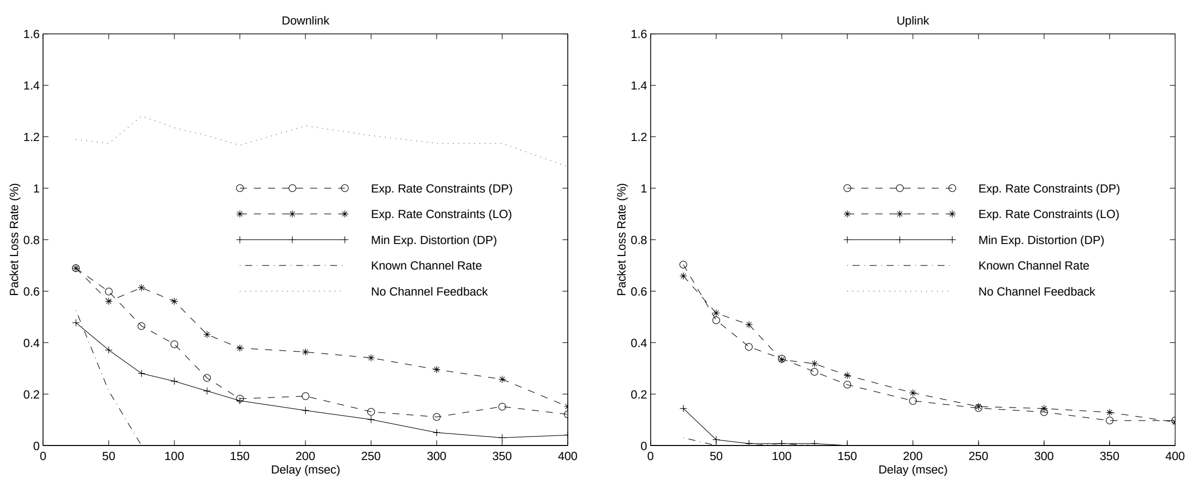

Fig. 9. N**-state Markov channel model:** Resulting packet loss rate under end-to-end delay constraint from 50 msec to 400 msec. Dynamic programming approach with minimal expected distortion criteria is used as rate-control algorithm.

of the channel rates is relatively modest. This indicates that the dynamics of the channel are too fast with respect to the response time of the rate control, thus most rate control approaches perform similarly (the algorithm without feedback still performs worse because it has no knowledge of the channel state, i.e., data is not recoded when there are channel losses). Note that by comparison exact knowledge of the channel behavior does result in improvements in the downlink channel. This can be justified by the longer average burst sizes and higher variances in burst sizes.

B. Complexity Comparison

It is also worth noting that Lagrangian optimization approach is much faster than the dynamic programming approach. Table IV summarizes the complexity comparison from our simulations. While the reported computation

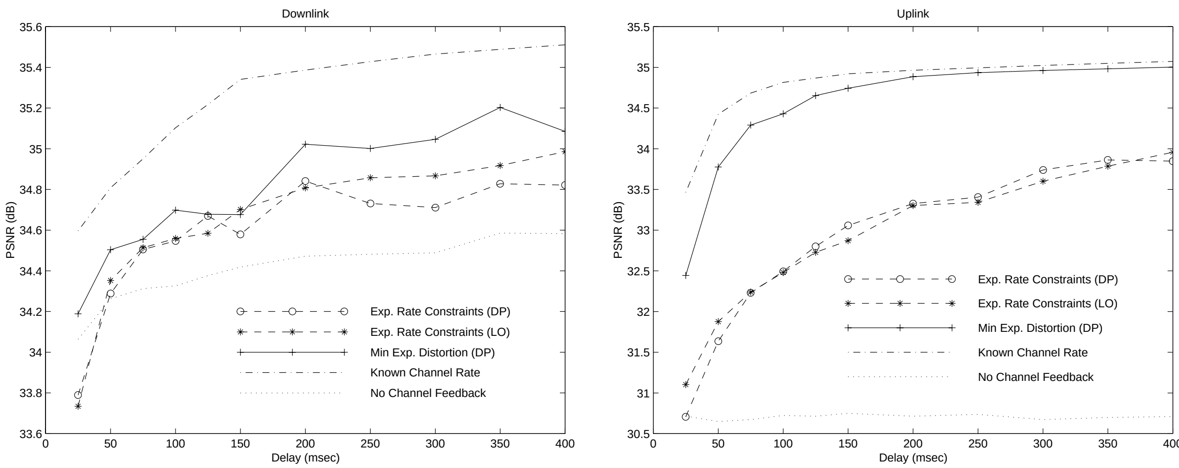

Fig. 10. **Two-state Markov channel model:** Resulting PSNR of the decoded video under end-to-end delay constraint from 50 msec to 400 msec. Dynamic Programming (DP) and Lagrangian Optimization (LO) are used for selecting encoding rates.

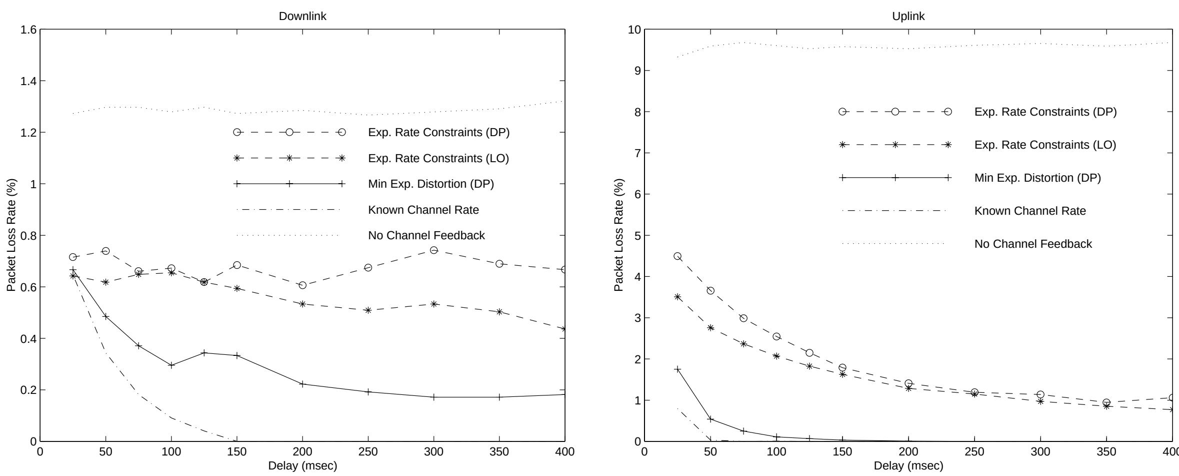

Fig. 11. **Two-state Markov channel model:** Resulting packet loss rate under end-to-end delay constraint from 50 msec to 400 msec. Dynamic programming approach with minimal expected distortion criteria is used as rate-control algorithm.

Delay (msec)	50	100	150	200	250	300	350	400
LO	1.1	1.4	1.8	2.5	3.0	3.8	4.4	5.1
DP	11.7	60.9	150.0	277.1	431.9	621.9	901.0	1241.4

TABLE IV

Complexity comparison in terms of CPU time (sec) for simulating the Lagrangian optimization (LO) and dynamic programming (DP) rate control approaches.

times depend obviously on the specific implementation we use, they do give an order of magnitude of the relative complexities. The basic observation to explain the difference in complexities is that in the Lagrangian approach the complexity depends linearly on the number of blocks considered, i.e., the end-to-end delay, while in the dynamic programming approach the complexity depends on the number of states in the trellis, which has a growth that is approximately quadratic on the number of blocks. Thus we can observe that comparing the high and low delay cases in Table IV an increase of a factor of 8 in delay results in increases in computation times of factors of 5 and 100 for the Lagrangian and DP approaches, respectively.

Note that all our discussion has concentrated on the video encoder, and thus we have considered only the complexity of the RD optimization at the encoder. We have not taken into account other computation requirements at the encoder (they would be the same for the two algorithms considered). Also, the decoder complexity would be the same regardless of the algorithm used, and in fact it would also be the same as in the case where a non-optimized, simple rate control approach is used.

C. Performance Degradation due to Inaccurate Channel Model and Long Feedback Delay

In Section IV-E we discussed that an inaccurate channel model and the existence of significant feedback delay may cause degradation of the proposed rate control approach that uses the probabilistic channel model and the channel feedback information. The following experiments allow us to assess the resulting performance degradation.

C.1 Mismatched channel model

We still assume that the underlying channel behavior for a downlink channel can still be accurately modeled as two-state model with transition probability $\boldsymbol{P}$ as defined in Table III, i.e.,

$$\boldsymbol{P} = \begin{bmatrix} 0.998965 & 0.001035 \\ 0.1720 & 0.8280 \end{bmatrix} \tag{42}$$

However, the encoder uses inaccurate two-state models that are different from the the correct model in transition probability as

$$\boldsymbol{P}_{high} = \begin{bmatrix} 0.998965 & 0.001035 \\ 0.043002 & 0.956998 \end{bmatrix} \tag{43}$$

(Higher error rate than $\boldsymbol{P}$.)

$$\boldsymbol{P}_{low} = \begin{bmatrix} 0.999742 & 0.000258 \\ 0.1720 & 0.8280 \end{bmatrix} \tag{44}$$

(Lower error rate than $\boldsymbol{P}$.)

The rate control scheme is base on dynamic programming to minimize the expected distortion. The performance of the mismatched channel model cases are compared to that with correct channel model as depicted in Fig. 14. We can see that the rate control with model $\boldsymbol{P}_{high}$ suffers from performance degradation by "over-reacting" to the feedback of bad channel condition and introducing additional distortion by using fewer encoding bits, while the rate control with model $\boldsymbol{P}_{low}$ suffers performance degradation by under-estimating the error rates; in this case the additional distortion is caused by packet losses, as the rate control model was too optimistic. In the above simulations we assume that there is no feedback delay (b=0) so that we can completely separate the two factors (i.e., delay and channel mismatch) affecting the performance of the algorithm

C.2 Feedback Delay

Fig. 12 and 13 show the resulting PSNR and packet loss rate with various feedback delay when the two rate control approaches, dynamic programming approach to minimize the expected distortion (Fig. 12) and Lagrangian optimization approach with expected rate constraints (Fig. 13), are used. It can be observed that, as expected, the performance of the rate control degrades, in terms of PSNR and packet loss rate, as the feedback delay increases.

D. Conclusions

In summary, in this paper, we have proposed rate control algorithms for robust video transmission over wireless channel with bursty channel errors. The rate control is integrated with the ARQ error control to comply with the delay constraints of the real-time video transmission. The channel feedback and channel model are used by the encoder to adjust video encoding rate subject to the change of channel condition. Our results indicate that using feedback results in lower packet losses and higher reconstructed PSNR.

Acknowledgments: The authors would like to thank the anonymous reviewers for their comments which helped improve the manuscript.

References

[1] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Low bit-rate video transmission over fading channels for wireless microcellular systems," *IEEE Trans. on Circ. and Sys. for Video Tech.*, pp. 1–11, Feb. 1996.

[2] M. Zorzi, R. R. Rao, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Trans. on Veh. Tech.*, vol. 46, pp. 445–455, May 1997.

[3] H. Ibaraki *et. al.*, "Mobile video communication techniques and services," in *Proc. of Visual Communic. and Image Proc., VCIP'95*, (Taipei, Taiwan), pp. 1029–1033, 1995.

[4] Y. Wang, Q.-F. Zhu, and L. Shaw, "Maximally smooth image recovery in transform coding," *IEEE Trans. on Comm.*, vol. 41, pp. 1544–1551, Oct. 1993.

[5] R. Stedman, H. Gharavi, L. Hanzo, and R. Steele, "Transmission of subband-coded images via mobile channels," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 3, no. 1, pp. 15–26, 1993.

[6] Y. J. Liu and Y. Q. Zhang, "A mobile data code division multiple access (CDMA) system with power control and its application to low-bit-rate image transmission," in *Proc. IEEE VTC'97*, 1993.

[7] S. Lin, D. J. Costello, and M. Miller, "Automatic repeat request error control schemes," *IEEE Communications Magazine*, pp. 5–17, 1984.

[8] N. Färber, E. Steinbach, and B. Girod, "Robust H.263 compatible video transmission over wireless channels," in *Proc. PCS'96*, pp. 575–578, 1996.

[9] E. Steinbach, N. Färber, and B. Girod, "Standard compatible extension of H.263 for robust video transmission in mobile environments," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 7, pp. 872–881, Dec. 1997.

[10] ITU-T, "Video coding for low bitrate communication." ITU-T Recommendation H.263; version 1, Nov. 1995, version 2, Jan. 1998.

[11] J. Mitchell, W. Pennebaker, C. E. Fogg, and D. J. LeGall, *MPEG Video Compression Standard*. New York: Chapman and Hall, 1997.

[12] V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards: Algorithms and Architectures*. Kluwer Academic Publishers, 2nd ed., 1997.

[13] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Signal Proc.*, vol. 36, pp. 1445–1453, Sept. 1988.

[14] M. Chen and A.N. Willson,Jr., "Rate-distortion optimal motion estimation algorithm for video coding," in *Proc. ICASSP'96*, (Atlanta, GA), pp. 2096–2099, May 1996.

[15] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. on Image Proc.*, vol. 3, pp. 26–40, Jan. 1994.

[16] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. on Image Proc.*, vol. 3, pp. 533–545, Sept. 1994.

[17] D. T. Hoang, *Fast and Efficient Algorithms for Text and Video Compression*. PhD thesis, Brown University, 1997.

[18] D. T. Hoang, E. L. Linzer, and J. S. Vitter, "Lexicographic bit allocation for MPEG video," *Journal of Visual Communication and Image Representation*, vol. 8, Dec. 1997.

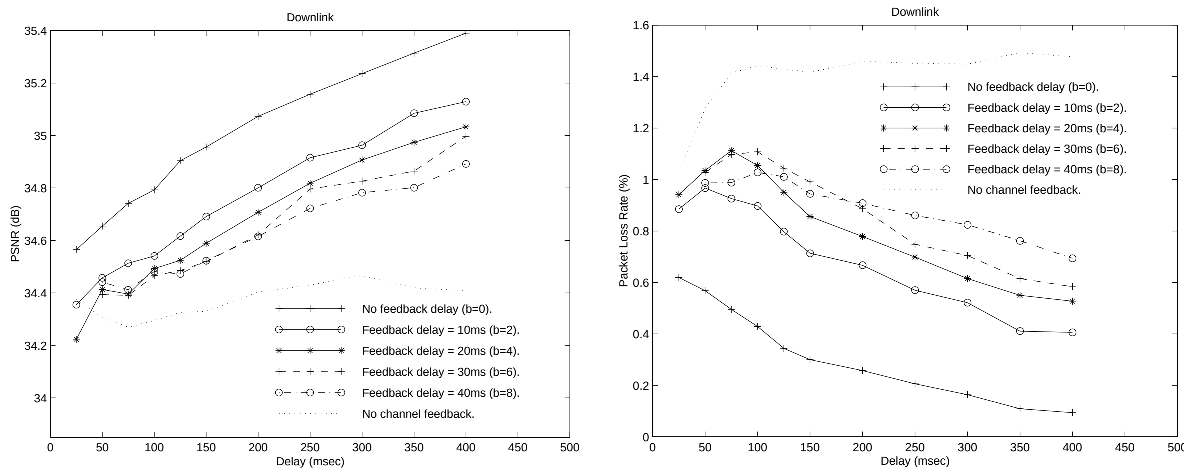

Fig. 12. **Performance Degradation Due to Feedback Delay:** Resulting PSNR and packet loss rate of the decoded video with various feedback delays. Rate control is based on dynamic programming approach to minimize expected distortion.

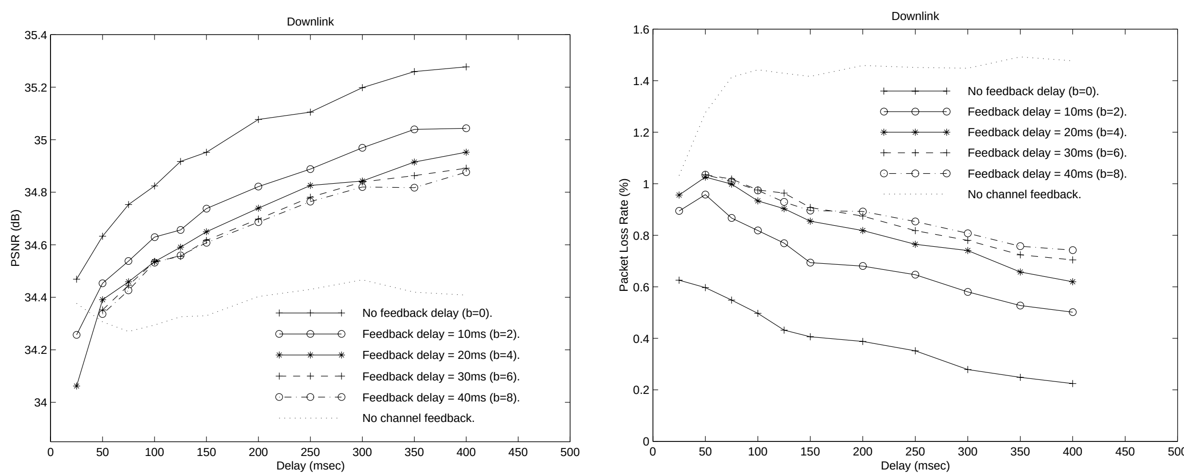

Fig. 13. **Performance Degradation Due to Feedback Delay:** Resulting PSNR and packet loss rate of the decoded video with various feedback delays. Rate control is based on Lagrangian optimization approach with expected rate constraints.

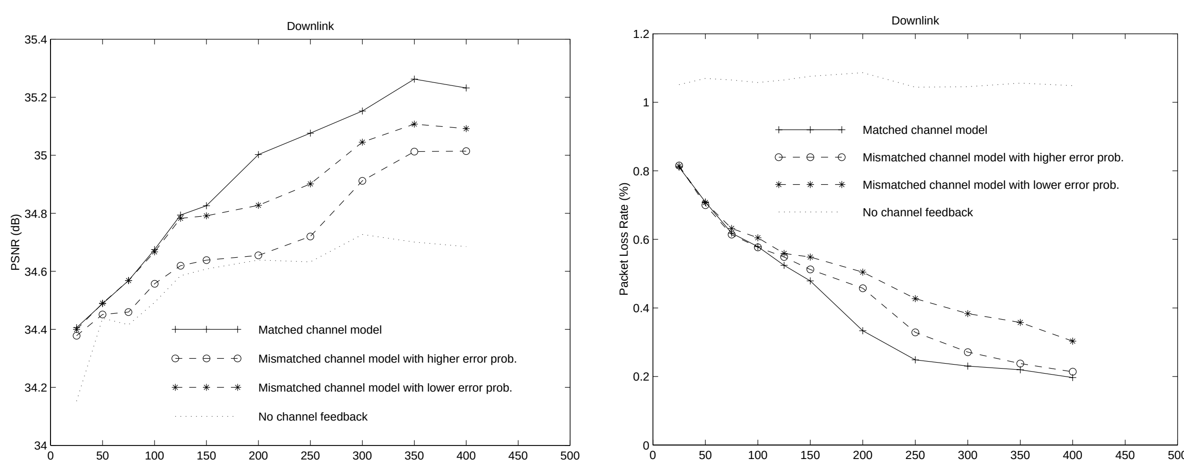

Fig. 14. **Mismatched Channel Models:** Resulting PSNR and packet loss rate of the decoded video when inaccurate channel models are used. Rate control is based on dynamic programming approach to minimize expected distortion.

[19] G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression, Optimal Video Frame Compression and Object Boundary Encoding.* Kluwer Academic Publishers, 1997.
[20] A. Ortega and K. Ramchandran, "Rate-distortion techniques in image and video compression: An overview," *IEEE Signal Processing Magazine*, vol. 15, pp. 23–50, Nov. 1998.
[21] C.-Y. Hsu, A. Ortega, and A. R. Reibman, "Joint selection of source and channel rate for VBR video transmission under ATM policing constraints," *IEEE J. on Sel. Areas in Comm.*, pp. 1016–1028, Aug. 1997.
[22] J.-J. Chen and D. W. Lin, "Optimal bit allocation for coding of video signals over ATM networks," *IEEE JSAC*, vol. 15, pp. 1002–1015, Aug. 1997.
[23] A. Ortega and M. Khansari, "Rate control for video coding over variable bit rate channels with applications to wireless transmission," in *Proc. of ICIP'95*, (Washington, D.C.), Oct. 1995.
[24] C.-Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over wireless channels," in *Proc. of Visual Communic. and Image Proc., VCIP'97*, (San Jose, California), pp. 1200–1211, Feb. 1997.
[25] H. Liu and M. El Zarki, "Adaptive source rate control real-time wireless video transmission," *ACM/Baltzer Mobile Networks and Applications Journal*, vol. 3, no. 1, pp. 49–60, 1998.
[26] R. H. M. Hafez and G. R. Rajugopal, "Adaptive rate controlled robust video communications over packet wireless networks," *ACM/Baltzer Mobile Networks and Applications Journal*, vol. 3, no. 1, pp. 33–47, 1998.
[27] A. R. Reibman and B. G. Haskell, "Constraints on variable bit-rate video for ATM networks," *IEEE Trans. on Circuits and Systems*, vol. 2, pp. 361–372, Dec. 1992.
[28] TIA/EIA/IS-95, "Mobile-station-base-station compatibility standard dual-mode wideband spread spectrum cellular system," Jan. 1993.
[29] ITU-T Recommandation I.363, "B-ISDN ATM Adaptation Layer (AAL) specification," Jun. 1993.
[30] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Robust low bit-rate video transmission over wireless access systems," in *Proc. of ICC'94*, vol. 1, pp. 571–575, May 1994.
[31] H. S. Wang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," in *Proc. ICUPC'94*, pp. 160–164, 1994.
[32] M. Zorzi, R. R. Rao, and L. Milstein, "On the accurancy of a first-order Markov model for data transmission on fading channels," in *Proc. IEEE ICUPC'95*, Nov. 1995.
[33] E. N. Gilbert, "Capacity of burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sept. 1960.
[34] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sept. 1963.
[35] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. Info. Theory*, vol. IT-13, no. 2, pp. 221–227, 1966.
[36] C.-Y. Hsu, *Rate control for video transmission over variable rate channels.* PhD thesis, University of Southern California, Aug. 1998.
[37] D. P. Bertsekas, *Dynamic Programming.* Prentice-Hall, 1987.
[38] A. Ortega, "Optimal rate allocation under multiple rate constraints," in *Data Compression Conference'96*, (Snowbird, Utah), Mar. 1996.
[39] ITU-T Recommendation H.261, "Video codec for audiovisual services at p×64 kbit/s," 1993.
[40] The Portable Video Research Group (PVRG), Stanford University, "PVRG-P64 CODEC v. 1.2." Available URL: ftp://havefun.stanford.edu/pub/p64/P64v1.2.tar.Z.

Chi-Yuan Hsu received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California (USC) in 1994 and 1998, respectively. He was a Research Assistant in the Signal and Image Processing Institute and the Integrated Media Systems Center at USC from 1994 to 1998. Since June 1998 he has been with Sony Semiconductor Company of America, San Jose, CA.

He received the 1997 IEEE Communications Society Leonard G. Abraham Prize Paper Award for a paper co-authored with A. Ortega and A. R. Reibman which appeared in IEEE JSAC, Aug. 1997. He also received an Outstanding Reseach Paper Award from the Dept. of Electrical Engineering-Systems, USC, in 1998. His research interests include video rate control, joint source-channel rate allocation, and video transmission over broad-band and wireless networks.

Antonio Ortega Antonio Ortega was born in Madrid, Spain, in 1965. He received the Telecommunications Engineering degree from the Universidad Politecnica de Madrid (UPM), Madrid, Spain in 1989 and the Ph.D. in Electrical Engineering from Columbia University, New York, NY in 1994. He was a research assistant at the Image Processing Group at UPM (1990). At Columbia he was a graduate research assistant at the Center for Telecommunications Research (1991-94) and was supported by a scholarship from the Fulbright commission and the Ministry of Education of Spain.

Since September 1994 he has been an Assistant Professor in the Electrical Engineering-Systems department at the University of Southern California. At USC he is also a member of the Integrated Media Systems Center (IMSC), an NSF Engineering Research Center, and the Signal and Image Processing Institute. In 1995 he received the NSF Faculty Early Career Development (CAREER) award. He is a member of IEEE, SPIE and ACM. He is a member of the IEEE Signal Processing Society Multimedia Signal Processing Technical Committee and the technical program co-chair for the 1998 Workshop on Multimedia Signal Processing. Since 1996 he is also an Associate Editor for the IEEE Transactions on Image Processing. In 1997 he was the recipient of the USC School of Engineering Northrop-Grumman Junior Research Award. He received the 1997 IEEE Communications Society Leonard G. Abraham Prize Paper Award for a paper co-authored with C.-Y. Hsu and A. R. Reibman which appeared in IEEE JSAC, Aug. 1997.

His research interests are in the areas of image and video compression and communications. They include topics such as adaptive methods for image/video coding, joint source-channel coding for robust video transmission, rate control and video transmission over packet wired or wireless networks.

Masoud Khansari Masoud Khansari received the B.A.Sc. and the Ph.D. degrees both in electrical engineering from the University of Toronto, Ontario, in 1988 and 1993, respectively. He then joined the INRS-Telecommunication at Verdun, Quebec. He spent the academic years of 1994 and 1995 at the University of California at Berkeley and Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland. Since June 1996, he has been a member of scientific staff at Hewlett-Packard Laboratories, Palo Alto, California.