Joint Source Coding and Data Rate Adaptation for Energy Efficient Wireless Video Streaming

Carlos E. Luna, *Associate Member*, Yiftach Eisenberg, *Student Member, IEEE*, Randall Berry, *Member, IEEE*, Thrasyvoulos N. Pappas, *Senior Member, IEEE*, and Aggelos K. Katsaggelos, *Fellow, IEEE*

Abstract—**Rapid growth in wireless networks is fueling demand for video services from mobile users. While the problem of transmitting video over unreliable channels has received some attention, the wireless network environment poses challenges such as transmission power management that have received little attention previously in connection with video. Transmission power management affects battery life in mobile devices, interference to other users, and network capacity. We consider energy efficient transmission of a video sequence under delay and quality constraints. The selection of source coding parameters is considered jointly with transmitter power and rate adaptation, and packet transmission scheduling. The goal is to transmit a video frame using the minimal required transmission energy under delay and quality constraints. Experimental results are presented that illustrate the advantages of the proposed approach.**

Index Terms—**Energy efficiency, power and rate control, video streaming, wireless video.**

I. INTRODUCTION

RAPID GROWTH in wireless networks is fueling the demand that services traditionally available only in wire-line networks, such as video, be available to mobile users. However, several important issues, such as transmitter power control, are unique to wireless networks and deserve special attention. In this paper, we consider the interaction of video compression and transmitter power and rate adaptation. Our goal is to efficiently utilize transmission energy while meeting the delay and video quality constraints imposed by a video streaming application.

In wireless networks, communication takes place over a time-varying and unreliable channel. Video transmission over unreliable networks has been an active field of research. Error resilience and error concealment have received considerable attention [1], [2]. A complimentary approach is to adapt the behavior of the video encoder to the conditions of the channel. For example, in [3], the objective is to minimize the expected distortion at the receiver subject to rate constraints derived from a stochastic model of the wireless channel and application delay constraints. In [4]–[7], the approach is to select the coding mode for each macroblock (MB) taking into account the probability of packet loss in the channel and the error concealment technique used by the decoder in order to reduce the expected distortion at the receiver.

In the papers cited above, the theme is to adapt the behavior of the video encoder and decoder to cope with the effects of a lossy and time-varying channel. Alternatively, through the use of transmitter power and rate adaptation, the characteristics of the wireless channel as seen by the video encoder can be changed. For example, increasing the transmission power can lead to higher throughput. However, this may lead to increased interference for other users or inefficient use of the available battery energy. The use of transmitter power and rate control to manage these tradeoffs has received considerable attention including [8]–[15]. Other approaches include putting a communication device into "sleep mode" when it is not required by an application [16]. Similar strategies have been studied within IEEE 802.11 [17].

For streaming traffic, an important consideration is meeting the delay constraints of the application. The problem of designing energy efficient transmission policies for randomly arriving traffic with delay constraints has been studied in [12], [13], [18], and [19]. The goal in these papers is to minimize the total amount of energy expended at the transmitter while meeting constraints on the delay experienced by the data. Policies that maintain an average buffer delay were studied in [12]. In [13], the basic tradeoff between average queueing delay and average transmission power was characterized. In [18] and [19], the problem of scheduling a set of packets by a given deadline with minimum energy is considered.

Streaming video traffic has a specific set of delay constraints that do not fit into the above settings. Furthermore, the arriving data is not entirely random, but depends in part on the source coding decision made by the video encoder. Our approach here is to jointly consider both the physical layer power control and scheduling along with adaptation of source coding parameters. We briefly mention several other approaches along similar lines. In [20] and [21], a joint source coding and power control (JSCPC) approach to allocating source rate and transmission power under bandwidth constraints for an individual user is considered. In [22], the goal is to adjust the source coding parameters and the allocation of transmitter power in order to spend the minimal amount of energy required to transmit a video sequence, over a fixed rate channel, subject to an expected distortion and delay constraint. The authors extend this work in [23] by accounting for both the expected value and the variance of the end-to-end distortion.

Fig. 1.   System block diagram.

Fig. 2.   Delay constraint components for a video packet.
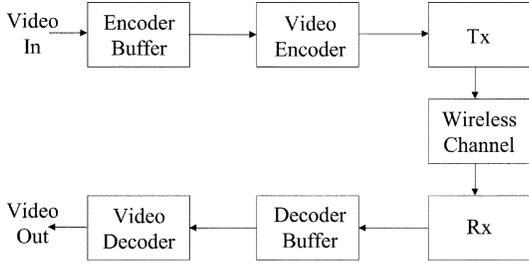
In the next section, we present our problem formulation in detail. In Section III, we present two algorithms based on dynamic programming (DP) for solving this problem. Several experimental results illustrating the tradeoffs in energy and distortion are presented in Section IV. Finally, Section V contains some concluding remarks.

II. PROBLEM FORMULATION

A block diagram of the system considered in this paper is shown in Fig. 1. Video frames are captured and stored in the encoder buffer. The video encoder reads video data from the encoder buffer and produces a stream of video packets that is transmitted over a wireless channel. We utilize in this work a hybrid motion compensated video encoder, such as the ones implemented by all existing video compression standards (e.g., H.263, MPEG-1-2, baseline MPEG-4). Each video packet is made up of a sequence of consecutive MBs and can be independently processed by the encoder or decoder. The transmitter (Tx) can dynamically allocate transmission rate and power at the physical layer for each packet in order to meet the delay constraints of the application and ensure reliable transmission. Several techniques for data rate adaptation have been incorporated into existing wireless standards (for example, see [24] for a survey of techniques that are currently used). In addition to deciding on the rate at which a packet is sent, the transmitter may schedule the time at which packet transmission begins. When the channel conditions are poor, this allows the transmitter to idle until a more favorable time. At the receiver (Rx), the incoming video packets are received and stored in the decoder buffer. The decoder reads video packets from this buffer and displays the video sequence in real-time. By real-time display, we mean that once the receiver begins displaying the received video, the display process continues uninterrupted, without stalling. If video data does not arrive on time to be displayed, then this data is considered lost. In this situation, if the receiver and transmitter are to operate at the same frame rate, then each frame must experience a constant end-to-end delay. We define the end-to-end delay as the amount of time between frame capture and display at the decoder. The size of the delay depends on the nature of the application. Interactive applications, such as video conferencing, require that the end-to-end delay be on the order of 200 ms. On the other hand, one-way applications, such as video on demand, can tolerate delays on the order of several seconds.
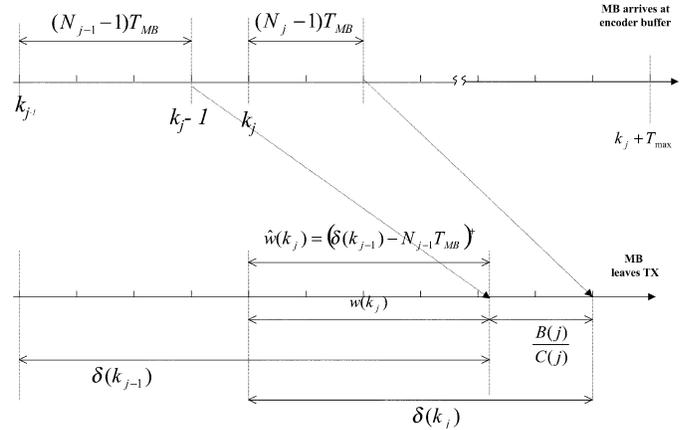
A. Delay Constraints

In this section, we translate the constant end-to-end delay constraint on a video frame into delay constraints for each video packet. The diagram in Fig. 2 illustrates the various components involved in these delay constraints. First, we consider the delay constraint at the MB level. Let M be the number of MBs in a video frame and k the MB index ($k = 0, \ldots, M - 1$). A video frame captured at time t must be displayed at time $t + \Delta T$, where ΔT is the required end-to-end delay for every video frame. Therefore, MB k and all following MBs, i.e., $k + 1, \ldots, M - 1$, in the frame must be available at the decoder in time to be decoded. That is, MB k must arrive before $t + \Delta T - (M - k)\Delta T_d$, where ΔT_d is the time required to decode one MB (assumed to be a known constant). Also, note that MB k becomes available at the encoder after the previous MBs in the frame, i.e., $0, \ldots, k - 1$, have been encoded. That is, MB k becomes available to the video encoder at time $t + k\Delta T_e$, where ΔT_e is the time required to encode a MB (also assumed to be a known constant). Therefore, the following must hold for the video frame to experience constant end-to-end delay:

$$\Delta T_{eb}(k) + \Delta T_c(k) + \Delta T_{db}(k)$$
$$= \Delta T - (k+1)\Delta T_e - ((M-k)\Delta T_d), \ \forall k \quad (1)$$

where $\Delta T_{eb}(k)$ is the encoder buffer delay, $\Delta T_c(k)$ is the channel transmission time and $\Delta T_{db}(k)$ is the decoder buffer delay for MB k. Given $\Delta T_{eb}(k)$ and $\Delta T_c(k)$, $\Delta T_{db}(k)$ represents the amount of time a MB waits in the decoder buffer before it is decoded. Therefore, in order to avoid stalling, we need $\Delta T_{db}(k) \geq 0$. Therefore, for each MB, the following constraint must be met:

$$\delta(k) \leq \Delta T - (k+1)\Delta T_e - ((M-k)\Delta T_d), \ \forall k \quad (2)$$

where $\delta(k) = \Delta T_{eb}(k) + \Delta T_c(k)$. This delay constraint can be enforced at the encoder for each MB. To simplify our discussion, we assume $\Delta T_e = \Delta T_d = T_{\text{MB}}$. This implies that the delay constraint on each MB, in (4), is a constant, $T_{\text{max}} = \Delta T - (M + 1)T_{\text{MB}}$. In the following, we refer to $\Delta T_{eb}(k)$ as $w(k)$, the waiting time before MB k is transmitted. A video packet is made up of a sequence of consecutive MBs and should arrive at

the decoder buffer in time to meet the delay constraints for all the MBs in the packet. Note that since we assume $\Delta T_e = \Delta T_d$, it is sufficient to meet the delay constraints for the first MB in a video packet.

The original video frame arrives at the encoder as a stream of MBs spaced every T_{MB} s. Each MB is then encoded and placed in a video packet. Consider a video packet j of size $B(j)$ bits and made up of N_j MBs. We denote the first MB in packet j by k_j, with $k_j = k_{j-1} + N_{j-1}$. Assume packet j is transmitted at rate $C(j)$ bits/s. The delay experienced by MB k_j, can then be expressed as,

$$\delta(k_j) = w(k_j) + \frac{B(j)}{C(j)} \qquad (3)$$

where $B(j)/C(j)$ is the transmission delay for packet j, and $w(k_j)$ is the waiting time for MB k_j. The waiting time, $w(k_j)$, is made up of three components. First, the packetization delay, which is the time for the rest of the MBs in the packet to become available, i.e., $(N_j-1)T_{\mathrm{MB}}$. The second component is any additional time the packet must wait for the preceding packet to finish its transmission. Finally, there is the scheduling delay $\tilde{w}(k_j)$, which is the time the transmitter waits to begin transmission of the packet. Therefore, $w(k_j)$ can be expressed as

$$w(k_j) = (N_j - 1)T_{\mathrm{MB}} + \left(\delta(k_{j-1}) - (N_{j-1} + N_j - 1)T_{\mathrm{MB}}\right)^+ + \tilde{w}(k_j). \quad (4)$$

The first term in (4) is the packetization delay and the second term corresponds to the waiting time before packet transmission can begin. It will be convenient to rewrite (4) as

$$w(k_j) = (N_j - 1)T_{\mathrm{MB}} + (\hat{w}(k_j) - (N_j - 1)T_{\mathrm{MB}})^+ + \tilde{w}(k_j) \quad (5)$$

where $\hat{w}(k_j) = (\delta(k_{j-1}) - N_{j-1}T_{\mathrm{MB}})^+$ is the waiting time for MB k_j if packet j consists of only one MB and there is no scheduling delay, i.e., $N_j = 1$ and $\tilde{w}(k_j) = 0$ in (3). This situation is illustrated in Fig. 2.

Our goal is to assign source coding parameters (e.g., quantization step size and coding mode), and communication resources at the physical layer (e.g., transmission rate, transmission power, and transmission schedule) to each video packet to ensure the delay constraints are satisfied and that acceptable video quality can be maintained while using the minimum required transmission energy.

B. Channel Model

In a wireless setting, communication takes place over a channel with a time-varying response. Impairments such as fading and multipath have a significant impact on the performance of the communication system. Several dynamic resource allocation techniques have been developed to combat these impairments [10], [25], [26]. The techniques presented in this paper only require knowledge of a function relating transmission rate and power to channel state information. Such a function can be obtained from an analytical model of the wireless channel or from empirical measurements.

We consider a specific example based on a slowly-varying wireless channel with frequency nonselective block fading, modeled as a finite-state Markov channel (FSMC) [27]. In this channel model, the fading process is modeled by a finite-state Markov chain with state space $\mathcal{H}$. The fading transitions occur every T_c s and are governed by the transition matrix $\mathbf{A}$ of the Markov chain. We assume that a function relating transmission power to the desired transmission rate and channel state information is known at the transmitter. The expected amount of energy required to transmit video packet k of size $B(k)$ bits at a rate of $C(k)$ bits/s can be expressed as

$$E(k) = \mathbb{E}_H\left\{ \sum_{l=\lceil \tilde{w}(k)/T_c \rceil}^{L_k-1} P(H_{k,l}, C(k))T_c \,\Big|\, H_{k,0} = h_{k,0} \right\} \quad (6)$$

where $P(H_{k,l}, C(k))$ is the required power, $\{H_{k,l}\}$ is the sequence of fading states during the packet transmission and L_k is the number of channel blocks it takes to transmit packet k at the given rate. In this expression, we are conditioning on the fading state $H_{k,0}$ at the start of the packet's transmission. We can express L_k as

$$L_k = \left\lceil \frac{B(k)}{C(k)T_c} \right\rceil + \left\lceil \frac{\tilde{w}(k)}{T_c} \right\rceil. \quad (7)$$

Thus, the expected energy required to transmit a packet depends only on the statistics of the channel and L_k. This expected cost can be computed off-line, given the channel model, and implemented as a table lookup at the transmitter.

As an example, we consider adapting transmission power to maintain a channel with a given capacity. During time-slot l, we model the channel over which packets are being sent as a band-limited additive white Gaussian noise (AWGN) channel with gain $\sqrt{h(l)}$. We assume that the gain stays fixed during each time slot and is assumed to be known at both the transmitter and receiver.

If the desired transmission rate for the kth packet is $C(k)$, we assume that the required transmission power at each time slot is the minimum power such that the channel over which this packet is sent has Shannon capacity $C(k)$, i.e.,

$$P(h_{k,l}, C(k)) = \frac{N_0 W}{h_{k,l}}\left(2^{C(k)/W} - 1\right) \quad (8)$$

where W is the bandwidth of the channel and $h_{k,l}$ is the channel gain, N_0 is the power spectral density of the noise.

From Shannon's coding theorem, (8) gives a lower bound on the transmission power required to reliably transmit at rate $C(k)$; moreover, for large enough packets, this bound will be approachable and will give a reasonable indication of the required power.

C. Optimization Problem

Video packet j is composed of N_j MBs. Each MB k is coded using a quantizer, $q(k)$ chosen from a finite set $\mathcal{Q}$, resulting in distortion $D(k)$ and rate $R(k)$ bits. The size of the video packet payload is given by $B(j) = \sum_{k=k_j}^{k_j+N_j-1} R(k)$. We wish to transmit the resulting video packet at rate $C(j)$ bits per second chosen from a finite set of allowable channel rates $\mathcal{C}$. Our goal is then to select the number of MBs in each video packet, the

coding parameters for these MBs, and a transmission rate and schedule for each packet with the objective of minimizing the total expected energy required to transmit the video frame subject to both an expected total distortion constraint and a delay per packet constraint. The expectations are taken with respect to the channel state denoted by the random process $H(k)$. We pose this the following constrained optimization problem

$$\min_{N_j,C(j),\tilde{w}(j),q(k)} \mathbb{E}_H \left\{ \sum_j E(j) \right\}$$

$$\text{s.t.}: \mathbb{E}_H \left\{ \sum_j \sum_{k=k_j}^{k_j+N_j-1} D(k) \right\} \leq D_T$$

$$\delta(k_j) \leq T_{\max}, \ \forall k_j \qquad (9)$$

where $\delta(k_j)$ and $E(j)$ are given by (2) and (6), respectively. The initial conditions $w(0)$ and $h_0(0)$ are the initial wait time and the initial channel state, respectively.

Increasing the number of MBs in a packet can result in increased source coding efficiency, which may lead to better energy-distortion tradeoffs. Furthermore, under adverse channel conditions the packetization delay effectively allows us to wait for more favorable channel conditions, which can result in additional energy savings [28], [29]. On the other hand, more MBs result in larger packets which reduce the ability to adapt the transmission rate in response to channel conditions.

III. PROPOSED ALGORITHM

In this section, we present a solution to the optimization problem in (9) based on Lagrangian relaxation and DP. First, we relax the distortion constraint. Thus, we introduce a Lagrange multiplier $\lambda > 0$ and solve the following relaxed problem:

$$\min_{N_j,C(j),\tilde{w}(k_j),q(k)} \mathbb{E}_H \left\{ \sum_j \left[E(j) + \lambda \sum_{k=k_j}^{k_j+N_j-1} D(k) \right] \right\}$$

$$\text{s.t.}: \delta(k_j) \leq T_{\max}, \ \forall k_j. \qquad (10)$$

This relaxed problem can be solved using techniques from DP [30]. By appropriately choosing λ, the problem of (9) can be solved within a convex-hull approximation by solving (10) [31]. The search for an appropriate choice of λ can be carried out by the bisection algorithm or a fast convex search technique.

A. DP Solution of Relaxed Problem

In this section, we describe in detail our solution to the relaxed problem of (10). Consider the situation where we want to transmit a video packet j, with initial MB $k_j = k$. Then, we need to specify the number of MBs, the quantizers and transmission rate for this video packet. These decisions are based on a state defined as

$$\mathbf{x}(k) = \begin{bmatrix} \hat{w}(k) \\ h(k) \end{bmatrix} \qquad (11)$$

where $\hat{w}(k)$ is given in (5) and $h(k)$ is the channel state when we consider MB k. Note that $\hat{w}(k) \in [0, T_{\max}]$ is real-valued and, thus, the resulting state space is infinite. For computational reasons, we quantize $\hat{w}(k)$ into a set of $N_W + 1$ values,
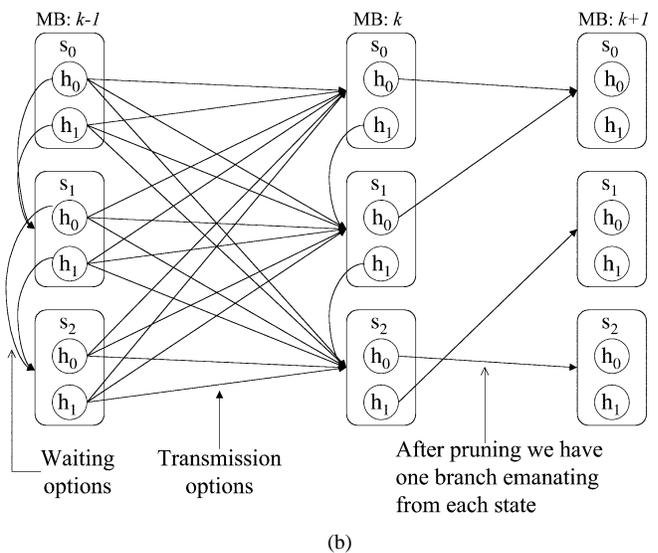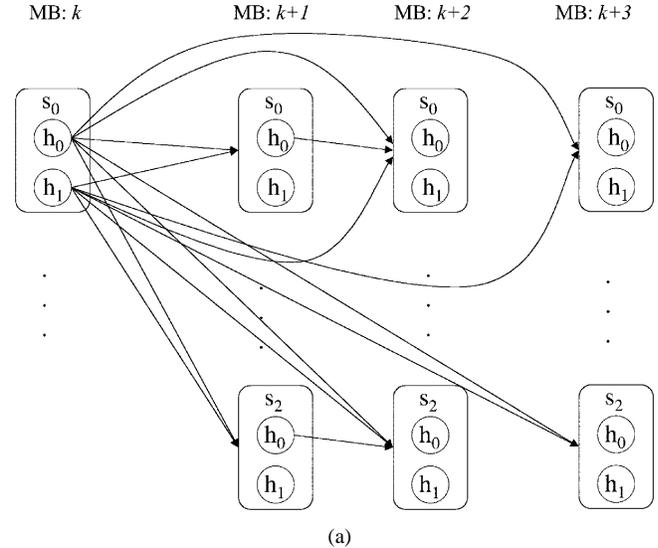

Fig. 3. (a) DAG formulation of energy minimization problem. Each branch leads to a deterministic value of $\hat{w}$. The channel state is determined by the channel statistics. (b) DAG formultaion for problem with packet transmission scheduling.

$\{s_0, \ldots, s_{N_W}\}$, as will be described later. The resulting optimization problem is equivalent to solving a stochastic shortest path problem for a directed acyclic graph (DAG) such as the one depicted in Fig. 3(a), for $N_W = 2$ and $\mathcal{H} = \{h_0, h_1\}$. In this diagram, four stages corresponding to MBs k to $k+3$ are shown. Each node corresponds to the situation where we start a packet with MB k. Each branch in the graph corresponds to a choice of N_j, a sequence of quantizers, transmission schedule, and rate. Let $\mathcal{U}(\mathbf{x}(k))$ be the set of feasible choices

$$\mathcal{U}(\mathbf{x}(k)) = \bigcup_{N_j=1}^{M-k-1} \mathcal{U}_{N_j}(\mathbf{x}(k)) \qquad (12)$$

where

$$\mathcal{U}_{N_j}(\mathbf{x}(k)) = \left\{ u(k) \in \mathcal{Q}^{N_j} \times \mathcal{C} \times \tilde{\mathcal{W}} \right.$$

$$\left. : \frac{B(j)}{C(j)} + w(k) \leq T_{\max} \right\}. \qquad (13)$$

Here, $\mathcal{C}$ represents the set of allowable channel transmission rates, $\mathcal{Q}^{N_j}$ represents the possible quantizer sequences of length N_j, and $\tilde{\mathcal{W}}$ represents the allowable transmission schedules, which are represented by the set of possible additional waiting times given as

$$\tilde{\mathcal{W}}(\mathbf{x}(k)) = \left\{ \tilde{w} = nT_c : n = 0, 1, \ldots, \left\lfloor \frac{T_{max} - w(k)}{T_c} \right\rfloor \right\}. \quad (14)$$

Thus, the set $\mathcal{U}_{N_j}(\mathbf{x(k)})$ contains all the feasible choices of N_j MBs starting with MB k when the system is in state $\mathbf{x}(k)$.

For each choice of $u(k) \in \mathcal{U}_{N_j}(\mathbf{x}(k))$, the cost incurred by MB k is given by

$$g(\mathbf{x}(k), u(k)) = E(j) + \lambda \sum_{k_j=0}^{N_j-1} D(k_j). \quad (15)$$

In Fig. 3(a), each choice of $u(k)$ is represented by a branch emanating from a node. The waiting time for the next packet is given deterministically by $u(k)$. However, the channel state depends on the transition probabilities from the statistical model of the channel.

We want to find a policy $\mu : (\mathbf{x}(k), k) \to \mathcal{U}(\mathbf{x(k)})$ that minimizes the total expected cost in (10). We solve this problem by using DP. We start the algorithm at $k = M - 1$, that is

$$J_{M-1}^*(\mathbf{x}(M-1)) = \min_{\mathcal{U}(\mathbf{x}(M-1))} \{E(M-1) + \lambda D(M-1)\} \quad (16)$$

is calculated. Then, for $k = M - 2, \ldots, 0$, we recursively define the *cost-to-go* functions $J_k^*(\mathbf{x}(k))$ as

$$J_k^*(\mathbf{x}(k)) = \min_{\mathcal{U}(\mathbf{x}(k))} \mathbb{E}_H \left\{ g(\mathbf{x}(k, u(k)) + J_{k+N_j}^*(\mathbf{x}(k+N_j)) \right\}. \quad (17)$$

In carrying out (17), all feasible combinations of packetization, quantizers, scheduling, and transmission rates are considered for each state. This optimization clearly eliminates all branches but one emanating from each node of the DAG. Given the initial state $\mathbf{x}(0)$, the optimal solution is obtained by backtracking. Clearly, $J_0^*(\mathbf{x}(0))$ is the optimal total expected cost of (10).

B. Quantizing $\hat{w}(k)$

Note that $\hat{w}(k)$ is continuous, which results in an infinite number of possible system states. We approximate the solution to the problem by quantizing $\hat{w}(k)$ and then applying DP to obtain the optimal solution to the resulting approximate optimization problem [30]. Let $\mathcal{S}$ be a finite subset of the nonnegative real numbers given by

$$\mathcal{S} = \{s_0, \ldots, s_{N_W}\} \quad (18)$$

with $s_l = (lT_{max})/N_W$. Then, we have

$$\hat{w}(k_{j+1}) = \left\lceil (\delta(k_{j-1}) - (N_{j-1} + N_j - 1)T_{MB})^+ \right\rceil_{\mathcal{S}} \quad (19)$$

where

$$\lceil x \rceil_{\mathcal{S}} = \min\{s \in \mathcal{S} \mid s \geq x\}. \quad (20)$$

Using this new definition of $\hat{w}(k)$, we apply the DP algorithm in (17) to obtain the optimal solution to the approximated problem. Finer quantization of $w(k)$ leads to better approximations to the optimal solution, at the cost of more computation. Note that the effect of this approximation is to restrict the set of feasible choices $\mathcal{U}(\mathbf{x}(k))$ for each system state. Thus, the resulting solution will be a conservative approximation to the optimal solution.

C. A More Efficient Algorithm

In this section, we give an alternative to the previous algorithm, which results in improved performance and faster computation. In the previous section, the decision of which choice $u(k) \in \mathcal{U}(\mathbf{x(k)})$ is made for each MB only at the time when the MB first can be transmitted. In this section, we allow the transmitter to defer this decision. First, we express the set $\mathcal{U}(\mathbf{x}(k))$ as

$$U(\mathbf{x}(k)) = \bigcup_{n=0}^{\lfloor T_{max} - w(k)/T_c \rfloor} \mathcal{U}_{N_j}^n(\mathbf{x}(k)) \quad (21)$$

where $\mathcal{U}_{N_j}^n(\mathbf{x}(k))$ is the feasible set of $u(\mathbf{x}(k))$ when there are N_j MBs in the packet and the transmitter waits n time slots

$$\mathcal{U}_{N_j}^n(\mathbf{x}(k)) = \left\{ u(\mathbf{x}(k)) \in \mathcal{Q} \times \mathcal{C} \times \{nT_c\} \right. $$
$$\left. : \frac{B(k)}{C(k)} + nT_c + w(k) \leq T_{max} \right\}. \quad (22)$$

The optimization problem can be solved using the following algorithm. We start the algorithm at $k = M - 1$, that is

$$J_{M-1}^*(\mathbf{x}(M-1)) = \min \left\{ \min_{\mathcal{U}_{N_j}^0(\mathbf{x}(M-1))} \right.$$
$$\left. \{E(M-1) + \lambda D(M-1)\}, J_{M-1}^*(\mathbf{x}'(M-1)) \right\} \quad (23)$$

for each possible system state $\mathbf{x}(M-1)$. Note that this procedure must be carried out in order of decreasing $w(k)$. Then, for $k = M - 2, \ldots, 0$, we recursively define the *cost-to-go* functions

$$J_k^*(\mathbf{x}(k)) = \min \left\{ \min_{\mathcal{U}_{N_j}^0(\mathbf{x}(k))} \mathbb{E}_H \left\{ g(\mathbf{x}(k, u(k)) \right.\right.$$
$$\left.\left. + J_{k+1}^*(\mathbf{x}(k+1)) \right\}, J_k^*(\mathbf{x}'(k)) \right\} \quad (24)$$

where $\mathbf{x}'(k)$ is given by

$$\mathbf{x}'(k) = \begin{bmatrix} w(k) + T_c \\ h(k+1) \end{bmatrix} \quad (25)$$

for all k.

Solving this problem is equivalent to finding a stochastic shortest path through a graph as depicted in Fig. 3(b) for $N_j = 1$. Note that in this graph, the option of waiting is represented by a branch that connects to a node in the same stage k. Solving this via DP is possible, as long as the system states for each MB are considered in order of decreasing $w(k)$.
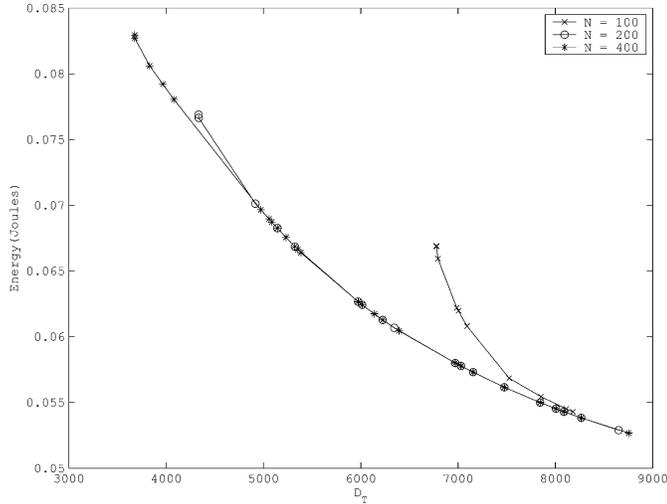
Fig. 4. Convex hull of operational energy distortion points with transmission at $C = 200$ kb/s, $T_{\max} = 166.7$ ms and varying values of N. Increasing N gives a more accurate approximation to the optimal solution.

IV. EXPERIMENTAL RESULTS

In this section, we present some experimental results that illustrate the tradeoffs studied in this paper. We consider transmission of the foreman sequence in QCIF format at 30 fps. The video sequence is encoded with the MPEG-4 implementation provided by MoMuSys. Each MB can be encoded using one of eight quantization parameters given by $\mathcal{Q} = \{2, 4, 8, 16, 20, 24, 28, 31\}$ and can be coded as INTRA or INTER modes. We consider transmission over a channel with bandwidth $W = 500$ kHz and AWGN with variance $N_0 W = 0.39$. The fading is modeled by a two-state Markov chain with state space $\mathcal{H} = \{0.9, 0.1\}$. We use a symmetric transition probability matrix of the form

$$\mathbf{A} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}. \tag{26}$$

A. Experiment I

In this experiment, we illustrate the effect of different problem parameters on the solution. First, consider the effect of the approximation error resulting from the quantization of the state space. We first consider transmission of the first frame of the foreman sequence with $T_{\max} = 100$ ms and $p = 0.7$ in (26). Sweeping the value of $\lambda > 0$ in (10), we obtain the convex hull of operational energy-distortion points. The results of this procedure are shown in Fig. 4 for $N = \{100, 200, 400\}$. We can see from the figure that decreasing the value of N, i.e., using a coarser quantization of the state space, results in a more conservative approximation to the solution of the problem in (10). Note that as the distortion threshold D_T is raised, the curves come closer together. This suggests that the approximation error decreases as D_T increases. Recall that the error introduced by the approximation affects the set of feasible choices defined by the delay constraint. Increasing D_T results in a situation with more flexible delay constraints and, thus, the effect of the approximation error is reduced. Similarly, lowering the value of D_T results in greater approximation error.
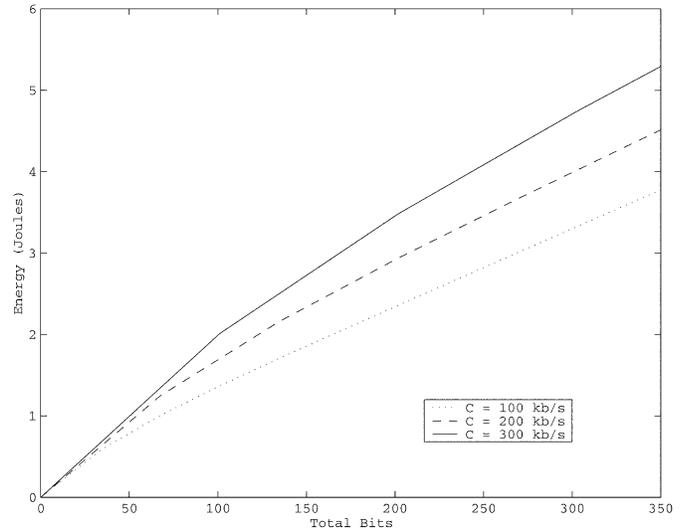

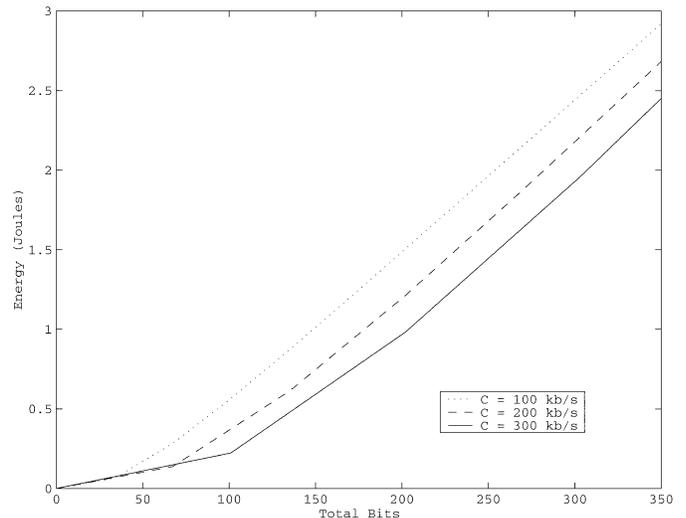Fig. 5. Required average transmission energy for different C (starting in good state).

Fig. 6. Required average transmission energy for different C (starting in bad state).

B. Experiment II

The following experimental results illustrate the tradeoffs of channel rate adaptation, transmission scheduling and packetization for a single frame. Initially, we consider a system with fixed packetization using 1 MB per video packet. The effect of packetization is presented in Section IV-B3.

1) Channel Rate Adaptation: First, consider the situation where a single packet of size B bits is to be transmitted at rate $C \in \mathcal{C}$, using the least amount of required energy. Figs. 5 and 6 show the total average transmission energy required to transmit a packet as a function of its length. In the first figure, the initial channel state corresponds to $h = 0.9$, i.e., the good state. This figure indicates that when transmission starts in the good state, it is advantageous to transmit at the fastest rate available. Another benefit of transmitting at a high rate is that this results in smaller values waiting times for future packets which eases the effect of the delay constraint. On the other hand, Fig. 6 depicts the situation with initial state with $h = 0.1$, i.e., the bad state.
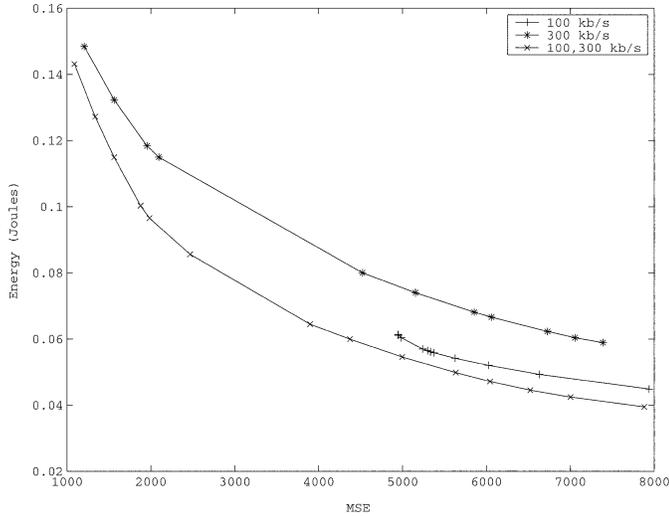
Fig. 7.   Convex hull of energy-distortion operational points.

Fig. 8.   Convex hull of operational energy distortion points with transmission with $C = \{100, 300\}$ kb/s, $N = 200$, and varying values of p.

This figure indicates that in the bad state, it is advantageous to transmit at the lowest rate possible. However, transmitting at a slow rate increases the waiting time for future packets. This tradeoff will become significant as the distortion threshold D_T is lowered, as illustrated by the experimental results presented below.

The effects of data rate adaptation are illustrated in Fig. 7. This figure depicts the convex hull of operational energy-distortion points for the first frame of the foreman sequence. These points were obtained by sweeping λ in (10). We consider $T_{\max} = 166.7$ ms, $p = 0.7$, and $N = 200$. The top two curves are for the case where the transmission rate is constant of either 100 or 300 kb/s. Notice that the higher the transmission rate the higher the required expected energy. Also note that the curve corresponding to 100 kb/s cannot meet the lower distortion thresholds under the delay constraints considered here. In the case where we allow data rate adaptation, energy savings can be obtained by increasing the transmission rate when the channel is in a good state and vice-versa decreasing the transmission rate when the channel is in a bad state. As we decrease the level of allowable distortion D_T, the curves come closer together. As we decrease the value of D_T, there is less opportunity to decrease the transmission rate and constant rate transmission policies become a better approximation to the optimal transmission policies.

Next, we take a closer look at the case with $\mathcal{C} = \{100, 300\}$ kb/s. Fig. 8 presents operational energy-distortion curves with varying values of p and with $N = 200$. As this figure indicates, increasing p leads to lower required energy. This is because the channel state is more likely to stay fixed longer during transmission of the entire packet. Hence, the initial channel state gives a better estimate of the energy required to send the packet.

Fig. 9 shows the expected fraction of packets that begin transmission in the good state. The case with $p = 0.5$ results in 50% of packets being sent in the good state. The curve corresponding to $p = 0.7$ shows the fraction of packets starting in the good state increases with D_T. Setting $p = 0.99$ results in a significant increase in the number of MBs transmitted in the good state.
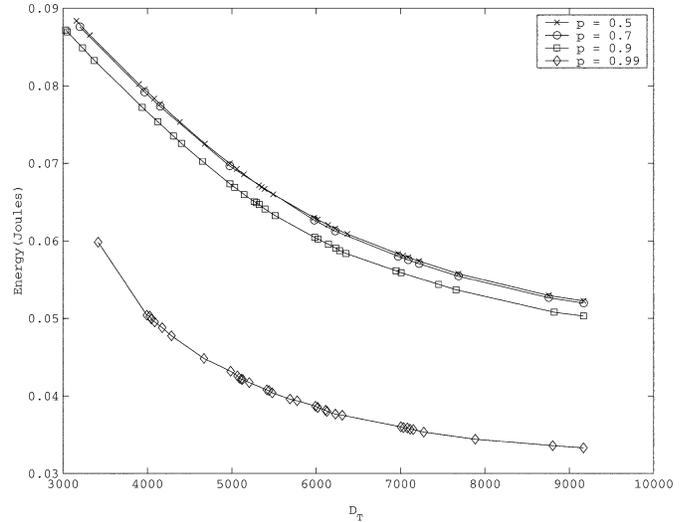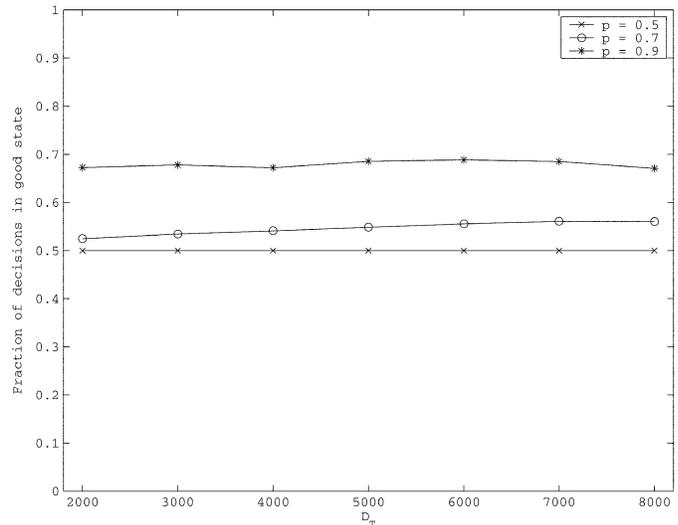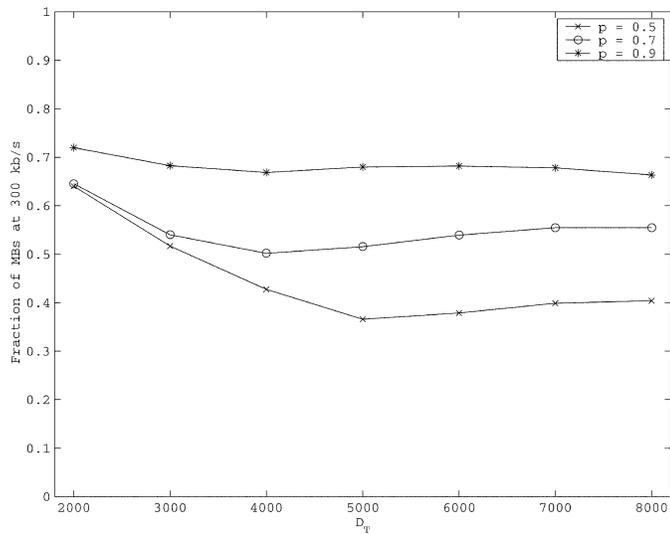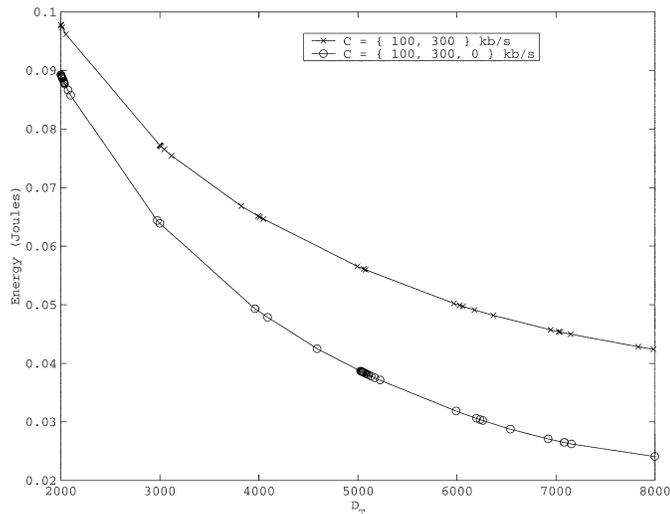

Fig. 9.   Expected fraction of MBs transmitted in the good channel state.

The expected fraction of packets transmitted using the high transmission rate is shown in Fig. 10. Comparing the curves for $p = 0.99$ in Figs. 9 and 10, we observe a very strong correlation between channel state and channel rate. In the case of $p = 0.7$, we observe the expected number of packets transmitted at high rate increase with increasing values of D_T for the range $D_T > 4000$. This correlates with the expected fraction of decisions in the good channel state increasing in Fig. 9. Note also in Fig. 10 that for values of $D_T < 4000$, the fraction of packets transmitted at the high rate increases with decreasing D_T in the range of $D_T < 4000$. The case with $p = 0.5$ shows similar behavior in Fig. 10.

2) Packet Transmission Scheduling: In this experiment, we illustrate the effect of packet transmission scheduling on the performance of the system. We consider the same situation as above. In Fig. 11, the convex hull of operational energy-distortion points is shown for $\mathcal{C} = \{100, 300\}$ kb/s with

Fig. 10.   Expected fraction of MBs transmitted at $C = 300$ kb/s.

Fig. 11.   Convex hull of energy-distortion operational points. The system with scheduling has $C = \{100, 300, 0\}$ kb/s.

Fig. 12.   Convex hull of energy-distortion operational points for system with different packetization schemes.

and without scheduling. We can see that the policies obtained with scheduling outperform the policies without scheduling.

The form of scheduling considered here consists of waiting for a channel time slot and then reconsidering the choice of transmission rate. If the channel is in a good state, waiting will result in a channel state transition and, therefore, there is a positive probability that the channel will be in the bad state when transmission begins. If transmission starts in the bad channel state, then a higher amount of transmission energy will be required. This can be seen by comparing Figs. 5 and 6. Therefore, waiting is not advantageous when the channel is in the good state. Similarly, waiting is advantageous when the channel is in the bad state.

Note that as the allowable distortion D_T decreases the curves come closer together. This is because as D_T is lowered the system has less opportunity to wait for better channel conditions. Thus, the policies that do not include scheduling becomes a better approximation to the optimal policy.
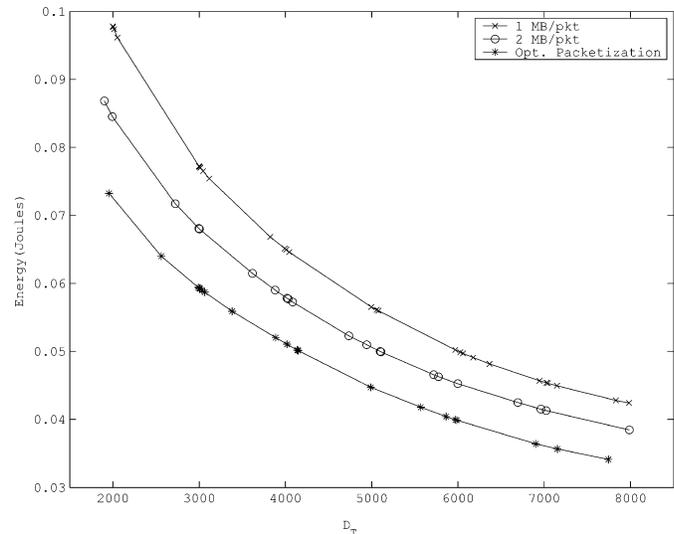
3) Packetization: Next, we present the effect of packetization on the solution of the problem. Increasing the number of MBs in a video packet, results in a form of scheduling, introduced by the packetization delay. Also, the amount of overhead required by the encoder/decoder is reduced. Each video packet can be independently processed by the encoder/decoder. To do this, additional information must be included in each video packet. Also, differentially encoding MBs within a packet gives more efficient R-D tradeoffs.

The convex hull of operational energy-distortion points are shown in Fig. 12 for a scheme which uses 1 MB/packet, 2 MB/packet, and one that optimizes between the two. The scheme with 2 MBs per packet is more energy efficient than the 1 MB/packet scheme. This is a result of increased source coding efficiency and the introduction of a limited form of scheduling. Recall that increasing the number of MBs in a packet introduces additional waiting time which may lead to more efficient energy-distortion tradeoffs. The proposed approach can choose between 1 or 2 MBs per packet by taking into account the channel state and the delay constraint for the MBs. The results of the optimization presented here indicate that as the level of allowable distortion decreases, the advantage of having a variable packetization scheme increases.

Fig. 13 shows the effect of packetization and scheduling combined. The system with scheduling and fixed packetization performs better for high values of D_T than the system with optimal packetization and no scheduling. As D_T is lowered, the advantage of packetization becomes more pronounced.

C. Experiment III

In the experiments presented so far, we have studied the optimal coding and transmission policies for a single frame. We now present results that illustrate the solution for multiple frames of a video sequence. We compare the energy consumption of our proposed approach to the energy required to transmit at a fixed rate an encoded sequence produced by a rate-distortion optimized MPEG-4 encoder with rate control.
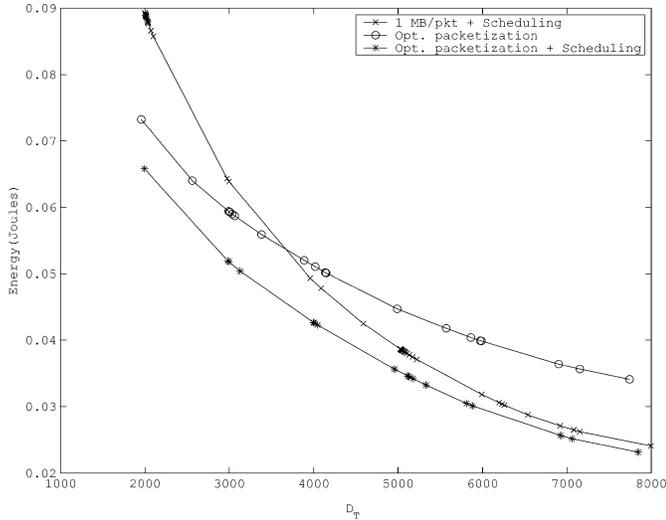
Fig. 13.   Convex hull of energy-distortion operational points for systems including packetization and scheduling.

Fig. 14.   Required expected energy for each frame.

Fig. 15.   Distortion per frame.

The reference system considered here consists of a TM5 rate controller and a R-D optimized video encoder. The rate controller determines a bit budget for each video frame R_{budget}. Given this bit budget, the R-D optimized encoder minimizes the total distortion of the frame subject to a rate constraint given by the bit budget. The resulting distortion is used as a distortion constraint in the proposed system. The deadline for arrival at the decoder buffer is given by the number of bits used to encode the video frame and the transmission rate of the channel C assumed fixed here. Therefore, the transmission delay for the video frame is given by $\sum_{k=0}^{M-1} B(k)/C$. From this transmission delay, we derive a delay constraint for each MB in the video frame. This delay constraint is given as the following.

In the proposed system, for each frame, the distortion and delay constraints are given by the reference system as described above. The formulation presented here allows the solution of the problem to be carried out for the whole video sequence. However, for computational reasons, we solve the problem independently for each video frame in the video sequence given the distortion and delay constraints of the reference system. For the reference system, we consider constant rate transmission with $C = 200$ kb/s and a 2 s buffer. We compare this reference system to a system that uses data rate adaptation and packet transmission scheduling. In the proposed system, we consider $\mathcal{C} = \{100, 200, 300\}$ kb/s with and without the option of scheduling. For each frame, we must find the appropriate value of the Lagrange multiplier λ.

The results for this experiment are shown in Figs. 14 and 15. In Fig. 14, we show the total expected energy required to transmit each video frame. The corresponding distortion levels for each frame are shown in Fig. 15. We can see from these figures that the system proposed here can match the PSNR for each frame of the reference system but uses significantly less transmission energy.

Using data rate adaptation yields average energy savings of 30.9% with respect to the system with constant rate transmission. The system that in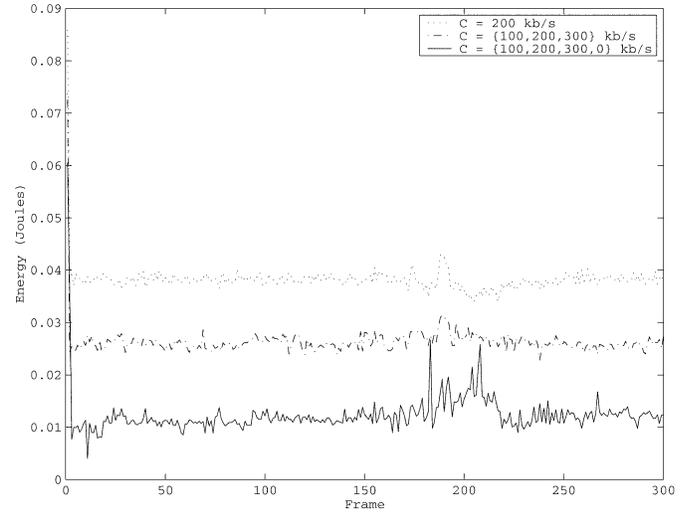corporates scheduling yields sav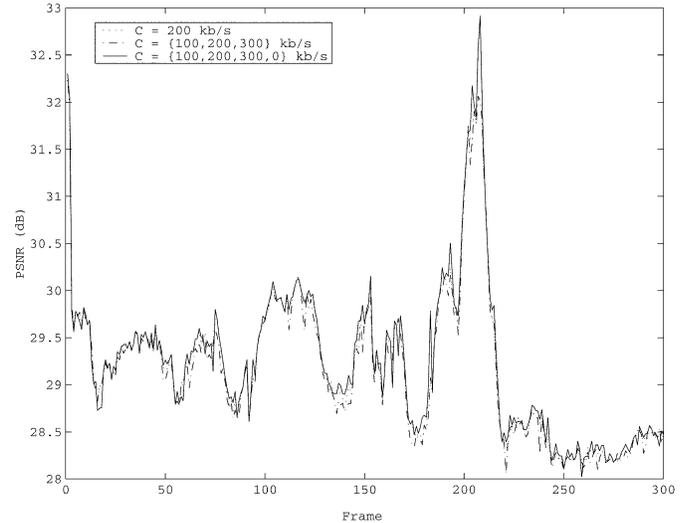ings of 67.7% on average. Note that around frame 200 there is a dip in the savings rate for both systems. This dip corresponds to a peak in the PSNR obtained by the reference system and, thus, imposes tighter constraints on the optimization. Note that the relative advantage of transmission scheduling is also reduced in this region.

V. CONCLUSION

In this paper, we have considered energy efficient wireless video streaming. The goal is to transmit a video sequence using the minimum required transmission energy subject to the video quality and delay constraints from the streaming application. Our formulation considers the tradeoffs in the selection of source coding parameters, packetization, transmitter adaptation (power and rate), and packet transmission scheduling. Two algorithms based on DP techniques have been presented. The experimental results presented here illustrate the tradeoffs involved.

The complexity of the resulting optimization problem can be very large. However, the reduction in required transmission energy makes this an attractive approach. A topic of further research is the design of more efficient algorithms to solve the problem. Additionally of great interest would be low complexity suboptimal approaches that can be devised based on the results of the optimization.

We have presented a comparison of our system to a reference system that transmits at a constant rate. In this reference system, the source coding decisions are done separately from the transmission decisions. The simulation results show that our system can achieve over 60% reduction in the required transmission energy. A direction for future research is the design of rate controllers based on the delay requirements of the video application in order to fully exploit the benefits systems with transmitter adaptation.

REFERENCES

[1] Y. Wang, G. Wen, S. Wenger, and A. K. Katsaggelos, "Error resilient techniques for video comunications," *IEEE Signal Processing Mag.*, vol. 17, pp. 61–82, July 2000.

[2] M. C. Hong, L. Kondi, H. Scwab, and A. K. Katsaggelos, "Video error concealment techniques," *Signal Processing: Image Commun.*, vol. 18, pp. 437–492, 1999.

[3] C. Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 756–773, May 1999.

[4] R. O. Hinds, T. N. Pappas, and J. S. Lim, "Joint block-based video source/channel coding for packet-switched networks," in *Proc. SPIE*, vol. 3309, Jan. 1998, pp. 124–133.

[5] R. O. Hinds, "Robust Model Selection for Block-Motion-Compensated Video Encoding," Ph.D. dissertation, MIT, Cambridge, MA, 1999.

[6] G. Côté, 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.

[7] 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.

[8] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[9] J. M. Rulnick and N. Bambos, "Mobile power management for wireless communication networks," *Wireless Networks*, vol. 3, no. 1, pp. 3–14, 1997.

[10] N. Bambos, "Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects," *IEEE Pers. Commun. Mag.*, vol. 5, pp. 50–59, June 1998.

[11] R. Berry, "Some dynamic resource allocation problems in wireless networks," presented at the SPIE ITCom, Denver, CO, Aug. 20–24, 2001.

[12] B. E. Collins and R. L. Cruz, "Transmission policies for time varying channels with average delay constraints," presented at the Allerton Conf. Communication Control and Computers, Monticello, IL, 1999.

[13] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *Trans. Inform. Theory*, vol. 48, pp. 1135–1149, May 2002.

[14] A. Chockalingam and M. Zorzi, "Energy efficiency of media access protocols for mobile data networks," *IEEE Trans. Commun.*, vol. 46, pp. 1418–1421, Nov. 1998.

[15] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. C. Chen, "A survey of energy efficient network protocols for wireless networks," *Wireless Networks*, vol. 7, no. 4, pp. 343–358, July 2001.

[16] R. Kravets and P. Krishnan, "Application-driven power management for mobile communication," *Wireless Networks*, vol. 6, no. 4, pp. 263–277, Sept. 2000.

[17] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)*, Spec. IEEE 802.11 Standard, 1998.

[18] B. Prabhakar, E. Uysal-Biyikoglu, and A. El Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. IEEE INFOCOM 2001*, vol. 1, Anchorage, AK, Apr. 22–26, 2001, pp. 386–394.

[19] A. El. Gamal, C. Nair, B. Prabhakar, E. Uysal-Biyikoglu, and S. Zahedi, "Energy-efficient scheduling of packet transmissions over wireless networks," in *Proc. INFOCOM 2002*, June 2002, pp. 1773–1782.

[20] Y. S. Chan and J. W. Modestino, "A joint source coding and power control approach for maximizing the capacity of CDMA wireless networks supporting heterogeneous video traffic," presented at the 11th Packet Video Workshop, Kyongju, Korea, 2001.

[21] ——, "Transport of scalable video over CDMA wireless networks: A joint source coding and power control approach," in *Proc. Int. Conf. Image Processing (ICIP)*, Thesaloniki, Greece, May 2001, pp. 973–976.

[22] Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Joint source coding and transmission power managment for energy efficient wireless video communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 8, pp. 411–424, June 2002.

[23] Y. Eisenberg, F. Zhai, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Variance-aware distortion estimation for wireless video communications," in Proc. ICIP, Barcelona, Spain, Sept. 2003.

[24] S. Nanda, K. Balachandran, and S. Kumar, "Adaption techniques in wireless packet data services," *IEEE Commun. Mag.*, vol. 38, pp. 54–64, 2000.

[25] E. Biglieri, C. Caire, and G. Taricco, "Coding and modulation under power constraints," *IEEE Pers. Commun.*, vol. 5, pp. 32–39, June 1998.

[26] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.

[27] H. S. Wang and N. Moayeri, "Finite-state Markov channel a useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 163–171, Feb. 1995.

[28] C. Luna, Y. Eisenberg, T. Pappas, R. Berry, and A. K. Katsaggelos, "Transmission energy minimization in wireless video streaming applications," in *Proc. 35th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2001, pp. 185–189.

[29] C. E. Luna, Y. Eisenberg, T. Pappas, R. Berry, and A. K. Katsaggelos, "Joint source coding and data rate adaptation for energy efficient wireless video streaming," presented at the IEEE Int. Packet Video Workshop, Apr. 24–26, 2002.

[30] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995, vol. 1.

[31] G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding*. Norwell, MA: Kluwer, 1997.

Carlos E. Luna (S'02–A'03) received the B.S. degree in electrical engineering from The Johns Hopkins University, Baltimore, MD, in 1993, and the M.S. and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, IL, in 1996 and 2002, respectively.

His current research interests include video transmission over packet networks, video compression, and transmission power management for wireless networks.

Yiftach Eisenberg (S'02) received the B.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, in 1999, and the M.S. degree in electrical engineering from Northwestern University, Evanston, IL, in 2001, where he is currently working toward the Ph.D. degree.

His research interests include efficient resource allocation in video communication systems and video quality metrics. He is exploring how video quality is related to energy consumption in wireless video communications.

Randall Berry (S'93–M'00) received the B.S. degree in electrical engineering from the University of Missouri-Rolla, in 1993, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1996 and 2000, respectively.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL. In 1998, he was on the technical staff at MIT Lincoln Laboratory, Advanced Networks Group. His primary research interests include wireless communication, data networks, and information theory.

Dr. Berry is the recipient of a 2003 NSF CAREER Award.

Thrasyvoulos Pappas (M'87–SM'95) received the S.B., S.M., and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1979, 1982, and 1987, respectively.

From 1987 to 1999, he was a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ. In September 1999, he joined the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL, as an Associate Professor. His research interests are in image and multidimensional signal processing. His recent work has been on perceptual image coding, video transmission over lossy channels, model-based halftoning, image and video segmentation, video processing for sensor networks, and audiovisual signal processing.

Dr. Pappas is the Electronic Abstracts Editor and an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING. He is the Chair of the IEEE Image and Multidimensional Signal Processing Technical Committee and a Member of the Multimedia Signal Processing Technical Committee. He served as Technical Program Co-chair for ICIP-2001, Thessaloniki, Greece. He is also Co-chair for the Conference on Human Vision and Electronic Imaging, sponsored by The International Society for Optical Engineers (SPIE) and IST.

Aggelos K. Katsaggelos (S'80–M'85–SM'92–F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively, both in electrical engineering.

In 1985, he joined the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL, where he is currently a Professor, holding the Ameritech Chair of Information Technology. He is also the Director of the Motorola Center for Communications and a Member of the Academic Affiliate Staff, Department of Medicine at Evanston Hospital. He is on the Editorial Board of Academic Press, Marcel Dekker: Signal Processing Series, Applied Signal Processing, and Computer Journal. He is the Editor of *Digital Image Restoratino* (New York: Springer-Verlag, 1991), coauthor of *Rate-Distortion Based Video Compression* (Norwell, MA: Kluwer, 1997), and coeditor of *Recovery Techniques for Image and Video Compression and Transmission* (Norwell, MA: Kluwer, 1998). He is the co-inventor of eight international patents.

Dr. Katsaggelos is a Member of the Publication Board of the IEEE Proceedings, the IEEE Technical Committees on Visual Signal Processing and Communications, and Multimedia Signal Processing, He has served as Editor-in-Chief of the IEEE *Signal Processing Magazine* (1997–2002). He is a Member of the Publication Boards of the IEEE Signal Processing Society, the IEEE TAB Magazine Committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1990–1992), an Area Editor for the journal *Graphical Models and Image Processing* (1992–1995), a Member of the Steering Committees of the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1997) and the IEEE TRANSACTIONS ON MEDICAL IMAGING (1990–1999), a Member of the IEEE Technical Committee on Image and Multidimensional Signal Processing (1992–1998), and a Member of the Board of Governors of the IEEE Signal Processing Society (1999–2001). He is the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), and an IEEE Signal Processing Society Best Paper Award (2001).