TRANSMISSION ENERGY MINIMIZATION IN WIRELESS VIDEO STREAMING APPLICATIONS

Carlos E. Luna, Yiftach Eisenberg, Thrasyvoulos Pappas, Randall Berry,
Aggelos K. Katsaggelos

Department of Electrical and Computer Engineering
Northwestern University
2145 Sheridan Road, Evanston, IL 60208
Phone: (847) 491-3039
Fax: (847) 491-4455
Email: {carlos,yeisenbe,pappas,rberry,aggk}@ece.nwu.edu

ABSTRACT

A major limitation of wireless communications is that mobile users depend on a battery with limited energy supply. Efficiently utilizing this energy is a key consideration in designing wireless networks. We consider a situation where a video sequence is to be transmitted over a wireless channel. The goal is to limit the amount of distortion in the received video while using the minimum required transmission energy and satisfying the delay constraints resulting from the streaming application. To accomplish this goal we jointly consider source coding and dynamic allocation of physical layer communication resources in a novel framework. In this setting, we formulate an optimization problem that corresponds to minimizing the transmission energy required to achieve an acceptable level of distortion subject to a delay constraint.

1. INTRODUCTION

A major limitation in wireless networks is that mobile users must rely on a battery with a limited supply of energy. Efficiently utilizing this energy is a key consideration in the design of wireless networks [1, 2]. In this paper, we consider a situation where a mobile user is transmitting a video sequence over a wireless channel. Previous work in this area has concentrated either on error resilient source coding methods (see [3] for a recent survey) or on channel coding to ensure reliable transmission of all the bits. In this paper, we jointly consider source coding and dynamic allocation of physical layer communication resources in order to transmit video at an acceptable quality using the minimum possible transmission energy, while satisfying delay constraints.

We consider video transmission for real-time display. 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.

In Fig. 1 a block diagram of the system considered here is shown. The video encoder takes raw video and produces a stream of video packets that are to be transmitted over a wireless channel. These video packets are buffered and then transmitted over the wireless channel. The transmitter (Tx) can dynamically allocate communication resources at the physical layer to 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 (see for example [4] for a survey of the currently available techniques). 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. The key point is that display proceeds continuously and that packets that do not arrive by their display time are considered lost.

The rest of this paper is organized as follows: In the next section we present the problem formulation in detail. In Section 3, we present the algorithm we propose to solve this problem. Section 4 presents experimental results. Section 5 concludes the paper.

2. PROBLEM FORMULATION

We consider a system where video is encoded using a block based motion compensated technique to produce a stream of video packets (e.g. MPEG-4). Each video packet is transmitted through the wireless channel and processed independently by the decoder. The size of the packets and their relationship to slices or group of blocks (GOBs) is not explored in this paper. Instead, we will initially consider a simple packetization scheme where each packet is made up of a single macro block (MB). We therefore use the terms MB and video packet interchangeably.

Let M be the number of MBs in a video frame and k the MB index. The original video frame arrives at the encoder as a stream of MBs spaced every T_{MB} seconds. We may express T_{MB} as a function of the frame rate and the number of MBs in a frame, that is, $T_{MB} = T_f/M$, where T_f is the duration of one frame in seconds. The video frame must experience a constant end-to-end delay which we denote by T_{max}. That is, if the frame is captured at time t, it must be received at the decoder buffer before time $t+T_{max}$. Thus, we impose a maximum delay constraint of T_{max} on each MB (refer to Fig. 2). Note that if a MB is not present at the decoder buffer before the delay of T_{max} it will be treated as lost.

The diagram in Fig. 2 illustrates the various components of the end-to-end delay for a video packet. In this figure, we assume that the encoding and decoding delay as well as the propagation delay through the wireless channel are constant and can be ignored. We only need to concern ourselves with the delay at the encoder buffer and the transmission delay. The delay experienced by each MB depends on the number of bits used to encode the MB, the channel rate at which we transmit the MB and also on how long the MB must wait in the encoder buffer before transmission can begin.

Our goal is to assign to each MB a choice of coding mode and quantization step size, at the source coding level, and a channel rate and transmission power at the physical layer in order to obtain good video quality. For a video packet, k, of size $B(k)$ bits that is transmitted at rate $C(k)$ we express the delay as,

$$\delta(k) = w(k) + \frac{B(k)}{C(k)}, \qquad (1)$$

where $w(k)$ is the amount of time the packet must wait in the buffer before transmission; we refer to this as the waiting time. The second term in Eq. (1) is the transmission delay for MB k. This is the amount of time it takes to transmit $B(k)$ bits at a channel rate of $C(k)$ bits per second. The waiting time depends on

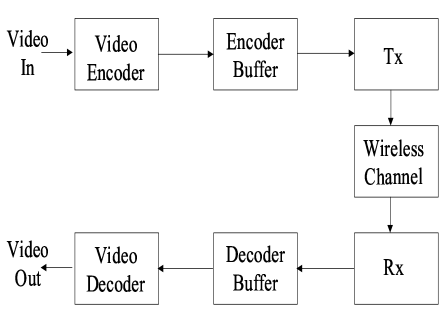

Fig. 1. System block diagram.

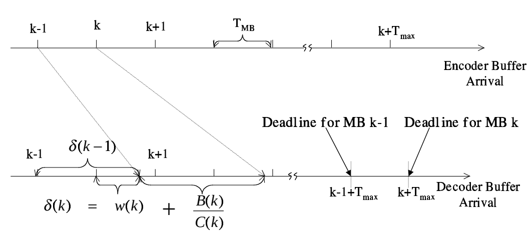

Fig. 2. Timing diagram.

the delay for the previous MB, $k-1$. Thus we can write

$$w(k) = (\delta(k-1) - T_{MB})^+, \qquad (2)$$

where $(x)^+ = x$ for $x > 0$ and 0 otherwise.

Now we turn our attention to the allocation of communication resources at the physical layer. We assume that each video packet is sent over a slowly varying wireless channel with frequency nonselective fading. We model the channel over which the kth packet is sent as a band-limited additive white Gaussian noise channel with gain $\sqrt{h(k)}$. We assume that the gain stays fixed during the transmission of the kth packet and randomly varies between packets in a manner that will be described later.[1] This gain is assumed to be known at the transmitter and the receiver. If the desired transmission rate for the kth packet is $C(k)$, we assume that the required transmission power is the minimum power such that the channel over which this packet is sent has capacity $C(k)$, i.e.,

$$P(h(k), C(k)) = \frac{N_0 W}{h(k)} \left(2^{\frac{C(k)}{W}} - 1 \right), \quad (3)$$

where W is the bandwidth of the channel and N_0 is the power spectral density of the noise. From Shannon's coding theorem, Eq. (3) 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. Thus the amount of energy required to transmit video packet k of size $B(k)$ bits at a rate of $C(k)$ bits per second can be expressed as

$$E(k) = P(h(k), C(k)) \frac{B(k)}{C(k)}. \qquad (4)$$

In this paper, we assume that the transmitter has perfect channel side information (CSI), i.e., $h(k)$ is known exactly. Furthermore, the fading coefficient remains constant for the duration of the transmission of the packet. We model the fading process by a finite state Markov chain with known transition probability matrix $\mathbf{A}$. We assume that the channel state transitions according to $\mathbf{A}$ after each packet transmission is complete.

[1] Of course this is a somewhat simplified assumption, because the transmission time of a packet varies from packet to packet. The ideas presented here can be extended to a more realistic model that takes this into account.

3. PROPOSED ALGORITHM

The problem addressed in this paper is that of transmitting a frame of video using the minimum amount of energy subject to quality and delay constraints imposed by the application.

Macroblock k is coded using a quantizer $q(k)$ from a finite set of allowable quantizers $\mathbb{Q}$ resulting in a video packet of $B(k)$ bits with distortion $D(k)$. This video packet is transmitted at rate $C(k)$ bits per second chosen from a finite set of allowable channel rates $\mathbb{C}$. We want to select a quantizer, $q(k)$, and channel rate, $C(k)$, for each MB 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 MB constraint. The expectations are taken with respect to the channel state denoted by the random process $h(k)$. We pose this as an optimization problem given below:

$$\min_{q(k),C(k)} \mathbb{E}_H \left\{ \sum_{k=0}^{M-1} E(k) \right\} \qquad (5)$$

subject to:

$$\mathbb{E}_H \left\{ \sum_{k=0}^{M-1} D(k) \right\} \leq D_T \qquad (6)$$

$$\delta(k) \leq T_{max}, \forall k, \qquad (7)$$

where $\delta(k)$ and $E(k)$ are given by Eqs. (1) and (4) respectively. The initial conditions $w(0)$ and $h(0)$ are the initial wait time and the initial channel state, respectively.

We introduce a Lagrange multiplier $\lambda > 0$ and solve the following relaxed problem:

$$\min_{q(k),C(k)} \mathbb{E}_H \left\{ \sum_{k=0}^{M-1} [E(k) + \lambda D(k)] \right\} \quad (8)$$

subject to the delay constraints of Eq. (7). This relaxed problem can be solved using techniques from Dynamic Programming

(DP) [5]. By appropriately choosing λ, the problem of Eqs. (5) - (7) can be solved within a convex-hull approximation by solving Eq. (8) and (7) [6].

3.1. DP Solution of Relaxed Problem

We define the system state as

$$\mathbf{x}(k) = \begin{bmatrix} w(k) \\ h(k) \end{bmatrix}, \tag{9}$$

where $w(k)$ is given in Eq. (2) and $h(k)$ is the channel state at time k.

Let $\mathbb{U}(\mathbf{x}(k))$ be the set of feasible choices of quantizers and channel rates when the system is in state $\mathbf{x}(k)$, i.e., the delay constraint of Eq. (7) is not violated. For $u(k) \in \mathbb{U}(\mathbf{x}(k))$, the cost incurred by MB k is given by,

$$g(\mathbf{x}(k), u(k)) = E(k) + \lambda D(k). \tag{10}$$

We solve the problem of Eq. (8) using DP. We start the algorithm at $k = M - 1$, that is,

$$J^*_{M-1}(\mathbf{x}(M-1)) = \min_{\mathbb{U}(\mathbf{x}(M-1))} \{E(M-1) + \lambda D(M-1)\} \tag{11}$$

is calculated. As it will be explained later, the values of $w(k)$ are quantized and therefore a fixed number of states is considered at each stage of the DP algorithm. Then for $k = M-2, \ldots, 0$, we recursively define

$$J^*_k(\mathbf{x}(k)) = \min_{\mathbb{U}(\mathbf{x}(k))} \mathbb{E}_H \{g(\mathbf{x}(k, u(k)) + J^*_{k+1}(\mathbf{x}(k+1))\}. \tag{12}$$

In carrying out (12) all combinations of $q(k)$ and $C(k)$ belonging to $\mathbb{U}(\mathbf{x}(k))$ are considered. This optimization clearly eliminates all branches but one emanating from each state of the trellis. 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 Eq. (8).

In the development above we have assumed that the set $U(\mathbf{x}(k))$ is not empty. If this set is empty, for some $\mathbf{x}(k)$, then we let $J^*_k(\mathbf{x}(k)) = \infty$.

Equation (11) can be extended to include a terminal cost to account for the waiting time imposed on the first MB of the next frame. In this manner, we can account for future frames in the video sequence.

In the problem above, $w(k)$ is continuous. We may approximate the solution to the problem by quantizing $w(k)$ and then applying DP to solve the approximate problem optimally [5]. Let $\mathbb{S}$ be a finite subset of the non-negative real numbers given by

$$\mathbb{S} = \{s_0, \ldots, s_N\}, \tag{13}$$

with $s_l = (l \times T_{max})/N$. Then we have

$$w(k+1) = \lceil w(k) + \frac{B(k)}{C(k)} - T_{MB} \rceil_{\mathbb{S}}, \tag{14}$$

where

$$\lceil w(k) \rceil_{\mathbb{S}} = \min\{s \in \mathbb{S} \mid s \geq w(k)\}. \tag{15}$$

Using this new definition of $w(k)$, we apply the DP algorithm in Eq. (12) 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.

4. EXPERIMENTS

In this section we present experimental results that will serve to illustrate the tradeoffs discussed in this paper. We consider the transmission of the first frame of the foreman sequence in QCIF format with maximum delay $T_{max} = 5T_f$. In these experiments we set $N = 200$ in Eq.(13). The video packets are produced by an MPEG-4 encoder operating in Intra mode. We consider a set of eight available quantization step sizes given by $\mathbb{Q} = \{2, 4, 8, 16, 20, 24, 28, 31\}$.

We consider a wireless channel with bandwith $W = 500kHz$ and additive white gaussian noise with variance $N_0 W = 0.39$. The fading is modeled by a two state Markov chain with state space $\mathbb{H} = \{0.81, 0.01\}$. We use a symmetric transition probability matrix of the form

$$\mathbf{A} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}, \tag{16}$$

with $p = 0.7$.

By sweeping λ in Eq. (8), we obtain the convex-hull of Energy-Distortion operational points. These results are shown in Fig. 3 for $\mathbb{C} = \{300, 100\}\ kb/s$, and constant rate transmission at 100 kb/s and 300 kb/s. For low video quality (high MSE), we choose to transmit at low rates more often and the performance of the variable rate curve matches that of the 100 kb/s curve. As the quality improves (lower MSE), we need to transmit more often at the high rate, thus the 300 kb/s curve approaches the variable rate curve. In the intermediate MSE region, the variable rate strategy requires less energy than either of the constant rate strategies.

The advantage afforded by the ability to choose the channel rate per packet depends greatly on the nature of the relationship between power and rate given by Eq. (3). For example, when W is large with respect to the transmission rate, this relationship is almost linear in $C(k)$. In this case, it is nearly optimal to transmit at the highest available rate regardless of the channel state.

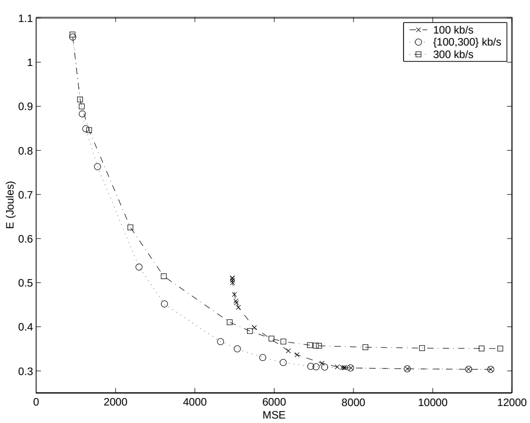

Fig. 3. Convex hull of Energy-Distortion operational points.

5. CONCLUSIONS

In this paper we have considered tradeoffs between transmission energy and image quality in a wireless video transmission application. Using a simplified model, we have solved the problem of minimizing transmission energy subject to quality and delay constraints imposed by the application. Quantizing the values of the waiting time results in an optimization problem which does not grow exponentially.

Extensions of the simple model considered here are currently under consideration. In these extensions, we consider channel models that account for the variations in packet transmission delay. In this manner, we can provide a more realistic channel model. We also consider situations where the transmitter is allowed to wait before transmitting a packet. In this manner, the transmitter increases the waiting time for a packet in order to transmit when channel conditions are more favorable.

6. REFERENCES

[1] N. Bambos, "Toward power-sensitive network architectures in wireless communications: concepts, issues, and design aspects," *IEEE Personal Communications Magazine*, June 1998.

[2] R. Berry, S. Finn, R. Gallager, H. Kassab, and J. Mills, "Issues in low power wireless communication," in *Proc. 3rd ARL Fed. Lab. Symposium, (College Park, MD) Feb. 2-4*, 1998.

[3] Y. Wang, G. Wen, S. Wenger, and A.K. Katsaggelos, "Review of error resilient techniques for video comm.," *IEEE Signal Proc. Mag.*, vol. 17, pp. 61–82, July 2000.

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

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

[6] G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1997.