# A LAGRANGIAN OPTIMIZATION APPROACH TO RATE CONTROL FOR DELAY-CONSTRAINED VIDEO TRANSMISSION OVER BURST-ERROR CHANNELS

*Chi-Yuan Hsu and Antonio Ortega*

Integrated Media Systems Center, Department of Electrical Engineering-Systems
University of Southern California, Los Angeles, California 90089-2564
E-mail:{chiyuanh, ortega}@sipi.usc.edu

## ABSTRACT

We propose a rate control algorithm based on Lagrangian optimization for video transmission over burst-error channels. In our rate control approach, the delay and channel capacity constraints in the video transmission are translated into rate constraints at the encoder. Given that a feedback channel is available, the rate control mechanism can dynamically adjust the video encoding rate to meet the changing rate constraints as the channel conditions vary. Lagrangian optimization is used to find the optimal bit-allocation for the input video frames under the rate constraints, with the objective of minimizing the overall distortion at the decoder. We show how the performance of the transmission system, as measured in terms of the received video quality or the data loss rate, can be improved when information about the channel state is available and the encoder has an *a priori* probabilistic model of the channel behavior.

## 1. INTRODUCTION

Increasing the robustness of transmission over a link subject to channel errors usually requires packetizing the information and using an error control scheme, such as Forward error correction (FEC) or Automatic Repeat reQuest (ARQ), and is particularly challenging if the error characteristics are bursty. Using FEC may require a substantial amount of coding overhead to effectively protect the data, thus resulting in overall lower channel utilization rate, even during periods when transmission is relatively error-free. If a feedback channel is available, error control based on ARQ can be used as an alternative to FEC, with the advantage of incurring overhead only during high error periods. However, schemes such as ARQ may not be suitable for delay-constrained applications, such as video transmission, because the additional delay caused by packet retransmission may result in excessive end-to-end transmission delay for a particular packet. Here we argue that ARQ can be used in such a delay-constrained scenario if the error control mechanism is integrated with the encoder rate control algorithm, so that the source encoder can dynamically adjuts its output rate as a function of the channel error characteristics, e.g., reducing the encoding rate during periods of high bit error rate.

In this paper we thus consider video communication systems with integrated rate control ARQ-based error control, and study their performance for transmission over burst-error channels, as for example mobile radio channels subject to multipath fading. We start by formulating the rate constraints imposed on each block of encoded video due to the real time operation of the system and the available, and time varying, channel capacity as in in [1, 2]. We then use a two-state Markov chain to model the occurrence of channel errors, and show how such a model and the Selective Repeat (SR) ARQ feedback acknowledgements can be used by the encoder to estimate the expected channel rates, and modify accordingly the rate constraints. We proposed a solution based on dynamic programming for a similar rate control problem in [2] but here propose a solution-based on Lagrangian optimization with much lower complexity. Our simulation results demonstrate the performance improvements achievable by using channel feedback, as compared to "open loop" approaches.

## 2. DELAY AND RATE CONSTRAINTS

In typical video communications systems, the end-to-end delay $\Delta T$ each frame experiences (from capture time at the encoder to display time at the decoder) consists of the following delay components:

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

Under conditions of real time capture and playback, $\Delta T$ has to be constant, i.e., frames arriving at the decoder after their scheduled display time are useless. While the channel delay may be variable in general here we assume it to be constant (as would be approximately true in a point-to-point wireless communication environment). We also assume that the encoding and decoding delays are constant. Assume that $\Delta T_e$, $\Delta T_c$ and $\Delta T_d$ are known constants, then the only variable delay components will be $DeltaT_{eb}$ and $\Delta T_{db}$ and from (1) their sum will be constant and equal to:

$$\Delta T_{eb} + \Delta T_{db} = \Delta T - \Delta T_e - \Delta T_c - \Delta T_d \quad (2)$$

Given the duration of a frame interval, $T_f$, the total number of frames in either the encoder or decoder buffers, $\Delta N$, will also be constant,

$$\Delta N = \frac{\Delta T_{eb} + \Delta T_{db}}{T_f}. \quad (3)$$

### 2.1. Rate Constraints

The source encoding rate is constrained by the available channel capacity and the end-to-end delay $\Delta T$. We will use $T_p$, the time it takes to transmit one packet of data, as our basic time unit. Thus $t$ the time index will be an integer. One video frame spans $F$ packet intervals with $F = \frac{T_f}{T_p}$ and the $n$-th frame is encoded and released to the encoder buffer at time $n \times F$, or equivalently at time $t$ frame $n = \lfloor \frac{t}{F} \rfloor$ is the last frame that is encoded and released into the encoder buffer. Due to the delay constraint $\Delta N$, this frame has to be transmitted to the decoder by time $(n + \Delta N) \times F$ for decoding. We also assume that at time $t$, frame $m$ is the frame that is currently being transmitted through the channel. Define $R(i)$ as the number of bits used for encoding frame $i$, and $R'(m)$ as the number of bits of frame $m$ that are still in the encoder buffer and waiting for transmission at time $t$. Denote $C(k)$ as the number of bits transmitted by the channel at time $k$. The condition for frame $i$, $i \in \{m+1, \ldots, n\}$, to arrive at the decoder in time for decoding is that all the data of frame $i$, as well as the previous frames in the encoder buffer, has to be transmitted by the due time $(i + \Delta N) \times F$, thus (see also Fig. 1):

$$R'(m) + \sum_{j=m+1}^{i} R(j) \leq \sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \qquad (4)$$

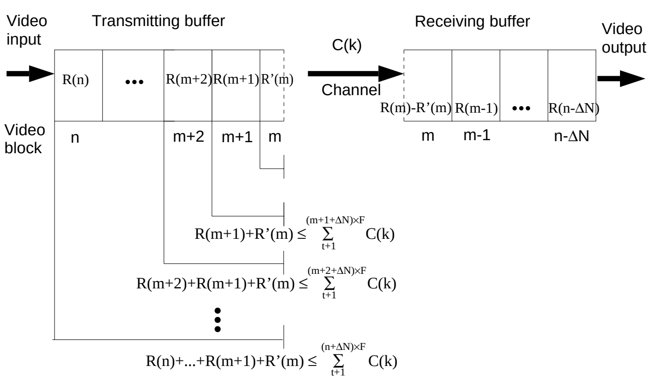



Figure 1: Rate constraints on the data in encoder buffer.

We assume that if part of frame $m$ has been transmitted the encoding rate for frame $m$ can no longer be changed, and thus we are interested in the rate constraints for remaining frames ($m + 1$ to $n$) in the encoder buffer. (1) can be rewritten by moving $R'(m)$ to the RHS as:

$$\sum_{j=m+1}^{i} R(j) \leq \left[ \sum_{k=t+1}^{(i+\Delta N)\times F} C(k) \right] - R'(m), \forall i \in \{m+1, \ldots, n\}. \quad (5)$$

In our rate control implementation, we assume that the encoding rates of those video frames which are still in the encoder buffer can be dynamically adjusted before transmission. In a DCT-based video compression scheme, a possible implementation of the system (see Fig. 2) would consist of having data quantized with different quantizers stored in separate buffers, so that the transmitter can select the data source from one of the buffers according to the rate control. Alternative approaches are possible, including those involving embedded bitstreams.

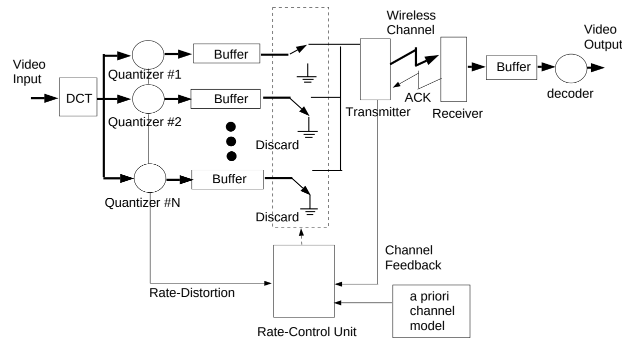
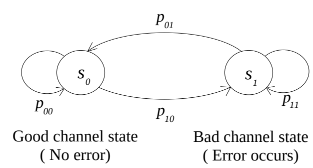



Figure 2: System block diagram.




| | Downlink | Uplink |
|---|---|---|
| $p_{00}$ | 0.99896 | 0.96618 |
| $p_{01}$ | 0.00103 | 0.03382 |
| $p_{01}$ | 0.17201 | 0.46945 |
| $p_{01}$ | 0.82799 | 0.53055 |

Figure 3: Markov chain model and transition probabilities.

Because of the end-to-end delay constraint, an erroneous packet need not be be retransmitted if the display time for all the data in that packet has been exceeded. In this worst case scenario the video data will be considered lost. We assume that the lost data in one block is replaced by its corresponding DC value at the decoder end (assuming that the DC value can be transmitted reliably by using FEC with high protection capability).

## 3. CHANNEL CHARACTERIZATION

Among many possible burst-error channel environments, here we focus on a a wireless CDMA spread spectrum system, consisting of two radio links, namely uplink (mobile-to-base) and downlink (base-to-mobile). Previous studies [3] show that a first-order Markov chain, such as the two-state Markov model by Gilbert [4] and Elliot [5], can provide a good approximation in modeling the error process at the packet level in fading channels. Here we use a simple two state model but other models, e.g. the $N$-state Markov model used in our previous work [2, 6], could also be used within our framework. In our model (see Fig. 3), 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 $s_1$. $p_{ij}$ for $i, j \in \{0, 1\}$ are the transition probabilities.

As shown in (5) the video encoder needs to comply with rate constraints that depend on the future channel rates. However in a lossy environment future channel rates are obviously not known and, thus we will have to estimate their values based on the current channel state and our model of channel behavior. Given the chosen Markov model with transition probability matrix:

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \tag{6}$$

we can define the state probabilities $\pi(k) = [\pi_0(k), \pi_1(k)]$ as the probabilities for the channel to stay in state $s_0$ and $s_1$ at time $k$ respectively. Assume that the observed channel state at time $t$ is $S(t)$, then the initial state probability at time $t$ can be written as:

$$\pi_n(t) = \begin{cases} 1, & \text{when } S(t) = s_n; \\ 0, & \text{otherwise.} \end{cases}, \quad \forall\, n \in \{0, 1\}. \tag{7}$$

The state probabilities $\pi(k)$ at time $k$ can be derived from $\pi(t)$ and $P$ as:

$$\pi(k) = \pi(t) \cdot \mathbf{P}^{k-t}. \tag{8}$$

Therefore the expected channel rate $E[C(k)]$ at time $k$ can be calculated as:

$$E[C(k)] = \bar{C} \times \pi_0(k) \tag{9}$$

and the approximated rate constraints can be derived by replacing $C(k)$ by the expected value $E[C(k)]$ in (5) as:

$$\sum_{j=m+1}^{i} R(j) \leq \left[ \sum_{k=t+1}^{(i+\Delta N)\times F} E[C(k)] \right] - R'(m), \tag{10}$$
$$\forall i \in \{m+1, \ldots, n\}.$$

## 4. OPTIMAL RATE CONTROL

Assume that video frames are encoded by quantizers from a quantizer set $\mathcal{Q}$. Denote $x_i \in \mathcal{Q}$ as the choice of quantizer for encoding frame $i$, and $R_{x_i}(i)$ and $D_{x_i}(i)$ as the associated encoding rate and distortion. The goal of rate control is to find a set of quantizer choices $\mathbf{x} = \{x_{m+1}, \ldots, x_n\}$ that will minimize the distortion for the blocks in the encoder buffer (given the expected channel rate):

**Formulation 1** *Find quantizer choice* $\mathbf{x}^*$ *at time* $t$ *such*

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_{j=m+1}^{n} D_{x_j}(j), \quad \textit{where } n = \left\lfloor \frac{t}{F} \right\rfloor. \tag{11}$$

*subject to the constraint set (10):*

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 R-D data of future frames. The optimization to solve Formulation 1 is performed on a sliding window basis, with quantizer choices possibly changing every time the window moves (due to the change in the video blocks considered, as well as changes in channel state). Thus the proposed solution cannot claim to achieve overall optimality since we perform a greedy allocation at each time $t$.

### 4.1. Lagrangian-based Solution

Using Lagrangian optimization for rate control under multiple rate constraints was previously studied in [7, 8]. In the Lagrangian optimization approach, the constrained optimization problem in Formulation 1 is equivalent to the unconstrained problem derived by introducing a non-negative Lagrange multiplier $\lambda_i$ associated with each rate constraint in (10) as:

**Formulation 2** *Find the quantizer choice* $\mathbf{x}^*$ *at time* $t$ *such*

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_{j=m+1}^{n} D_{x_j}(j) + \sum_{i=m+1}^{n} \lambda_i \cdot \left( \sum_{j=m+1}^{i} R_{x_j}(j) \right), \tag{12}$$

where we introduce $n-m$ Lagrange multipliers to replace the $n-m$ constraints in Formulation (11). Define $\lambda'_i$ as:

$$\lambda'_i = \sum_{j=i}^{n} \lambda_j, \quad \forall i \in \{m+1, \ldots, n\}. \tag{13}$$

then (12) can be rearranged as:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_{j=m+1}^{n} \left( D_{x_j}(j) + \lambda'_j \cdot R_{x_j}(j) \right) \tag{14}$$

Since $\lambda_{m+1}, \ldots, \lambda_m$ are all non-negative values, from (13) we have $\lambda'_{m+1} \geq \lambda'_{m+2} \geq \ldots \geq \lambda'_n$. The problem then is to find out the appropriate values of each $\lambda_i$ such that no constraint is violated. Because the mapping $\{\lambda_{m+1}, \ldots, \lambda_t\} \rightarrow \{\lambda'_{m+1} \ldots, \lambda'_n\}$ is one-to-one, it is equivalent to finding the appropriate non-negative values of $\{\lambda'_{m+1}, \ldots, \lambda'_n\}$. Define $J_i(\lambda'_i, x_i)$, the cost for frame $i$, as:

$$J_i(\lambda'_i, x_i) = D_{x_i}(i) + \lambda'_i \cdot R_{x_i}(i), \forall i \in \{m+1, \ldots, n\}. \tag{15}$$

If each video frame is encoded as intra-frame, the quantizer for each frame can be independently chosen by minimizing the cost for each frame $J_i(\lambda'_i, x_i)$ as:

$$x_i^* = \arg\min_{x_i \in \mathcal{Q}} J_i(\lambda'_i, x_i), \ \forall i \in \{m+1, \ldots, n\}. \tag{16}$$

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 [7] a similar problem is solved by iteratively increasing the lower bounds on the multipliers, defined as $\{\Lambda'_{m+1}, \ldots, \Lambda'_m\}$, such that the violation of rate constraints can be prevented, and adjusting the values of $\{\lambda'_{m+1}, \ldots, \lambda'_m\}$ until an optimal bit allocation, where none of the constraints is violated, is found. We use the algorithm of [7] to solve problem as formulated in (16). The algorithm is described as follows:

**Step 0** Initially the quantizer choices $\hat{\mathbf{x}} = \{\hat{x}_{m+1}, \ldots, \hat{x}_n\}$ are obtained by using a single Lagrange multiplier $\lambda'_n$ for all blocks in (16), subject to only one constraint: $\sum_{j=m+1}^{n} R(j) \leq \left[\sum_{k=t+1}^{(n+\Delta N)\times F} C(k)\right] - R'(m)$.

**Step 1** If $\hat{\mathbf{x}}$ is such that all rate constraints in (16) are met, then $\hat{\mathbf{x}}$ is the optimal solution $\mathbf{x}^*$ for Formulation 2. 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 $\Lambda'_v = \min \lambda'_v$ for the video segment from frame $m+1$ to frame $v$ which just prevents violation of the rate constraint:
$\sum_{j=m+1}^{v} R(j) \leq \left[\sum_{k=t+1}^{(v+\Delta N)\times F} C(k)\right] - R'(m)$.

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

**Step 3** Go to Step 1. Repeat until all the rate constraints in (10) are met.

Refer [7] for detailed description of the algorithm and the proof of optimality.

## 5. RESULTS AND CONCLUSIONS

We simulate the burst-error channel by using the Markov model defined earlier, and apply the Lagrangian-based rate control at the encoder[1] In order to allow a better response to channel state changes, we define sets of 3 Macro Blocks (MB) as our basic coding unit (i.e. each group of three MBs has the same quantizer). We also select the frame rate such that each packet interval equals to duration of 3 MBs.

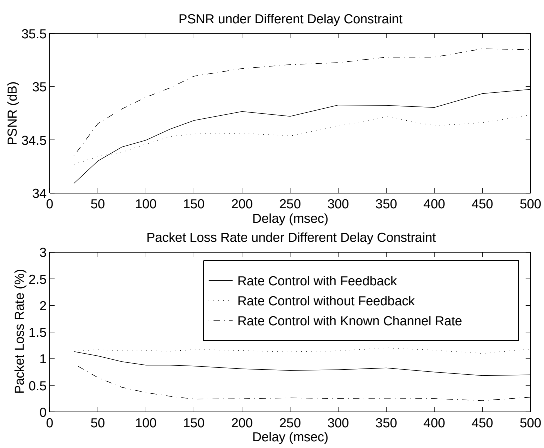



Figure 4: PSNR and Packet Loss Rate in the Downlink transmission under delay constraints from 25 to 500 msec.

For comparison purposes, we also simulate the same communication system and rate control without the real-time channel feedback, i.e., we estimate all future channel rates to be equal to the nominal channel rate weighted by the probability of error. We also apply our rate control algorithm in the unrealistic scenario where the encoder has deterministic knowledge of future channel conditions. This serves as a bound on the achievable performance. We use the channel model to generate 7 different channel realizations, and apply the rate control on the same transmitting video sequence. The resulting video distortions and packet

[1]The test video sequence is "Susie" from frame 1 to frame 100 in QCIF format (176 × 144 pixels for each frame). Each frame is encoded as intra-frame by H.261 encoder at the quantization step size chosen from four values: 12, 14, 20 and 30.

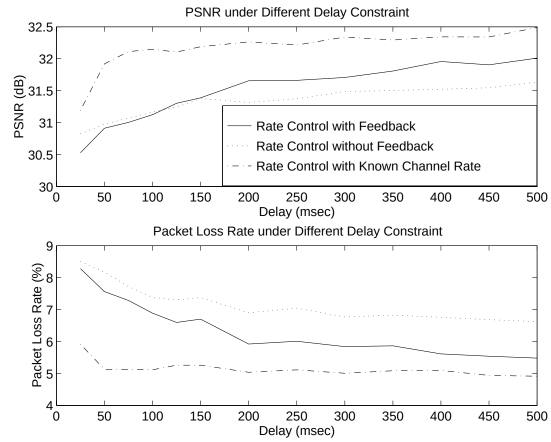



Figure 5: PSNR and Packet Loss Rate in the Uplink transmission under delay constraints from 25 to 500 msec.

loss rates are averaged over all these realizations. Fig 4 shows the resulting video distortion (measured in PSNR) and the packet loss (due to excessive delay) in the downlink channel under various delay constraints. Fig 5 shows the result for the uplink channel.

Our results indicate that using feedback results in lower packet losses and higher reconstructed PSNR. Also, the Lagrangian-based rate control introduced here was shown in our simulations to be up to two orders of magnitude faster than the dynamic programming approach proposed in our previous work [2].

## 6. REFERENCES

[1] 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*, Oct. 1995.

[2] 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*, 1997.

[3] 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.

[4] E. N. Gilbert, "Capacity of burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sept. 1960.

[5] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sept. 1963.

[6] A. Ortega and M. Khansari, "Rate control for video coding over variable bit rate channelswith applications to wireless transmission," in *Proc. the 2nd. Intl. Conf. on Image Proc., ICIP'95*, 1995.

[7] A. Ortega, "Optimal rate allocation under multiple rate constraints," in *Data Compression Conference'96*, 1996.

[8] 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.