RATE CONTROL FOR VIDEO TRANSMISSION OVER MULTIPLE CHANNELS IN ATM NETWORKS

W. Xu, Y. Sermadevi, and S. S. Hemami

Cornell University, School of Electrical and Computer Engineering, Ithaca, NY 14853
email: {wxu, swamy, hemami}@ece.cornell.edu

ABSTRACT

Variable bit rate (VBR) transmission of video over asynchronous transfer mode (ATM) networks requires rate control to select both source and channel rates without violating buffer and network traffic constraints. When the peak transmission rate is limited on a single channel for delay-sensitive applications, the source can be partitioned over multiple channels with possibly unequal conditions. In this paper, it is shown that the rate-control algorithms for multiple channel transmissions can be directly extended from their single channel counterparts, with appropriate modifications. A rate-control algorithm is implemented that selects source and multiple channel rates for each video frame in order to minimize the distortion variation, subject to en-/decoder buffer and channel traffic constraints.

1. INTRODUCTION

Many state-of-art compression algorithms produce variable bit rate (VBR) video streams in order to achieve best compression efficiency and visual quality for complex video contents. Due to limited resources, video sequences transmitted over networks must comply with network policing constraints along with physical buffer sizes at the encoder and decoder. One of the most commonly used policing mechanisms for asynchronous transfer mode (ATM) networks, for example, is leaky bucket (LB) [1]. Therefore, transmission of video streams over ATM networks requires rate control to select both source and channel rates to achieve the best video quality at the decoder, without violating any of the constraints imposed by the networks and en/decoders.

The problem of joint selection of source and channel rates for VBR video transmission has been considered in [2][3], where the source and channel rates are selected either heuristically [2], or through a search in finite sets [3]. The optimal video quality (defined as a function of frame distortions) is obtained within the space expanded by the valid source and channel rates.

With the dramatic increase of demand for video applications over networks, the available bandwidth for each application stream can be very limited. The peak transmission rate provided by a network connection can be close to or even below the video average rate at the desired quality and delay requirements. This implies that the valid space of source and channel rates can be so small that the optimal solution cannot provide good video quality. When multiple network connections with small bandwidths are available, a video stream can be partitioned over these channels to obtain a logical channel with a higher bandwidth [4].

In this paper, the joint rate control problem is investigated assuming multiple channels with different traffic constraints are available. In [5], the problem is formulated assuming large enough physical buffers are available. Thus the constraints were only imposed by networks. Expensive computation complexity and dramatically increased dimensions have prohibited the problem from being further generalized. Here, it is shown that the formulation in single-channel cases can be directly extended to multiple-channel cases, with modifications on buffer state updates and LB constraints. Due to an increased dimension in the search space of solutions, fast implementations of the algorithms are desired. Linear programming techniques are applied to solve for the optimal source rate and multiple channel rate allocation.

The paper is organized as follows. Section 2 demonstrates the extension of buffer constraints from single-channel to multiple-channel cases. Section 3 presents the general formulation of joint source and multiple channel rate selection. Section 4 briefly introduces the implementation techniques based on linear programming. Section 5 shows the experimental results, and section 6 concludes the paper.

2. VBR BUFFER CONSTRAINTS

In this section the constraints imposed by both the networks and en-/decoder buffers are briefly introduced. Extensions of these constraints are made to multiple channels.

Assume there are N channels available with their traffic policing constraints. The channel conditions may vary over time. For real-time applications, it is assumed that the channel parameters at intervals corresponding to independent coding units, such as GOPs in MPEG coding, are known to the rate-control algorithm.

In contrast to single channel applications, N channels have different propagation delays T^j_{prop} ($j = 1, \dots N$). Thus the encoder and decoder buffers must store $\triangle M$ frames at any time, where

$$\triangle M = \max_j (\mathcal{T} - T^j_{prop})/T_{fr} \quad (1)$$

$$= (\mathcal{T} - \min_j T^j_{prop})/T_{fr}, \quad (2)$$

$\mathcal{T}$ is the constant end-to-end delay in order to maintain non-interrupted decoding at the decoder, T_{fr} is the frame interval, $j = 1, \dots N$. The maximum provides a conservative value of the number of frames the buffers must be able to store. Because the channels are transmitting portions of a single video sequence in parallel, fewer frames may be stored in either encoder or decoder buffers in practice.

For the purpose of illustration, each channel is policed by a leaky bucket (LB) ($\bar{C}^j$, L^j_{max}) ($j = 1, \dots N$), where $\bar{C}^j$ and L^j_{max} are the average drain rate and the LB size of channel j respectively. Other policing functions can also be applied such as constant rates, or sliding-window mechanisms. At the i-th frame interval, the encoder outputs E_i bits to the encoder buffer and channel rate C^j_i for channel j is selected to transmit bits from the encoder buffer to the decoder buffer.

Constraints for the buffers and LBs have been derived in [2][3]. The encoder and decoder buffer states are denoted as B^e_i and B^d_i at the end of frame interval i, with the initial buffer fullness as B^e_0 and B^d_0, respectively. LB state is denoted as L^j_i for channel j at interval i. Their time updates at interval i are as follows

$$B^e_i = B^e_{i-1} + E_i - \sum_{j=1}^{N} C^j_i \quad (3)$$

$$B^d_i = B^d_{i-1} + \sum_{j=1}^{N} C^j_i - E_{i-\triangle M} \quad (4)$$

$$L^j_i = \max\{0, L^j_{i-1} + C^j_i - \bar{C}^j\} \quad (5)$$

where $i = 1, \dots n$, n is the number of frames in the coding unit. Note that instead of using one channel rate, the sum of all the channel rates is used in (3) and (4) because multiple channels are transmitting the source bits in parallel. The constraints imposed by the buffers are therefore

$$\begin{aligned} 0 \le B^e_i &\le B^e_{max} \\ 0 \le B^d_i &\le B^d_{max} \\ L^j_i \le L_{max} \qquad j &= 1, \dots N \end{aligned} \quad (6)$$

3. PROBLEM FORMULATION

As mentioned above, the channel parameters are assumed known for the current coding unit (GOP, for example) with n frames to encode. Let d_i be the distortion of the i-th frame ($i = 0, \dots n-1$). Assume that $R_i(d_i)$, the rate-distortion curve for the i-th frame, is: piecewise linear, monotone decreasing, convex and continuous.

In order to achieve video quality with small variations, the video coding problem is formulated in the minmax form as follows:

$$\min_{E_i, C^j_i} \max_i d_i \quad (7)$$

$$\text{subject to} \quad \text{constraints in } (6) \quad (8)$$

$$\textstyle\sum_{i=1}^{N} E_i \le E_{total} \quad (9)$$

$$0 \le C^j_i \le C_{max} \quad (10)$$

where E_{total} is the total number of bits allowed to code the unit, C_{max} is the maximum channel rate. The rate-distortion function follows the curves $R_i(d_i)$. Compared to that for single-channel applications, the formulation is to optimize over multiple channel rate selections C^j_i at each interval. Additionally, in order to allow for channel variations for the following coding unit, more constraints can be added which reset the decoder buffer and LB states to certain values after the last frame interval, such as their initial states or half the buffer fullness.

4. IMPLEMENTATION

The above derivation shows that the rate-control algorithms for multiple channel transmissions can be directly extended from their single channel counterparts, taking into account the changes in the updates of buffer states with multiple LB constraints.

The rate control algorithm in [6] is implemented to jointly select source and multiple channel rates in order to minimize the distortion variation in the coded sequence. The algorithm employs an iterative method to solve the rate control problem for a predictive coder. At each step in the iteration, frame-level distortion-rate curves are gathered assuming fixed reference frame allocations. The resulting curves are assumed to be independent. A linear relaxation of the rate constraints and a piecewise linear approximation of the frame-level distortion-rate functions are used to formulate and solve the problem using Linear Programming (LP) techniques.

5. EXPERIMENTAL RESULTS

A number of experiments have been performed to evaluate the performance of the rate-control algorithm given multiple channels. The standard SIF sequences *Football* and *Flower Garden* were used in the experiments. The decoder buffer is of size 96 kbits. The decoding starts at the end of the 3rd interval. The leaky buckets are assumed half full initially when the allocation starts. The maximum channel rate

C_{max} is 1 Mbits/sec. The sequences are encoded in MPEG with a GOP size of 15 at 30 frames/sec. The average rate of the sequence is set to 0.8 Mbits/sec.

5.1. Constant Channel Conditions

In this experiment 3 channels are available. The channel LB descriptors are held constant as (80 kbits/sec, 32 kbits), (240 kbits/sec, 96 kbits), (400 kbits/sec, 160 kbits), respectively. Figure 1 (a) gives the rates and the PSNR of the encoded *Football* frames over the optimization iterations. The variation range of PSNR is decreased from 6.7 dB to 0.9 dB. Figure 1 (b) shows the allocated channel rates corresponding to each frame interval. When the channels have constant descriptors, their rate allocations follow similar trends depending on the encoded frame types. The curves vary around their corresponding drain rates. Experiments with *Flower Garden* display similar results.

5.2. Varying Channel Conditions

Two channels are assumed available in this experiment. Their LB descriptors remain exactly the same for the first 14 frame intervals as (360 kbits/sec, 144 kbits). Starting from the 15th interval, the drain rate of the first channel decreases to 36 kbits/sec while that of the second channel increases to 720 kbits/sec. Figure 2 shows the selection of the encoded frame rates and channel rates for the sequence *Football*. The sequence PSNR displays similar performance to that when the channel conditions are constant. Its variation decreases from 6.7 dB to 1.0 dB. Because the rate control algorithm is aware of the future channel variations, the rates allocated to the two channels differ starting from the first interval. This difference increases after the 6th interval and becomes dramatically obvious after the 14th interval. Experiments with *Flower Garden* show similar results.

5.3. Comparison with One Channel Transmission

In this experiment a comparison is performed on the video quality between the transmission using a single channel and that using two parallel channels for both sequences. The single channel (labelled S) has LB descriptor (720 kbits-/sec, 288 kbits); while each of the two channels (labelled P and Q) has LB descriptor (360 kbits/sec, 144 kbits). For a fair comparison, channel S has a maximum channel rate C_{max} twice of either of channel P and Q, at 2 Mbits/sec.

The experiment shows that the encoded quality (in terms of PSNR) is exactly the same whether using one channel to transmit the sequence or using two parallel channels. The PSNR of the encoded *Flower Garden* frames is between 25.4 dB to 27.0 dB. The PSNR of *Football* is between 28.6 dB to 29.5 dB. However, the maximum allocated channel rates are 1.46 Mbits/sec and 1.16 Mbits/sec for *Flower Garden* and *Football* respectively over the tested intervals, when the sequence is transmitted over channel S; while the maximum channel rates are 0.73 Mbits/sec for *Flower Garden* and 0.62 Mbits/sec for *Football* in each of channel P and Q. This is intuitively correct because channel S requires higher channel rates in order to achieve comparable performance to that using two parallel channels.

6. CONCLUSION

In this paper the problem of jointly selecting source and channel rates is investigated under the scenario that one video application is transmitted over multiple ATM connections in parallel. It is shown that rate-control algorithms for the problem can be extended from those for single channel applications, with an increased dimension in the constraints. An example rate-control algorithm using LP techniques is implemented and demonstrates that the performance in terms of video PSNR using multiple channels is exactly the same as that using a single channel when the combined effect of LB parameters are equivalent. Experiments with various video sequences also show that the algorithm can successfully take into account the variations in channel conditions and adjust the rate allocation between multiple channels accordingly. The algorithm can be further improved when the channels have loss behaviors and channel coding is applied. A straightforward solution is to consider the coding redundancy as a proportional factor and decrease the channel LB descriptors accordingly.

7. REFERENCES

[1] A. R. Reibman and A. W. Berger, "Traffic descriptors for VBR video teleconferencing," *IEEE/ACM Trans. Networking*, vol. 3, pp. 343–50, April 1991.

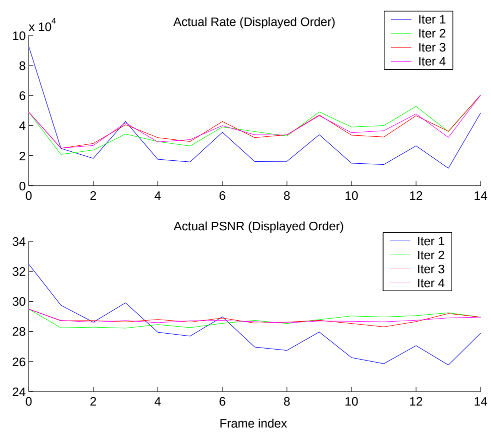
[2] A. R. Reibman and B. G. Haskell, "Constraints on variable bit rate video for ATM networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 2, no. 4, pp. 361–72, Dec. 1992.

[3] C.Y. Hsu, A. Ortega, and A.R. Reibman, "Joint Selection of Source and Channel Rate for VBR Video Transmission Under ATM Policing Constraints," *IEEE JSAC*, vol. 15, no. 6, pp. 1016–28, Aug. 1997.

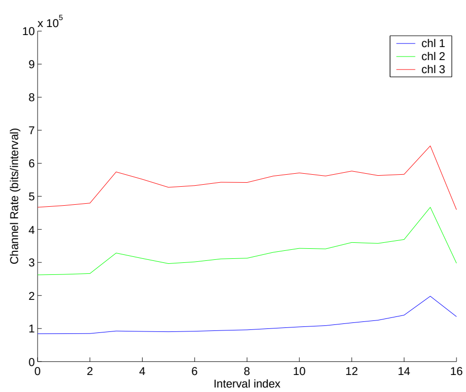
[4] W. Xu and S. Hemami, "Efficient Partitioning of Unequal Error Protected MPEG Video Streams for Multiple Channel Transmission," in *Proc. ICIP*, Rochester, NY, Sept. 2002, vol. 2, pp. 721–4.

[5] S. Servetto, K. Ramchandran, K. Nahrstedt, and A. Ortega, "Optimal Segmentation of a VBR Source for its Parallel Transmission over Multiple ATM Connections," in *Proc. ICIP*, Santa Barbara, CA, Oct. 1997, pp. 5–8.

[6] Y. Sermadevi and S.S. Hemami, "Linear Approximations for Rate Control in Video Coding," in *Proc. ICIP*, Barcelona, Spain, Sep. 2003.

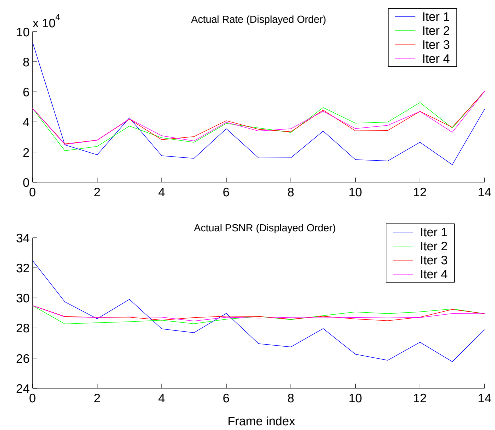


(a) Encoded frame rate and PSNR of *Football*.

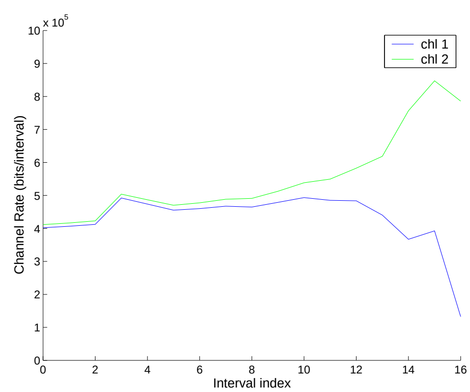


(b) Allocated channel rates for each frame interval.

Fig. 1. The optimal frame and channel rate allocation and the corresponding frame PSNR for the sequence *Football* using LP techniques. The sequence is transmitted through 3 parallel channels with LB descriptors (80 kbits/sec, 32 kbits), (240 kbits/sec, 96 kbits), (400 kbits/sec, 160 kbits).



(a) Encoded frame rate and PSNR of *Football*.



(b) Allocated channel rates for each frame interval.

Fig. 2. The optimal frame and channel rate allocation and the corresponding frame PSNR for the sequence *Football* using LP techniques. The 2 channels have LB descriptors (360 kbits/sec, 144 kbits) from interval 0 to 13, with a change to (36 kbits/sec, 144 kbits) and (720 kbits/sec, 144 kbits) respectively from interval 14 to 16.