JOINT SOURCE CHANNEL CODING WITH HYBRID FEC/ARQ FOR BUFFER CONSTRAINED VIDEO TRANSMISSION

Rohit Puri, Kannan Ramchandran
Dept. of Electrical and Computer Engg.
University of Illinois, Urbana-Champaign
Urbana, Illinois

Antonio Ortega
Integrated Media Systems Center
University of Southern California
Los Angeles, California

Abstract - We propose an Automatic Repeat Request (ARQ) / Forward Error Correction (FEC) scheme for synchronous transmission of video over a binary symmetric constant rate channel. The approach consists of jointly allocating source and channel rates to video blocks from a given admissible set subject to the buffer or equivalently end-end delay constraints. The channel codes used are the popular class of powerful FEC codes known as Rate-Compatible Punctured Convolutional (RCPC) Codes. The method used involves independent coding of the video units and optimization of the end-to-end expected delivered video quality. The existence of a return channel is assumed through which the decoder informs the encoder about the success/failure of the transmission. In the event of a failure, incremental parity information is sent to the decoder for correcting errors and a reallocation performed at the encoder. The simulations done point out the efficacy of the proposed scheme.

INTRODUCTION

Variable bit rate methods for source coding of video have received significant attention in the past in view of their compression efficiency and network bandwidth utilization. The problem of optimal allocation of rates for delay constrained video has been solved for both the cases of noise free [1] and noisy channels [2]. In [2], the channel considered was a burst error wireless channel and the error control technique employed was conventional Automatic Repeat Request (ARQ) wherein based on feedback from the decoder, packets received in error are retransmitted.

In this paper, we employ a hybrid ARQ/FEC scheme using Rate-Compatible Punctured Convolutional (RCPC) codes [3] and investigate the potential advantages of integrating it into an ARQ based scheme to improve the delivered video quality at the receiver. RCPC codes are a family of efficient channel codes with rate compatibility restriction which implies that all the code bits of a high rate code are used by the lower rate codes. Such schemes have been considered for image transmission [4]. However there has been no treatment in literature on the use of these codes in the context of delay constraints (which result in equivalent buffer constraints). Intuitively, a hybrid scheme based on RCPC codes has an edge over conventional ARQ in the sense that during retransmissions following an error, there is no need to resend the entire packet but only incremental parity information as proposed in [3]. Because of the rate compatibility of the channel codes, any parity sent previously (as a forward error correction component or as a retransmission), can be combined with the incremental parity to facilitate error removal at the cost of lesser bandwidth.

SYSTEM, PROTOCOL AND CHANNEL DESCRIPTION

The proposed video transmission scheme has applications in video on demand and other storage and video server applications. The system consists of an encoder (figure (1)) which also performs the role of a rate controller. The rate distortion characteristics of each of the video units are assumed available to the encoder. Based on this knowledge, the encoder allocates source rates and then channel rates to the video units. The transmission units are variable length packets. A typical packet consists of a source-channel code for a video unit and an outer error detection code such as a 16-bit CRC for the source code and finally a header.

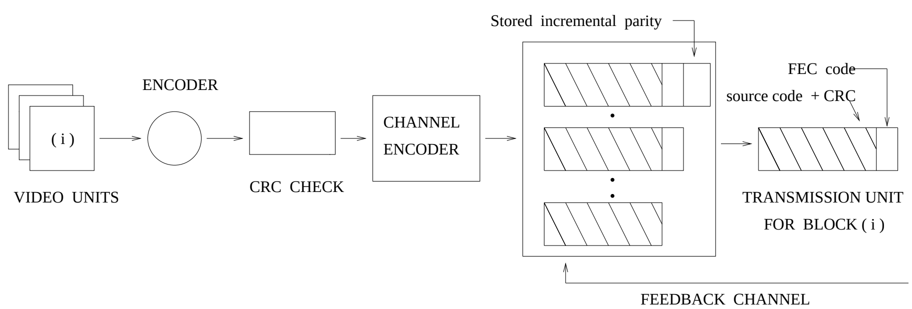

Figure 1: Encoder: Rate Control using hybrid FEC / ARQ

On receiving the bits, the decoder attempts to correct the channel errors and recover the packet. The success or failure in correcting the channel errors is determined by the error detection mechanism (assumed to be perfect) and the result is conveyed back to the encoder through the feedback channel. Upon success the encoder proceeds with the transmission of the next unit. Else, the encoder, optimally switches to a stronger channel code and transmits only the incremental bits needed for the chosen code. The decision to send more bits is taken only if the deadline for the display of the unit has not been reached and the procedure goes on.

The channel model used is that of a constant bit rate, stationary, binary symmetric channel with instantaneous feedback. The rate allocation algorithm developed is particular to this simple model. However, it still highlights the salient features of the buffer constrained joint source channel optimization and can serve as a first step towards analysis of the problem using more complicated channel models. The framework can however, be easily extended to the case of feedback with delay.

PROBLEM FORMULATION

Definitions and Notations

Video Sequence: $(V_n)_{n=1,\ldots N}$ is the sequence of transmission units to send. The units may be video frames, or subbands of a frame, or groups of frames; the only assumption being that each unit being sent is coded independently of others.

Source/Channel Rates: $S_k(i)$ and $C_j(i)$ denote respectively the source and channel rates in bits for the i^{th} unit of quality k and strength j respectively where $k = 1, \ldots.K$ and $j = 1, \ldots.J$. From the Rate-Distortion characteristic of the i^{th} unit, $S_k(i)$ is associated with mean squared Distortion $D_k(i)$. For convolutional channel codes, code strength j is associated with a rate factor $r(j)$ (so that $C_j(i) = r(j).S_k(i)$) and with a probability of bit error $p_b(j)$. The channel codes used are convolutional codes so that for $j_1 \geq j_2$, $r(j_1) \geq r(j_2)$ and $p_b(j_1) \leq p_b(j_2)$. The probability of bit error for a particular code strength depends only upon the signal to noise ratio in the channel. For our scheme we assume that the channel is stationary and the signal to noise ratio and hence $p_b(j)$ for each code strength j is known to the encoder.

Expected Distortion per Transmission Unit: Define probability of frame error $p_f(i)$ for the i^{th} unit as:

$$p_f^{k,j}(i) = 1 - (1 - p_b(j))^{S_k(i)}; \tag{1}$$

The Expected (Mean) Distortion for the i^{th} unit is defined as:

$$ED^{k,j}(i) = (1 - p_f^{k,j}(i)).D_k(i) + p_f^{k,j}(i).E(i) \tag{2}$$

where $E(i)$ denotes the energy of the i^{th} unit i.e. the mean square distortion encountered when zero rate is assigned to the i^{th} unit. This is clearly a pessimistic view since here a unit received in error is considered as lost fully.

Rate Constraints: Time is measured in terms of video unit times. Let N be the total number of units to be sent. Let

$$R^{k,j}(i) = S_k(i) + C_j(i) \tag{3}$$

denote the rate assigned to the i^{th} unit. Then if C denotes the channel bandwidth in bits per video unit time and L is acceptable latency of the application then at time n if unit m is currently being transmitted and $R^{'}(m)$ bits of it are present in the encoder buffer we need:

$$R^{'}(m) + \sum_{l=m+1}^{i} R^{k,j}(l) \leq (L+i-n).C \qquad for \;\; i = m+1, ...N \tag{4}$$

Statement of the Problem

Formulation (For the case of instantaneous feedback): Let the current time be n and the unit being transmitted be m. Let $R^{'}(m)$ bits of unit m be in the encoder buffer. Now if the m^{th} unit is lost before time $n+1$, then the optimization problem can be formulated as:

For all $i \in m+1, ...N$ obtain $k(i)$ and $j(i)$ such that $k(i) \in 1, ...K$ and $j(i) \in 1, ...J$ and that

$$\sum_{i=m+1}^{N} ED^{k(i),j(i)}(i) + ED_{new}^{k(m),j_{higher}}(m) \tag{5}$$

is minimized subject to the constraints (4) where $R^{'}(m)$ in (4) is replaced by $R^{'}(m) + I(j_{low}, j_{higher})$ where $j_{higher} \in j_{low}, ...J$. The latter term is the incremental parity term for the m^{th} unit. For the case of delayed feedback, additional terms appear in (4).

Every time an error occurs, the encoder tries to send an incremental channel code for the unit in error so that (5) is minimized. So there is a dynamic reallocation for all the units which are yet to be transmitted. Since all the units up to n (current time) will already have been encoded, and there is a possibility that during the reallocation lesser number of bits are made available for them than the current assignment, we see that implementation is most elegantly supported if all the units are source coded in an embedded form. In this case taking away bits is equivalent to "chopping" off some bits off the source code.

Optimization Framework

The solution to the above problem is found using Dynamic Programming. In the trellis generated the stages represent the time in terms of video unit times and the states are the encoder buffer occupancy states (B_e) (related at successive time instants by: $B_e(i+1) = B_e(i) + R(i+1) - C$) as shown in figure 2. The constraints in (4) are implemented by restricting the number of states in each stage to $L.C$. The cost associated with a branch from node a in stage n to a node b in stage $n+1$ is given by the expected distortion associated with the rate $r(n+1)$ for the $(n+1)th$ unit $(r(n+1) = s(n+1) + c(n+1))$ where $r(n+1) = b + C - a$. Because of independent coding of the video units and the fact that the channel is a binary symmetric channel (memoryless) the cost of a branch is independent of the path it came from and hence we can use dynamic programming techniques. The metric in (5) represents the total cost of a path from the initial to the final stage and is to be minimized over all possible paths.

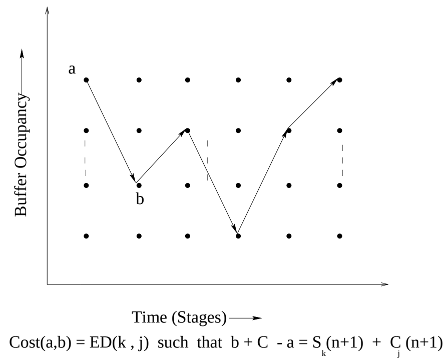

Figure 2: Trellis representing encoder buffer occupancy at different stages.

EXPERIMENTAL RESULTS

For the simulations, Said and Pearlman's (embedded) SPIHT coder [5] was used in the intraframe mode and the family of RCPC codes used (memory M=4 and puncturing period P=8) was obtained from [3]. The rate distortion characteristics used were a parameterized version of those for the "Football sequence" when coded as a sequence of images using the SPIHT coder. The simulations were run for 100 frames with end-end latency L=1. The value of L chosen was motivated by a desire to obtain representative performance for practical sequences (which typically have much larger lengths so that the end-end latency would typically be a small percentage of the playing time).

The video units (obtained after subdividing the output frames from the video player) and their sizes were chosen so that after source coding the video units the average transmission packet size was close to that needed for maximum throughput for the ARQ scheme for each value to signal to noise ratio in the channel. This would offer a conservative comparison. The hybrid scheme consistently outperforms the ARQ scheme for all ranges of bit error rates (figure (3)) with a striking difference for bit errors in the range (10^{-3} to 10^{-5}). A reason for large gains at high BERs is that because of the strict buffer size used (L=1), the ARQ scheme at high BERs is unable to get much throughput across.

CONCLUSIONS AND FUTURE WORK

We formalized the problem of optimal joint source-channel rate control for stored video and presented an algorithm to compute the optimal control which can be trivially extended to real time transmission. We considered

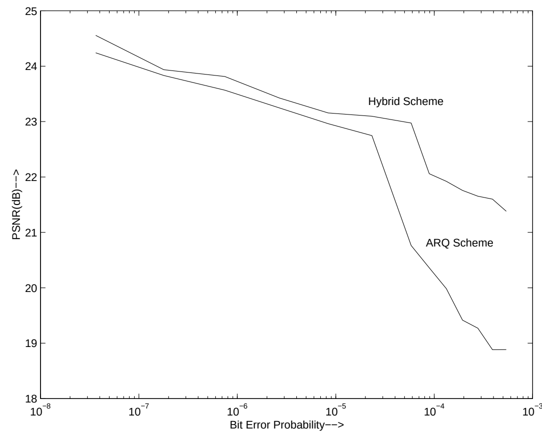

Figure 3: Average PSNR at the decoder as a function of signal to noise ratio in the channel for the hybrid and the ARQ scheme at a target transmission rate of about 1bpp.

the scenario where the encoder has knowledge of the channel state (signal to noise ratio) and the channel is stationary. It is to be noted that the scheme is optimal even for the case of time varying channel when the state is known to the encoder. Thus an extension of this work would be to incorporate a channel state estimator whose output is available to the encoder.

References

[1] A. Ortega, K. Ramchandran and M. Vetterli, "Optimal Trellis Based Buffered Compression and Fast Approximations," *IEEE Trans. Image Processing*, vol. 3, pp. 26–40, Jan. 1994.

[2] C. Y. Hsu, A. Ortega and M. Khansari, "Rate Control for Robust Video Transmission over Burst-Error Wireless Channels," IEEE Journal on Selected Areas in Communications, Special Issue on Multi-Media Network Radios, to appear, 1998.

[3] J. Hagenauer, "Rate Compatible Punctured Convolutional (RCPC) Codes and Their Applications," *IEEE Trans. Communications*, vol. 36, pp. 389–399, Apr. 1988.

[4] V. Chande, H. Jafarkhani and N. Farvardin , "Joint Source-Channel Coding of images for Channels with Feedback," in *I. T. workshop* SanDiego, California, Feb. 1998,pp. 50–51.

[5] A. Said and W. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees"' *IEEE Trans. Circ. Sys. Video Tech.*, vol. 6, pp. 243–250, June 1996.