Error Control for Video Multicast using Hierarchical FEC

Wai-tian Tan, Avideh Zakhor *
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, CA 94720
e-mail: {*dtan avz*}*@eecs.berkeley.edu*

Abstract

Bit-rate scalable video compression with layered multicast has been shown to be an effective method to achieve rate control in heterogeneous networks. In this paper, we further propose the use of hierarchical FEC as an error control mechanism that allows receivers to individually trade-off latency for received video quality. The scheme is efficient since FEC packets are used to protect only the more important data layers and is multicast only to receivers that need them, thereby improving network utilization. Furthermore, there is no loss in error correcting capability by using hierarchical FEC when maximum distance separable codes are used. Actual MBONE experiments are performed to evaluate the performance of the proposed scheme.

1 Introduction

The deployment of IP multicast in the MBONE has provided a convenient and efficient way of disseminating video information to multiple recipients over the Internet. However, since the recipients may have different delay tolerances and are typically connected to the source via paths of different delay, bandwidth and loss characteristics, traditional approaches to flow and error controls based on source adaptations may not be effective. A more elegant alternative is the receiver-driven approach which allows receivers to decide on the communication parameters [1]. For instance, by carrying different layers of scalable video on different multicast groups, receivers can individually subscribe or unsubscribe to the appropriate layers of multicast groups to effect flow control [1, 2].

On the error control side, most work has focused on providing end-to-end reliability based on retransmissions [3] or hybrid schemes employing forward error correction (FEC) and retransmissions [4]. Because the cost of retransmissions in a multicast setting is high, such schemes are often justified only for low bandwidth applications that require reliability, such as the shared white-board in the MBONE tools or financial data distribution. Retransmissions are inappropriate for real-time applications, especially when round trip time is long. In contrast, pure FEC based schemes are simpler, but can be wasteful as redundancy packets are transmitted to all receivers regardless of whether they are needed or not.

In this work, we will investigate the application of hierarchical FEC as an error control method for scalable video multicasting. By hierarchical FEC, we mean embedded FEC or redundancy streams, each of which belongs to a different multicast group. In such a way, subscribing to more groups corresponds to higher level of protection. There are two advantages to using hierarchical FEC. First, each receiver can independently adjust the desired level of protection based on past reception statistics. This receiver-driven approach to FEC is more flexible than most existing FEC schemes for multicast where the source determines a set of redundancy packets to be multicast to every recipient [4]. Second, each receiver will subscribe to only as many redundancy layers as necessary, reducing overall bandwidth utilization. Furthermore, the hierarchical nature of the FEC layers ensures minimum network utilization through sharing of common streams. An effective means to counteract burstiness in packet losses is suggested in [5] for packet audio where coarsely quantized copies of audio segments are staggered in time and transmitted at later times than the original segments to provide resiliency to traffic burstiness. The audio data and its lower quality copies are transmitted using one multicast group, providing little adaptability to receiver requirements. Following [5], we delay the transmission of FEC layers and stagger them in time to alleviate the effects of bursty losses. The additional latency associated with the FEC layers creates a trade-off between higher levels of protection and increased latency, and also acts as a disincentive to unnecessarily use FEC. However, unlike [5], we send each layer of FEC to a different multicast address so that different receivers can customize their subscription levels to their needs.

The use of scalable video has three implications. First, it is possible to achieve flow control by adjusting the number of data layers based on observed channel conditions. Second, it is possible to drop data layers to make room for FEC layers to avoid increase in overall transmission rate. Third, it is possible to strategically provide FEC only to the most important data layer to reduce the amount of FEC traffic. The optimal partition of the available bandwidth between data and FEC depends on the error resiliency of the scalable video compression. Since less error-resilient compression schemes suffer more when packets are lost, they tend to allocate larger portions of the available band-

*This work was supported by AFOSR contract F49620-96-1-0199, Sun Microsystems, Philips, Hughes Research Laboratories, and California State Program MICRO.

0-7803-5467-2/99/ $10.00

width to FEC. This generally causes the quality of the received video to be lower. As a result, error resiliency is desired even when FEC is used for error control. For the experiments in this paper, we use an error resilient, scalable video compression that we have recently developed for best-effort packet networks [6], and propose a new frame-rate scalable extension to it to provide better picture quality at lower bit rates.

In Section 2, we introduce hierarchical FEC as an error control mechanism for video multicast. The selection of multicast rate control algorithm, error control code for hierarchical FEC, and scalable compression for layered multicast are described in Sections 3, 4 and 5 respectively. The receiver algorithm to partition the available bandwidth between the data and FEC layers is described in Section 6. Finally, actual MBONE experiments are described in Section 7 followed by a conclusion.

2 Flow Control using Hierarchical FEC

In a multicast environment, besides bandwidth heterogeneity, different users may experience different packet loss rates and may have different degrees of tolerance to latency. For example, in live multicast of a lecture, participants who want to ask questions and interact with the lecturer desire stringent real-time constraints on the video while passive viewers may be willing to sacrifice latency for higher video quality. Because of the high bandwidth required for video traffic, generating a separate stream for every possible set of user requirements and channel loss rate is infeasible. Instead, we wish to cater to different users with varying latency requirements by providing a framework in which users can individually trade-off latency for quality.

In our proposed scheme, a scalable video source not only sends video packets to multiple multicast groups as in layered multicast, it also constructs hierarchical FEC layers for the data layers. Fig. 1 illustrates our proposed scheme using 3 layers of data and 2 layers of hierarchical FEC:

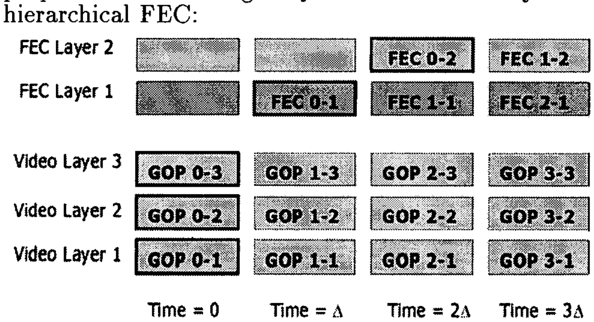

Figure 1: *Timing diagram of hierarchical FEC scheme*

where GOP-x-y denotes the y-th data layer corresponding to the x-th GOP. Similarly, FEC-x-y denotes the y-th FEC layer generated for the x-th GOP, and can be used to protect any single data layer. Note that more than one FEC layer can be used to protect a single data layer.

Addresses D_1, D_2, D_3 are used to carry layered video data and addresses R_1, R_2 are used to carry redundancy packets. All layers of video data corresponding to the same group of pictures (GOP) that are compressed together are transmitted in the same time slot. As a result, users with stringent delay requirements can simply subscribe to D_1, and possibly D_2 and D_3 depending on the bandwidth that is available to them. Users with higher tolerance of latency who are also experiencing packet losses can subscribe to addresses R_1 and/or R_2 to obtain redundancy to repair lost packets at the cost of extra delay. The reason for spreading the redundancy layers for the same GOP across different time slots is to achieve interleaving, which is an effective means of combating bursty losses often experienced in the Internet [5]. Fig. 2 summarizes the behavior for different users. We see that FEC packets are only transmitted to branches that lead to subscribers requiring FEC, thereby reducing overall network load.

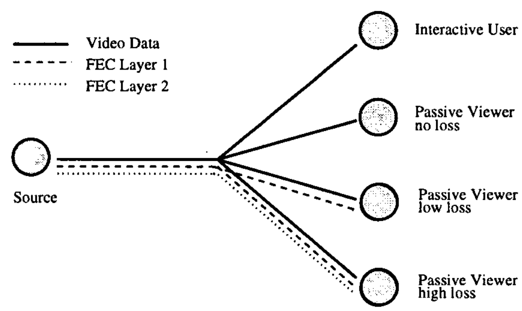

Figure 2: *Hierarchical FEC tailored to different receivers*

3 Rate Control

A central problem in layered multicast is determining the rate control algorithm at the receivers. Receiver-driven layered multicast (RLM) is an early scheme that drops a layer when congestion is detected and "probes" for more bandwidth when losses are absent [1]. Since the act of probing for bandwidth may cause congestion, leading other receivers to mistakenly reduce their subscription levels, the difficult tasks of coordinating probes and communicating their results are necessary [1, 7]. An alternative approach that does not involve probing, but instead calculates the bandwidth directly using measured quantities at the receiver is described in [8]. The scheme is based on a TCP-friendly equation [9] that estimates the average throughput T of a TCP connection under packet loss rate p and round trip time rtt:

$$T = \frac{1.22 \times MTU}{rtt \times \sqrt{p}} \tag{1}$$

where MTU, the maximum transport unit, is a constant that is often taken to be 500 bytes.

Because the problem of rate control in a layered multicast setting is still an area of active research and that our proposed hierarchical FEC scheme does not depend on any specific rate control scheme, we adopt a simple rate control algorithm which is sufficient for the scale of experiments we are performing. Specifically, our rate control scheme is similar to [8] in that every receiver estimates the available bandwidth according to Equation 1 in an uncoordinated fashion. To avoid unnecessary reaction to network transients, the bandwidth calculation is performed relatively infrequently, with a period of 50 seconds. The packet loss rate p is measured over the whole period, and over all data and FEC layers while rtt is updated only once every 5 seconds to reduce the load on the source. Clearly, it is possible to augment our rate control approach by adding synchronization mechanisms as in [7].

Given a dynamic bandwidth constraint obtained by periodically applying Equation 1, the receiver then decides on the optimal number of data versus FEC layers to subscribe. The partition of bandwidth between data and FEC depends on the characteristics of the scalable video compression method used and is described in Section 6.

4 Maximum Distance Separable Code

An effective FEC scheme will allow recovery of lost data packets with as little redundancy packets as possible. This is achieved when the loss of any m data packets are recoverable from any m parity packets. Such properties are often referred to as maximal distance separable (MDS) in algebraic coding theory. An example of MDS code is the Reed-Solomon code. Fig. 3 illustrates the use of a systematic (n, k) MDS code to generate $n-k$ FEC packets, which are partitioned to form FEC layers. Let f_i denote the number of packets in the i–th FEC layer. Since punctured MDS codes are also MDS [10], the first FEC layer can correct up to f_1 packet losses, the maximum any error control code with f_1 redundancy packets can achieve. Similarly, the first l FEC layers achieve maximum error correcting capability by being able to correct up to $\sum_{i=1}^{l} f_i$ lost packets. Hence, there is no loss in error correcting capability by using MDS codes for hierarchical FEC.

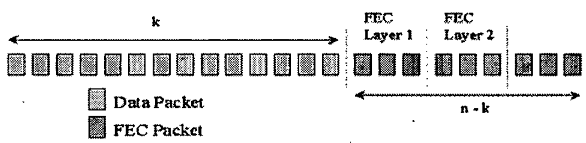

Figure 3: *Generating FEC layers using MDS codes*

5 Scalable Video Compression

For actual experiments over the MBONE, we have extended the error-resilient, scalable video codec in [6] to lower bit rates by making it both frame-rate and SNR scalable. The codec is chosen because it is (a) specifically designed to limit error propagation due to packet losses, (b) finely scalable, and (c) admits real-time software-only encoding and decoding, and (d) has comparable compression performance compared to non-scalable compression schemes such as MPEG-1

The ability to limit the effects of packet loss is important since even with the use of hierarchical FEC, packet losses may still be present due to traffic burstiness or because real-time constraints have precluded the use of FEC layers. Furthermore, a more error-resilient scheme requires less FEC at the same loss rate. Under our proposed video compression scheme, when a packet is lost, errors do not propagate to other packets in the same layer and only propagate to one packet in the upper layer. As is shown in Fig. 4, this is in contrast to most scalable video compression schemes which assume transport prioritization and produce packets that are highly dependent.

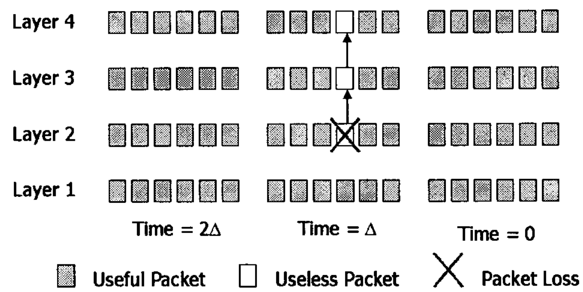

Figure 4: *Data dependency for chosen scalable video*

We now describe the frame rate scalable extension to the error-resilient, scalable compression algorithm based on 3-D subband coding in [6]. The motivation for this extension is to generate video with higher quality pictures at lower bit rates such as 50 *kbps* by allowing the frame rate to decrease. This is in contrast to the approach taken in [6] where the frame rate is constant regardless of bit rate. Even though temporal subband filtering provides multiple temporal resolutions of the same video, the temporally low-pass video is obtained by omitting temporal high-pass subbands. As a result, even at very high bit rates, the pictures reconstructed from temporal low-pass subbands only will be different from the original frames [11]. For example, when the Haar filter is applied to two video frames F_1 and F_2, the low frame rate video reconstructed from the temporal low-pass subbands will only asymptotically approach the averaged frame $(F_1 + F_2)/2$, rather than frames F_1 or F_2, as bit rate increases.

Our proposed approach to achieve, say, two levels of temporal scalability is to first sub-sample the input frames by 2 and then apply regular 3-D subband coding to compress the odd frames up to a predetermined rate R. When bit rate reaches R, we perform 3-D subband filtering on the even frames as well. Instead of coding the odd and even frames as two independent threads, better compression performance can be obtained by predictively coding the low temporal subbands of even frames from the reconstructed low temporal subbands of the odd frames at rate R. Once the even frames have been coded at the same quality as odd frames at rate R, we continue refining the subbands according to the same rate-distortion technique as in [6]. In such a way, videos at full and half frame rates are free from temporal subband filtering distortion. The scheme can be extended to obtain more than two frame rates, say 1/4, 1/8 or 1/16-th frame rate, at the cost of increased complexity.

Table 1 shows the compression performance of the proposed scheme at different frame rates and bit rates, and using different GOP sizes, for the

Mother and Daughter sequence at 30 *fps* and size 320×224. Generally, compression efficiency increases with GOP size due to higher prediction gains, particularly at lower frame rates.

Frame rate	Bit Rate	GOP 4	GOP 8	GOP 16
30 fps	400 kbps	34.14	35.38	35.51
15 fps	200 kbps	32.15	34.31	34.79
7.5 fps	100 kbps	29.88	31.42	32.65

Table 1: *PSNR for Mother sequence at different frame rates and bit rates and using different GOP sizes.*

6 Partitioning of bandwidth between Data and FEC

In this section, we consider the problem of dividing the available bandwidth between data and FEC layers at the receivers. The related problem of how to optimally generate the FEC layers for each data layer at the sender is not considered since it requires *a priori* knowledge of the reception statistics of the receivers. Generally, a source will generate FEC layers for each data layer. However, since the first data layer is generally much more important than other data layers, we make the simplifying assumption that only the first data layer is protected by FEC to reduce receiver complexity. Each receiver then individually partitions its available bandwidth between data and FEC layers for the first data layer so as to minimize the expected distortion given measured packet loss rate and *a priori* knowledge of the rate-distortion characteristics of the video material.

Assume constant packet sizes. When d layers of data are subscribed by a receiver, the expected distortion is given by:

$$D = \sum_{i=0}^{d} p_i \cdot D_i \tag{2}$$

where p_i is the probability that there is enough data and FEC packets to reconstruct the first i data layers but not data layer $i+1$. D_i is the distortion associated with decoding i data layers. Assuming a total of f FEC layers are subscribed and that each data layer and FEC layer contains N_d and N_f packets per GOP respectively, the bandwidth constraint of B packets per GOP results in:

$$f \cdot N_f + d \cdot N_d \leq B \tag{3}$$

The quantities D_i are given by the rate-distortion characteristics of video source and do not depend on the observed packet loss rate p. The D_i can be explicitly computed at the source and communicated to the receivers at the beginning of the session. Alternatively, one can classify video sources according to their characteristics and then generate representative values for D_i for every class. The probabilities p_i depend on p and the data dependencies between packets in the scalable bit-stream. For data dependencies given by Fig. 4, the probability that a first layer packet is lost and cannot be recovered by FEC is given by:

$$p_0 = p \sum_{j=0}^{N_d-1} \binom{N_d + fN_f - 1}{j} (1-p)^j p^{N_d+fN_f-1-j} \tag{4}$$

The other p_i are given by:

$$p_i = \begin{cases} p(1-p_0)(1-p)^{i-1} & \text{for } i \neq d \\ 1 - \sum_{w=0}^{d-1} p_w & \text{for } i = d \end{cases} \tag{5}$$

The optimal number of data layers d^* is then computed as:

$$d^* = \underset{d \in [0, \lfloor B/N_d \rfloor]}{\min \arg} \left(\sum_{i=0}^{d} p_i D_i \right) \tag{6}$$

7 Experimental Results

In this section, we will describe actual MBONE experiments using the proposed hierarchical FEC scheme. Multicast experiments are performed using UC Berkeley as the source and two machines at Georgia Tech and ISI in Los Angeles as receivers. The paths from Berkeley to ISI and Georgia Tech traverses 12 and 13 links respectively with average round trip times of 28 and 80 *ms* respectively. There is a total of 7 common links for the two paths. The video source is an action scene from the movie *Raiders of the lost Ark* at 12 fps and size 320 × 224. A total of 4 data layers each of 100 *kbps* and 4 FEC layers each of 50 *kbps* to protect the first data layer are generated. In Fig. 1, every GOP spans 4 video frames or 1/3 second, and consists of 14 packets of size 320 bytes per data layer, and 7 packets of size 320 bytes per FEC layer. Therefore, an additional latency of 4/3 seconds is introduced when a user decides to subscribes to all FEC layers.

We perform two experiments. In the first experiment, both receivers have stringent real-time requirements so that no FEC layers will be used. In the second experiment, which is run immediately after the first one, receivers are assumed to be non-interactive and hence can sacrifice latency and some data layers to yield bandwidth to FEC layers to achieve increased reliability.

In our experiments, we found the link from Berkeley to Georgia Tech to be free from network congestion and hence no packet losses were observed. As a result, the receiver at Georgia Tech subscribed at the maximum number of 4 data layers and no FEC layers in both experiments. The link from Berkeley to ISI however, had an average packet loss rate of over 18% and given a *rtt* of about 28 *ms*, Equation 1 results in an average throughput of 400 *kbps* for ISI. In the first experiment, all of the 400 *kbps* were allocated to data whereas in the second experiment, 200 *kbps* was allocated to data and the other 200 *kbps* to FEC. The partition of bandwidth between FEC and data was determined by the measured packet loss rate in order to minimize expected distortion as described in Section 6. The rate-distortion characteristics of the video is precomputed at the source and released to the receivers before the experiment.

To compare the two experiments in a meaningful way, it is necessary that channel conditions remain more or less constant during the experiments. Fig. 5 shows the packet loss rates experienced at the different data layers in the two experiments for the connection from Berkeley to ISI. It is observed that the actual observed packet loss rates do not change significantly between the two experiments and that the packet loss rate across layers is more or less constant.

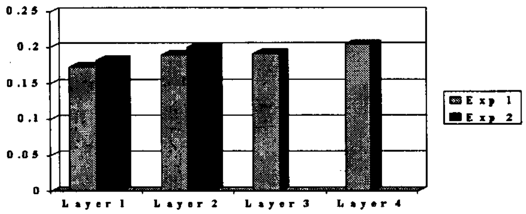

Figure 5: *Observed packet loss rate at different data layers (Berkeley to ISI)*

Fig. 6 shows the the fraction of packets received per GOP for experiment 2 for the Berkeley to ISI connection. We observe that when measured over intervals of 1/3 second, packet losses show much burstiness, which necessitates the delay in sending the FEC layers.

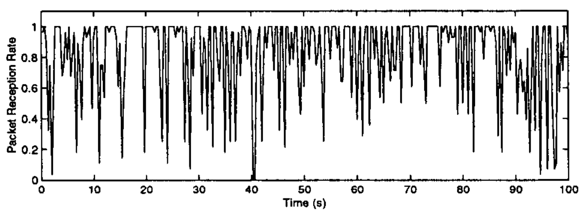

Figure 6: *Overall packet loss rate for experiment 2 (Berkeley to ISI)*

Each experiment is run for a duration of 300 GOP or approximately 100 seconds. For experiment 2, there are 130 GOPs out of 300 in which at least one first layer packet is lost. Of the 130 GOPs, 116 are corrected by the use of FEC. The decoded video trace for experiments 1 and 2 for ISI are shown in Fig. 7. The average MSE for experiments 1 and 2 are 333 and 85 respectively. Even though both experiments consume the same bandwidth, the use of FEC for the first layer enhances the reception quality of experiment 2 at the expense of an extra 4/3 second of latency.

8 Summary and Conclusions

We have proposed the use of hierarchical FEC as an error control method that allows each receiver to individually trade-off latency for better reception quality under layered multicast. The scheme is efficient in that FEC is provided only for the most important layer of the video data, and that there is no loss in error correcting capability by imposing a hierarchical FEC structure. A new frame rate scalable extension to an error-resilient, scalable video compression algorithm in [6] is also introduced to provide better picture quality at low rates. MBONE experiments were performed to demonstrate the potential benefits of scheme.

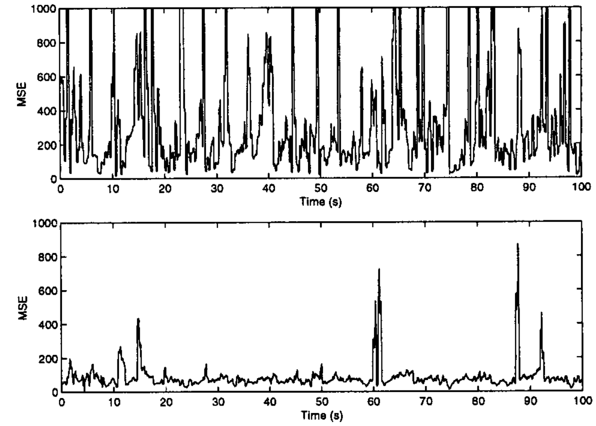

Figure 7: *MSE trace for video in experiments 1 (top) and 2 (bottom) from Berkeley to ISI.*

References

[1] S.McCanne, V.Jacobson and M.Vetterli, "Receiver-driven Layered Multicast," *Proc. Sigcomm*, pp 117-130, Oct. 1996

[2] D.Taubman and A.Zakhor, "A Common Framework for Rate and Distortion Based Scaling of Highly Scalable Compressed Video," *IEEE Trans. CSVT*, Vol. 6, No. 4, pp 329-354, August, 1996.

[3] S.Floyd, V.Jacobson, C.Liu, S.McCanne and L.Zhang, "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing," *IEEE Trans. Networking*, Vol. 5, No. 6, pp 784-803, Dec. 1997.

[4] J.Nonnenmacher, E.Biersack and D.Towsley, "Parity-based Loss Recovery for Reliable Multicast Transmission," *IEEE/ACM Trans. Networking*, Vol. 6, No. 4, pp 349-361, Aug. 1998.

[5] J.Bolot and A.Vega-Garcia, "The case for FEC based Error Control for Packet Audio in the Internet," *ACM Multimedia Sys.*, 1997.

[6] W.Tan and A.Zakhor, "Real-time Internet Video Using Error Resilient Scalable Compression and TCP-Friendly Transport Protocol," *IEEE Trans. Multimedia*, Vol. 1, No. 2, pp 172-186, June 1999.

[7] L.Vicisano, L.Rizzo and J.Crowcroft, "TCP-like Congestion Control for Layered Multicast Data Transfer," *Proc. InfoComm*, Vol. 3, pp 996-1003, March 1998.

[8] T.Turletti, S.Parisis and J.Bolot, "Experiments with a Layered Transmission Scheme over the Internet," *INRIA Technical Report*, http://www.inria.fr/RRRT/RR-3296.html

[9] M.Mathis, J.Semke, J.Mahdavi and T.Ott, " The Macroscopic Behavior of the TCP Congestive Avoidance Algorithm," *Computer Comm. Review*, Vol. 27, No. 3, July 1997.

[10] S.Wicker, *Error Control Systems for Digital Communication and Storage*, Prentice-Hall, 1995.

[11] G.J.Conklin and S.S.Hemami, "Evaluation of Temporally Scalable Video Coding Techniques," *Proc. ICIP*, Vol. 2, pp 61-64, Oct. 1997.