Using Interleaving to Ameliorate the Effects of Packet Loss in a Video Stream

Mark Claypool and Yali Zhu
Department of Computer Science
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609, USA
{claypool,yaliz}@cs.wpi.edu

Abstract

With the rapid improvement in computer and network technologies, high-bandwidth streaming multimedia applications are now possible on the Internet. However, the Internet does not provide the necessary Quality of Service (QoS) guarantees needed to support high-quality, real-time multimedia transmission, causing Internet multimedia applications to suffer from delay, jitter and loss. Among these, loss, typically caused by network congestion, can degrade the perceptual quality of multimedia streams the most. We propose a video interleaving approach that ameliorates the effects of frame loss by spreading out the bursty effects of loss. The sender first re-sequences data before transmission to help distribute loss, and returns the data to their original order at the receiver. We apply our approach to MPEG and evaluate the benefits of interleaving to perceptual quality with user studies. Our results show that interleaving adds a small amount of delay and bandwidth overhead, while significantly improving the perceptual quality of Internet video.

1 Introduction

Emerging new technologies in embedded systems and network protocols along with the explosive growth of the Internet provide great opportunity for distributed multimedia applications. However, today's Internet does not provide the necessary Quality of Service (QoS) guarantees that are needed to support high-quality, real-time video transmission. Multimedia data transmitted over the Internet often suffers from delay, jitter, and data loss. Data loss, in particular, can be extremely damaging to video since the intra-frame dependencies needed to achieve high-compression rates in video exacerbate the data loss when primary frames are lost. Unlike traditional applications, multimedia applications can tolerate some data loss. A small gap in a video stream may not significantly impair media quality, and may not even be noticeable to users. However, too much data loss can result in unacceptable media quality.

A number of techniques exist to repair packet loss in a media stream [1]. These techniques have proven to be effective for audio stream data loss but may have yet to be applied to video. In particular, we propose a video interleaving approach to reduce the damage to a video stream from packet loss. Interleaving assumes that better perceptual quality can be achieved by spreading out bursty packet losses in a media flow. In other words, several small gaps degrade quality less than a big gap in a multimedia flow. For example, consider the simple text-based case with a frame consisting of several characters of information:

WorcesterPolytechnicInstitute

Assume that during network transmission several characters in the frame get lost:

terPolytechnicInstitute

The first word is then very hard to reconstruct. However, if the original frame is interleaved as:

otlhnuWsocItreynstcrtiteePeci

Then after applying the same loss to the interleaved frame, the frame can be reconstructed as:

WrceserPoytecnicIstitute

and it is much easier to "interpolate" the missing letters. This same idea has been applied to audio streams as a loss recovery technique [2], but, to the best of our knowledge, has yet to be applied to video. In addition, it is known that the human visual system can smooth over imperfections better than the human auditory system, thus a small gap in a video stream maybe less noticeable than a small gap in an audio stream, suggesting that interleaving may be even more effective for repairing video than for repairing audio.

In our research, we design and implement an interleaving strategy for video streams. The sender resequences the video stream before transmitting, so that original adjacent units are separated by a guaranteed distance in the transmitted stream, and the receiver returns them to their original order. The re-sequencing guarantees that, in event of loss of a primary frame, the resulting gaps in the stream will be spread out. Our interleaving technique is applied to raw frames instead of encoded frames order to make it independent of specific compression standards.

To evaluate our approach, we design a methodology for measuring the effects of packet loss on video and the possible benefits from our repair technique from the user's perspective. We start from sample MPEG clips recorded from a variety of television programming, and decompress them to get the raw video frames. We apply our repair approach on these raw frames, simulate loss and re-compress the movie clips, which are then shown to users. Users evaluates each movie clip based on the perceived quality (PQ). We also analyze the system overhead, including bandwidth and delay caused by interleaving and compare interleaving to other methods of video repair.

The rest of this paper is organized as follows: Section 2 proposes our video interleaving approach in detail with analysis on system overhead; Section 3 describes our methodology for testing PQ, and discusses the user study results; Section 4 provides a brief comparison of video interleaving to other repair approaches; and Section 5 draws conclusions and briefly discuss possible future work.

2 Approach

Multimedia streams are compressed before being transmitted over the network. The MPEG (Motion Picture Expert Group) achieves high compression rates by exploiting temporal redundancies of subsequent pictures [3]. MPEG distinguishes 3 main frame types of frames for image encoding: the I-frame, P-frame, and B-frame. I-frame stands for Intra-coded frame and is self-contained to support fast, random access; the compression rate of the I-frame is the lowest. P-frame stands for Predictive-coded frame and the encoding and decoding of the P-frame requires the information of the previous I-frame and/or all previous P-frames; the compression rate of the P-frame is higher than that of the I-frame. B-frame stands for Bi-directionally predictive-coded frames and the encoding and decoding of the B-frame requires the information of the previous and following I- and/or P-frame; the compression of the B-frame is the highest. The dependency relationship between the frames is illustrated in Figure 1 for an encoding pattern (Group of Pictures, or GOP) of IBBPBBPBB.

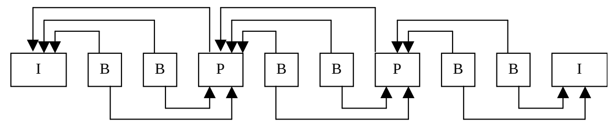

Figure 1: MPEG Frame Dependency Relationship

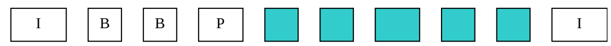

Figure 2: Loss of the Second P-Frame

During transmission of an MPEG video stream over network, a lost I- or P-frame means the dependent P- or B-frames cannot be decoded, leaving a large gap in the video stream playout. For example, if we use the same GOP pattern in Figure 1, the loss of the first P-frame in a GOP will make all the frames after the first I-frame undecodable, and even the loss of the second P-frame will result in a gap of 5 consecutive frames, as shown in Figure 2. This loss propagation appears to the user as bursty packet loss, although only one frame is actually lost by the network.

To repair video in the presence of frame losses, especially consecutive frame losses, we apply a video repair technique called interleaving. The basic idea of interleaving is to uniformly spread out large gaps in the video stream into several smaller gaps. In this way the effect of the loss of multiple consecutive frames is ameliorated, and the perceptual quality will be increased. At the sender, frames in a video stream are first interleaved, with the original consecutive frames being separated by a specific distance that is given by the interleaving algorithm. After arriving at the receiver, frames are then reconstructed back to their original order. If consecutive loss occurs in the interleaved stream during transmission or as a result of single loss propagation, after reconstruction at the receiver, a big gap in the stream caused by the consecutive loss or propagated loss will be spread out into several small gaps.

A parameter to interleaving is the *distance* separating the smaller gaps. For example, with a *distance* of 2, and a GOP pattern of IBBPBBPBB (the GOP size is 9), the interleaving stream will look like:

1 3 5 7 9 11 13 15 17 **2 4 6 8 10 12 14 16 18** ...

where a number indicates the position of one frame in a video stream. With a *distance* of 5, the interleaved stream will look like:

1 6 11 16 21 26 31 36 41 **2 7 12 17 22 27 32 37 42** ...

In the above sequence, each underscored line indicates a group of 9 frames generated by adding the value of *distance* to the number of a previous frame in the stream. This method is repeated until there are 9 frames in the group. After that, a new group starts and the number of the next frame is then picked from the smallest available frame number. The next frame is obtained by adding the value of *distance* to the number of the current frame, until the new group fills up with 9 frames. The number 9 is chosen because it is the number of frames in a GOP and even if all 9 frames in a GOP are lost in the interleaved stream, in the reconstructed stream, those losses are guaranteed to be non-consecutive, separated by the value of *distance*.

Our interleaving approach is also combined with repetition error recovery [1] in which a lost frame is recovered by repeating the previous consecutive frame. For example, with the GOP pattern of IBBPBBPBB, Figure 3 demonstrates whole-interleaving with *distance* of 2 and the result of the first P-frame lost within a video stream, and Figure 4 shows the same P-frame lost in the same video stream without interleaving.

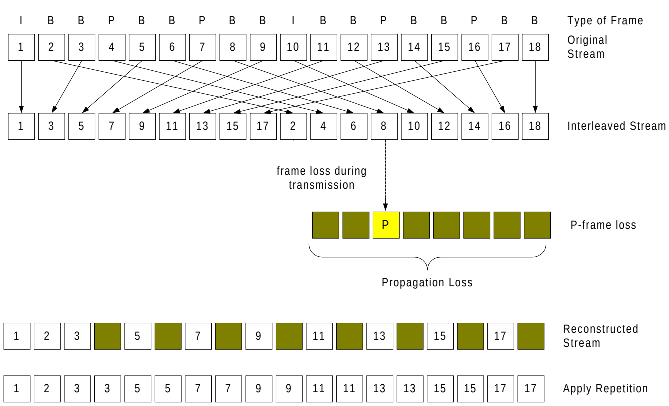

Figure 3: Effects of the First P-frame Lost When Using Interleaving

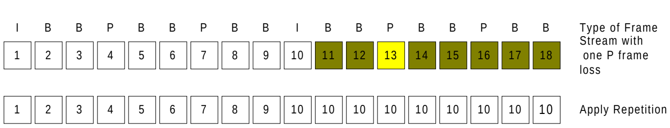

Figure 4: Effects of the First P-frame Lost Without Interleaving

2.1 Overhead

In compressing a sequence of pictures, less motion (smaller differences) between consecutive frames leads to better compression and smaller file sizes, and vice versa. In our video interleaving approach, the larger the interleaving distance the larger the overhead in terms of bandwidth, since originally consecutive frames are further separated by the value of *distance*. Based on measurements using MPEG video clips from Section 3, at an interleaving *distance* of 2, we get an approximate bandwidth overhead of about 15%, and as the distance is increased to 5, the bandwidth overhead increases to about 20%.

To apply interleaving, delay in the video stream is necessary, and the number of delayed frames depends on the interleaving *distance*. Figure 5 shows the process of re-sequencing frames at the sender with interleaving *distance* of 2 and 9 frames in a GOP. From the figure we can see that 2×9 frames need to be delayed to complete the resequence process at the sender. At 30 frames per second, this is approximately 500 milliseconds of delay. Generally speaking, the amount of delay overhead for interleaving is $distance \times size_{GOP}$ frames. The larger the interleaving *distance*, the higher the delay.

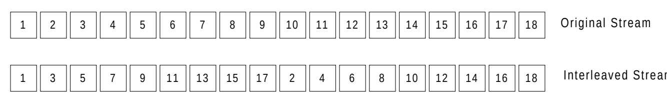

Figure 5: Process of Interleaving at the Sender (top: Original Stream, bottom: Interleaved Stream)

Both sender and receiver require additional processing on the video stream which cause extra delay overhead. However, for re-broadcast video (not live) the processing overhead at the sender can done off-line.

3 Evaluation

Since it is the end-user that determines whether a service or application is a success, it is vital to carry out subjective assessment of the multimedia quality afforded by our video repair approach [4]. In our work, we evaluate the effects of interleaving on the quality of video streams through studies in which users evaluate video quality.

In this section, we first describe the implementation of our interleaving technique, and the process of simulating loss and repair. We then present our methodology to evaluate the effects of video repair by user studies on perceptual quality, followed by the analysis on the results.

For our video interleaving system, video clips used in our implementation were recorded from a variety of television programming and then encoded into MPEG format.[1] The encoding tool we used was the Berkeley MPEG-1 Video Encoder [5], and the decoding tools we used were the Berkeley MPEG-2 player [5] and the Microsoft Media Player. We first broke the original *.mpg* file into separate *.ppm* files using the MPEG decoder, one file for each frame in the video stream. Next, we applied the interleaving algorithm as appropriate to the *.ppm* sequences. Then, we encoded the modified *.ppm* sequences using MPEG encoder to generate the *.mpg* file for transmission over the network. We applied a randomly generated loss rate to the video stream between sender and receiver, simulating a lossy network. After the receiver gets the video stream, the interleaving algorithm was again applied to the video stream to recover from any packets lost.

We chose four loss rates for examination: 2%, 5%, 10% and 20%, which we call the *raw* loss rate. For example, if 10 out of 100 frames are lost through the network, the raw loss rate is 10%. However, as described in Section 2, the loss of an I- or P-frame can results the frames that are dependant on it useless, which results in an even higher frame loss rate shown to the end user.

We did a user study to testing the subjective quality of video interleaving. The study lasted for two weeks during which time 32 users, primarily college undergraduate and graduate students, participated in the study. The study was carried on an Intel PIII (800 MHz) running SuSE Linux 6.4 i686, with the Berkeley MPEG-2 player. We verified that the client machine could play MPEG clips at 30 frames per second. Users were required to watch a sequence of MPEG movie clips and evaluate the quality of the movie clip with scores between 0, the worst quality, and 100, the best quality. After gathering all the scores from the users, we examined the data and analyze the effects of video interleaving on the perceptual quality of video streams.

Figure 6 shows the comparison in quality between interleaving repaired video clips and unrepaired clips. The mean quality scores are plotted along with 95% confidence intervals. At loss rates of 5%, 10% and 20%, interleaving improves the quality of the video clips. At 5% loss, the average score for the unrepaired clips is 59. By using interleaving, the average score improved to 76 (about a 25% improvement). The improvement from interleaving is about 39% at a loss rate of 10% and 38% at a loss rate of 20%. At a loss rate of 2%, although the average score for the interleaved clips is higher than that of the non-interleaved clips, the 95% confidence intervals for the two values overlap. A larger user study with more participants may be necessary to statistically separate the perceptual quality of the two loss rates. Unlike redundancy [6], interleaving still performs well at a loss rate of 20% in the presence of consecutive loss.

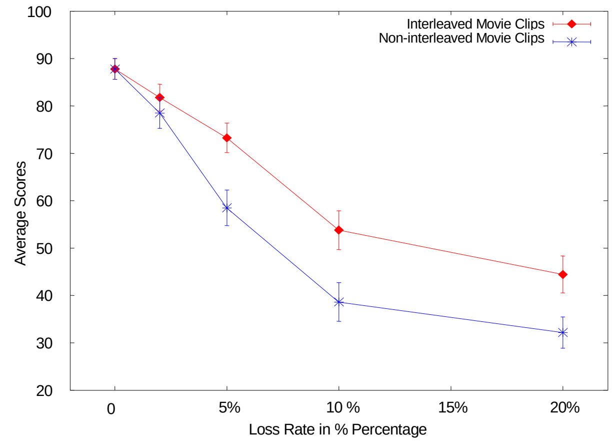

Figure 6: Effects of Loss Rates on Perceptual Quality Using Interleaving

In our user study on interleaving, we also chose two interleaving distances, 2 and 5, to test the effects of the *distance* value on interleaving. At loss rates of 5% and 20%, movie clips are interleaved using a *distance* of 2 and a *distance* of 5. The comparison of those two interleaving distances, together with the non-interleaved clips, is shown in Figure 7. For both loss rates, the unrepaired clips have worse perceptual quality, with interleaving with $distance = 2$ having the highest average score, and interleaving with $distance = 5$ being ranked in the middle. At a loss rate of 5%, the $distance = 5$ interleaving seems only slightly better, since the average scores of unrepaired and $distance = 5$ interleaving are almost the same, with the latter one slightly higher. At a loss rate of 20%, the $distance = 5$ interleaving does have some effect on improving perceptual quality, but the $distance = 2$ interleaving still performs better. These results show that increasing the interleaving distance value may not be helpful to the performance of the interleaving algorithm. Moreover, increasing the interleaving distance dramatically increases latency, as described in the next section.

4 Repair Method Comparison

Our video interleaving approach seeks to repair video by re-ordering the original video sequence. By adding redundant data to a media stream [6, 7, 8, 9, 10, 11, 12], *Forward Error Correction (FEC)* can also be used to repair the damage to the video stream due to packet

[1] All clips can be downloaded from http://perform.wpi.edu/downloads/#movies

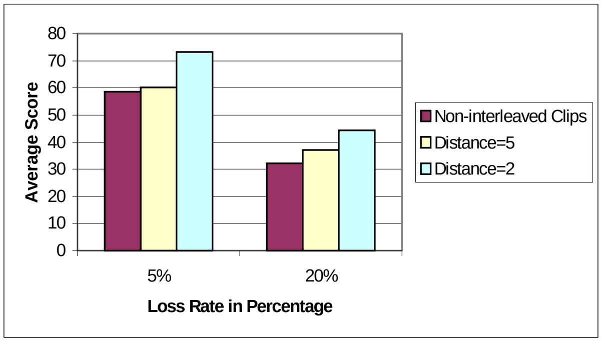

Figure 7: Effects of Interleaving Distance on Perceptual Quality

loss. *Media-dependent* FEC piggy-backs redundant video frames within the transmitted video stream in order to repair lost frames. At the sender, frames are compressed into two quality versions. High quality frames are sent to the receiver as the primary frame, while low quality frames are piggy-backed with the next primary frame as the redundant frames. In the case the primary frame is lost, the corresponding low quality secondary frame is used to replace it. *Media independent* FEC takes K original data packets, produces $(N - K)$ redundant packets, and sends the N packets out. If any K or more packets are received, then all the original packets can be completely reconstructed.

The amount of redundant data added for both media-dependent and media-independent FEC has direct impact on the effectiveness of the repair and the overhead. Research in media-dependent FEC [6] suggests a typical low-quality secondary frame would require about 1/10th as much space as the primary frame. Research in media-independent FEC [13] suggests a typical MPEG GOP would have 4 redundant packets for each 28 original packets in an I-frame and 2 redundant packets for each 8 original packets in a P-frame.

In addition to data overhead, when utilized, all repair approaches also add delay to the playout of the video frames. Interleaving adds a delay of the interleaving distance times the number of frames between I frames (the GOP size). Media-independent FEC adds a delay equivalent to the number of redundant frames added per GOP. Media-dependent FEC adds a delay of only one frame when repairing, but adds no additional delay if there is no loss.

For a quantitative comparison of video repair approaches, we assume full-motion video 30 frames per second, a typical GOP of IBBPBBPBBPBBPBB, a fixed GOP rate with packets spread evenly over the GOP, mean frame sizes based on [14] and use the average quality values reported in this paper and in [13, 6]. Table 1 summarizes the results.

Repair Method	Degradation at 2% Loss	Degradation at 5% Loss	Byte Ovrhd	Delay Ovrhd
Unrepaired	13%	35%	-	-
Intrlvng	8%	17%	15%	540ms
Dep. FEC	8%	20%	10%	33ms
Ind. FEC	3%	18%	18%	130ms

Table 1: Comparison of Video Repair Methods

All three repair technologies provide significant reduction in quality degradation under loss. Media independent FEC is somewhat more effective at low loss rates than either media dependent FEC or interleaving while at moderate loss rates interleaving is slightly better than either type of FEC. Media dependent FEC has the lowest overhead of the three approaches, slightly so for bytes and significantly so in terms of delay. Overall, FEC is more attractive for interactive video applications (such as a video conference) than is interleaving because FEC adds less overhead to the playout delay.

In addition, not shown in the above analysis is the robustness in the presence of bursty loss. Interleaving can withstand burst losses up to the size of the GOP times the interleaving distance since packets are spread out over subsequent GOPs. Media-independent FEC cannot withstand such large loss bursts without incurring significantly more byte and delay overheads. Media dependent FEC cannot withstand packet loss bursts at all without increasing the delay overhead.

5 Conclusions and Future Work

In this paper we present video interleaving as a repair technique to ameliorate the effects of network data loss for video streams. We design a methodology to investigate the benefits to users of our interleaving approach for MPEG and analyze the bandwidth and delay overheads.

Our interleaving approach re-sequence units in a video stream at the sender, transmits them through network to the receiver, and reconstruct the video to its original order. In this way, large gaps in a video stream due to consecutive frame losses can be spread out to several small gaps, thus improving the perceptual quality of the video stream. In the situation of random loss, our interleaving algorithm can improve the perceptual quality by about 25% at a loss rate of 5%, and about 40% at loss rates of 10% and 20%. We also examined the interleaving distance, which is the distance of the small gaps after one large gap is spread out. We find interleaving with a distance of 2 has much better performance than interleaving with a distance of 5.

Interleaving adds an additional delay of the interleaving distance times the number of frames between primary frames. This extra delay may make interleaving more attractive to non-interactive video applications. Interleaving changes the similarity relationship between consecutive frames in a video stream, and results in about a 15% bandwidth overhead. This overhead may be decreased by trading off some quality, as explored in [15].

In this paper, our interleaving methods are applied to a video stream only. A more challenging and useful approach may be to apply interleaving on video and audio streams together. Another possible area for future study is to more completely compare different video loss recovery techniques, such as redundancy, retransmission and interleaving. Since each technique results in different overhead in delay and bandwidth, the effects of this overhead on the network, including network congestion, link utilization and goodput, may be worth further evaluation.

References

[1] Carlos Perkins, Orlin Hodson, and Vicky Hardman, "A Survey of Packet-Loss Recovery Techniques for Streaming Audio," *IEEE Network Magazine*, Sep/Oct 1998.

[2] C. Perkins and O. Hodson, "Options for Repair of Streaming Media," *The Internet Society RFC 2354*, 1998.

[3] J. Mitchell and W. Pennebaker, *MPEG Video: Compression Standard*, Chapman and Hall, 1996, ISBN 0412087715.

[4] A. Watson and M. A. Sasse, "Measuring Perceived Quality of Speech and Video in Multimedia Conferencing Applications," in *Proceedings of ACM Multimedia Conference*, Bristol, UK, Sept. 1998, pp. 55 – 60.

[5] Universiy of California Berkeley, "Berkeley MPEG-2 Encoder and Player," Internent site `http://bmrc.berkeley.edu/frame/research/mpeg/`.

[6] Yanlin Liu and Mark Claypool, "Using Redundancy to Repair Video Damaged by Network Data Loss," in *Proceedings of IS&T/SPIE/ACM Multimedia Computing and Networking (MMCN)*, Jan. 2000.

[7] T. Nguyen and A. Zakhor, "Distributed Video Streaming with Forward Error Correction," in *12th International Packet Video Workshop*, Apr. 2002.

[8] J-C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-Based Error Control for Internet Telephony," in *Proceedings of IEEE INFOCOM*, Mar. 1999.

[9] C. Padhye, K. Christensen, and W. Moreno, "A New Adaptive FEC Loss Control Algorithm for Voice Over IP Applications," in *Proceedings of IEEE International Performance, Computing and Communication Conference*, Feb. 2000.

[10] K. Park and W. Wang, "QoS-Sensitive Transport of Real-Time MPEG Video Using Adaptive Forward Error Correction," in *Proceedings of IEEE Multimedia Systems*, June 1999, pp. 426 – 432.

[11] I. S. Reed and G. Solomon, "Polynomial Codes Over Certain Finite Fields," *Journal of the Society of Industrial and Applied Mathematics (SIAM)*, vol. 8, no. 2, pp. 300–304, June 1960.

[12] Luigi Rizzo, "Effective Erasure Codes for Reliable Computer Communication Protocols," *Computer Communication Review*, Apr. 1997.

[13] Huahui Wu, Mark Claypool, and Robert Kinicki, "Adjusting Forward Error Correction for TCP-Friendly Streaming MPEG," Tech. Rep. WPI-CS-TR-02-29, CS Department, Worcester Polytechnic Institute, Dec. 2002.

[14] M. Krunz, R. Sass, and H. Hughes, "Statistical Characteristics and Multiplexing of MPEG Streams," in *Proceedings of IEEE Infocomm*, Apr. 1995, pp. 455–462.

[15] Yali Zhu, "Using Interleaving to Ameliorate the Effects of Packet Loss in a Video Stream," M.S. thesis, Worcester Polytechnic Institute, June 2001, Advisor: Mark Claypool.