An Adaptive Protection Framework for MPEG Base Layer Video Streaming Over Internet

Chih-Yen Lin¹, Chyohwa Chen¹, Wen-Jen Ho²

{¹Department of Electronic Engineering of National Taiwan University of Science and Technology;
D8902101@mail.ntust.edu.tw, cchen@et.ntust.edu.tw},

{²Multimedia Technologies Laboratory of Institute for Information Industry; wenjen@iii.org.tw}

Abstract

Video streaming over the Internet or wireless environment is an imperative application. However, packetized video frames suffer from random losses when they are streamed over the Internet. Packet losses degrade video quality, even if applications adopt error resilience techniques. To provide more guaranteed video quality, we re-investigate the PET system [7] and propose an adaptive optimization-based protection assignment framework based on PET for streaming MPEG base layer video under a given bandwidth constraint. Our framework incorporates considerations of both the variant bandwidth issue and the packet loss issue on the Internet. The results of our experiments demonstrate our proposed framework achieves the goal of adaptive optimization assignment for MPEG video streaming.

1. Introduction

Video streaming over the Internet is a recent and particularly active research area because of its successful deployment in commercial applications and its central role in any applications involving multimedia content. Currently, popular video contents are coded using the DCT-based method, e.g. H.261, H.263, MPEG1, MPEG2, MPEG4, H.264, etc. Besides DCT transform coding, these video codecs achieve high compression ratio using time domain motion-compensated prediction. The more predictive frames within a group of picture (GOP) are, the higher compression ratio the codecs can achieve. However, these highly compressed video bit-streams are very sensitive to any bit-stream error [12, 13].

Since the Internet is based on IP, which provides an inherently unreliable "best-effort" service, streaming over the Internet suffers from unavoidable bandwidth variation and packet losses. Statistics shows that streaming video over the Internet experiences considerable packet losses, with a common 1~5% packet loss rates, and sometimes up to 30% or more (ICIP 2002, Civanlar's Tutorial). Our experiments indicate it is not feasible to provide MPEG streaming over the Internet if the application does not incorporate any kind of error control. There are two kinds of error control mechanisms: using time to exchange reliability (e.g. ARQ approach), and using space to

exchange reliability (e.g. FEC)[1]. The first mechanism is not suitable for the real-time, delay-constrained video streaming applications. In this paper, we will explore the second mechanism based on notions drawn from unequal loss protection (ULP)[7], specifically PET [7]. PET (Priority Encoding Transmission) is an ULP system that, for each message, uses different FEC protection to different message fragments, according to the priority of the fragments. However, the original PET system did not specify how to determine the level of FEC for each fragment. Leicher [8] applied PET in vic [9] for video transmission over a packet erasure channel. The method is briefly described as follows. The MPEG video stream is divided into coding blocks, or messages, with each coding block being a GOP. Frames are classified according to the MPEG frame types, namely intra-coded (I) frames, predictively coded (P) frames, and bidirectionally predictively coded (B) frames. A manual assignment of FEC rates to each type of frames is then performed according to their priorities. Since Leicher's PET system uses a fixed assignment, it cannot cope with the variable bandwidth issue effectively. A general approach for adaptive FEC assignment is to estimate and use only the bandwidth available and optimizes for packet-loss protection under the bandwidth constraint [3,5,6]. However, previous proposals assume their inputs are progressive bit-streams. Since the MPEG video bit stream is not progressive, their works are not immediately applicable for MPEG. In the paper [4], the authors propose to re-order the MPEG video stream into an embedded form. However the effectiveness of the proposed reordering scheme is time-consuming and the performance is less than a previous simple PET like proposal [2]. In this paper we propose a direct approach to an adaptive optimization assignment framework. Our framework can apply to a ULP system that does not have an optimizing assignment approach, e.g. Leicher's PET system, and HIPP [2].

The remainder of this paper is organized as follows. Section 2 describes the related subjects of ULP framework that includes the definition of ULP system and the definition of utility-cost function for our framework. Section 3 describes our optimization framework and its application in an ULP system. Section 4 evaluates our optimization framework via experiments. Section 5 concludes the paper.

2. The Unequal Loss Protection Framework

Our framework is applied to the PET [7] system. Our framework provides a systematic approach to optimize the assignment of FEC protection for the message types. Since the MPEG video stream is not a progressive bit-stream, it is not possible to assign the FEC rate by applying the works in [3,10]. Since the rate-distortion (RD) relationship is, in general, a complex function of bandwidth, GOP structure, the packetization process, and packet loss pattern, it is hard to derive an accurate RD curve reflecting all of the above factors. Therefore, previously proposed ULP algorithms focus on the bandwidth component only. In trying to design a more accurate framework reflecting all of the factors, we do not estimate the RD curve directly, but make use of a utility-cost function. In this section, we will briefly describe the notion of an ULP system first, and then define the utility-cost function required for our framework.

2.1 The PET ULP system

The PET ULP system is illustrated through the coding block a message, a GOP of frames, in figure 1. Each message is a GOP since a GOP is an independent video subsequence, representing the boundary of error propagation. In the coding block derived from a GOP, N denotes the number of packets used to carry the encoded GOP source data and FEC parity data. L denotes the size of each individual packet. As shown in the figure, a message is partitioned into m message fragments, each of size denoted by r_i . A priority is assigned to each message fragment, and each fragment is protected using different amounts of parity data according to its priority. The system is designed in such a way that for a high priority fragment, only a small number of packets must be received before it can be decoded. In PET, the priority is in the form of fractions assigned to each message segment. If the designated fraction of the N transmitted packets is received, the message fragment can be decoded completely. Thus, high priority message segment is designated with a low fraction. Similarly, in our system, an assignment is a vector $\bar{X}$ of dimension m , $\bar{X} = \langle x_1, x_2, \dots, x_m \rangle, \forall 1 \leq x_i \leq N$. For message fragment i , also called the message fragment at level i , our algorithm calculates the value x_i , which determines the amount of protection associate with the fragment. The amount of protection used for the fragment is $r_i \times N / x_i$. Note that when $x_i = r_i / L$, all parity data is assigned to protect the first fragment only. The general problem is to find an assignment that optimizes expected received quality under a given rate budget R^* , satisfying the constraint $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_m$.

2.2 The Definition of Utility-cost Function

The relationship between the quality and cost for each message is denoted by its utility-cost function, which is defined for each independent GOP. The cost for achieving the quality at a given level i is the cumulative bit-rate required to correctly decode the message segments from the lowest level to the i 'th level. The utility at that level is the quality gained from correctly decoding the message segments up to that level, defined using the average PSNR (PSNR) value, computed as the average of per frame PSNR value over the entire GOP. The PSNR value of each individual frame is computed in a standard manner from the mean squared error (MSE) using formula (1) [11].

Figure 1. An ULP coding block

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}} \quad (1)$$

In order to reflect receiver's decoding actions in the face of errors, the utility-cost function is computed in two steps. The first step is to designate number of levels and the frame loss compensation pattern at each level. As an example, consider a MPEG4 encoder that produces a sequence of encoded data frames with GOP structure "I₁B₁B₂P₁B₃B₄". Three priority levels are distinguished: I type, P type, and B type. Assume the frame size distribution for GOP, is that the size of the single I frame is 1517 bytes, the single P frame is 1312 bytes, and summation of all the B frames is 970 bytes. Referring to table 1, the first column shows that the number of levels is 3, with corresponding frame loss compensation actions performed as shown in the second column. That is, if the decoder can only decode up to level 1, perhaps because it receives only the single I frame, its action is to substitute the single I frame as replacement for all of the remaining frames in the GOP. In the second step, a utility is computed according to the designated frame-loss-compensation pattern for each level, using equation (1) in terms of PSNR. The results are as illustrated in table 1. In general, three factors must be considered: 1) the order of the frame importance (e.g. $I > P > B$), 2) the dependency between the frames (e.g. $I_{i,1} \supset P_{i,1} \supset P_{i,2} \supset \dots \supset P_{i,k}$), and 3) the importance of the data (e.g. motion vector > texture data, lower frequency symbol > high frequency symbol). Finally, we note that the utility-cost function is discrete, and in general not concave.

Level	Frame loss compensation	Cost	PSNR	Label of utility
0	(Blank frame*6)	0	13.1dB	U_0
1 (I)	$\{I_1, I_1, I_1, I_1, I_1, I_1\}$	1517	23.5dB	U_1
2 (P)	$\{I_1, I_1, I_1, P_1, P_1, P_1\}$	2829	27.7dB	U_2
3 (B)	$\{I_1, B_1, B_1, P_1, B_1, B_1\}$	3799	30.1dB	U_3

Table 1. The example of Utility-cost function

3. An Adaptive Optimization Assignment Framework

In this section, we describe our optimization framework, and a simple algorithm for optimization assignment is also provided. The system is designed to operate in real time in response to feedback of conditions of the transmission channel. The architecture for our adaptive optimization framework is shown in Figure 2. The optimization assignment module takes four inputs and outputs an optimization assignment $\bar{X}$. N is the desired number of packets to be used by the ULP system to transmit a coding block. p_j denotes the channel error probability distribution

that j out of N packets are lost, where $0 \leq j \leq N$. R^* is the bandwidth budget for the particular GOP, in consideration. Let the size of GOP, be R . We consider the situation where $R^* > R$ in this paper. The last input is the utility-cost function for the GOP as specified by the user.

Define the expected utility $E[U]$ as follows:

$$E[U] = \sum_{i=0}^N p_i f_u(N-i) \quad (2)$$

$$f_u(k) = \begin{cases} U_0, k < x_1 \\ U_i, x_1 \leq k < x_i \\ U_m, x_m \leq k \end{cases} \text{ where } 0 \leq k \leq N \quad (3)$$

Where $f_u(k)$ in formula (3) returns the utility value when k packets are received by consulting the utility-cost function.

The optimization algorithm must search for the assignment that maximizes $E[U]$. Since the utility-cost function is non-concave, there is no efficient procedure to compute the best assignment. When the number of levels is small, an exhaustive search optimization algorithm is sufficient. The problem is difficult if N and the number of levels are large. Currently, a simple pruning algorithm is used, by pruning off infeasible regions in the search space and then exhaustively search for the best assignment in the feasible regions. At system initialization, $L = R^* / N$ is computed as the packet length. Clearly the case when all of the protection is assigned to I frames only is a possible assignment, i.e. $\bar{X} = \langle x_1 = \frac{r_1}{L}, x_i = 0, i \neq 1 \rangle$. However, we desire an assignment where all frame types are placed into the coding block. That is, the x_i 's must be

$1 \leq x_i \leq N, i = 1 \dots m$. With our desire to provide prioritized protection for more important segments, the x_i 's must also satisfy $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_m$, with the assignment $x_1 = x_2 = x_3 = \dots = x_m$ ruled out since it violates our desire to provide unequal protection. All assignments considered must also satisfy the following equation, which states that the summation of each individual portion of L used for each fragment i must not exceed the packet length L :

$$\sum_{i=1}^m \left\lceil \frac{r_i}{x_i} \right\rceil \leq L \quad (4)$$

Figure 2. The system architecture

4. Experiments

Two experiments are performed to evaluate our proposed framework in this section. In the first experiment, we compare the performance of the original Leicher's PET system (PET-orig) with our optimized Leicher's PET system (PET-opti). In the second experiment, we vary the bandwidth budget in several scenarios to illustrate the effectiveness of our ULP framework in adapting to different bandwidth budget. In our experiments, we use the foreman QCIF with 285 frames, with frame-rate being 30 fps. The sequence consists of 19 GOPs, whose structure is "IBBPBBPBBPBBPBB". Each GOP consists of 15 frames. The average bit rate of the sequence is 100kbps.

4.1 Experiment I

In this experiment, we compare our framework to Leicher's PET system, which employs a fixed overall redundancy rate of around 30% over its message data rate. Using an exponential PMF loss model to simulate the network loss distribution, we explore six different mean packet loss rates scenarios: 0.05, 0.1, 0.15, 0.2, 0.25, and 0.30. The results are as shown in Figure 3. It can be seen our framework outperforms Leicher's PET system for over 1dB in most cases.

Figure 3. The comparison between PET-orig and PET-opt

4.2 Experiment II

In this experiment, we explore scenarios where the redundancy available to our system is 10%, 20%, 30%, 40%, 50% and 60% over the original data rate. The results are as showed in Figure 4. It can be seen our algorithm achieves the desired capability of adaptation to different degrees of packet loss and bandwidth budget. The redundancy of Leicher's PET system is over 30% and our framework still outperforms Leicher's PET system.

Figure 4. The different bandwidth budget for our framework

5. Conclusion

In summary, an adaptive optimization-based FEC assignment framework is proposed. Our framework improves upon the performance of the PET system significantly. We provide a systematic and direct approach towards the optimization of PET-like systems, in contrast to other proposals which assume progressive bit-streams and require re-ordering of MPEG bit-streams. The proposal is shown to be adaptive to bandwidth variability and optimizes the assignment. We also investigated the more cost-level for refining the result of optimization. However, it cannot obtain a significant result. So we believe three cost-level (as Leicher's PET system) is enough for practical applications and the complexity of optimization algorithm will be bounded in an acceptable range. For future work, we intend to investigate more efficient optimization algorithms for larger N values and their graceful degradation properties.

Acknowledgements

The authors would like to thank reviewers for helpful comments.

References

- [1] Dapeng Wu, Yiwei Thomas Hou, WenWu Zhu, Ya-Qin Zhang, Jon M. Peha, "Streaming Video over the Internet: Approaches and Directions", IEEE Transaction on Circuits and Systems for video technology, Vol. 11, No.1 February 2001.
- [2] Jill M. Boyce, "Packet loss resilient transmission of MPEG video over the Internet", Signal Processing: Image Communication, 15(1999) 7-24, September 1999.
- [3] Alexander E. Mohr, Eve A. Riskin, and Richard Ladner, "Unequal Loss Protection: Graceful Degradation over Packet Erasure Channels through Forward Error Correction." IEEE Journal of Selected Areas in Communications Special Issue on Error-Resilient Image and Video Transmission, 18(6): 819-828, June 2000.
- [4] Justin Goshi, Alexander E. Mohr, Richard E. Ladner, Eve A. Riskin, and Alan Lippman, "Unequal Loss Protection for H.263 Compressed Video.", Submitted to IEEE Transactions on Circuits and Systems for Video Technology, June 2002.
- [5] Rohit Puri, Kang-Won Lee, Kannan Ramchandran and Vaduvur Bharghavan. "An Integrated Source Transcoding and Congestion Control Paradigm for Video Streaming in the Internet." IEEE Transactions on Multimedia, Vol.~3, No.~1, pp~18-32, March 2001.
- [6] Rohit Puri, Kang-Won Lee, Kannan Ramchandran and Vaduvur Bharghavan. "Forward Error Correction (FEC) Codes Based Multiple Description Coding for Internet Video Streaming and Multicast." Signal Processing: Image Communication, Vol.~16, No.~8, pp~745-762, May 2001.
- [7] Andres Albanese, Johannes Bloemer, Jeff Edmonds, Michael Luby, Madhu Sudan, "Priority Encoding Transmission", IEEE Transactions on Information Theory, Vol. 42, No. 6, November 1996.
- [8] Christian Leicher, "Hierarchical Encoding of MPEG Sequences Using priority encoding transmission (PET)." Technical Report TR-94-058, The International Computer Science Institute, November 1994.
- [9] <http://www.icsi.berkeley.edu/PET/icsi-pet.html>
- [10] Alexander E. Mohr, Eve A. Riskin, and Richard E. Ladner., "Approximately Optimal Assignment for Unequal Loss Protection.", Proceedings of the 2000

International Conference on Image Processing (ICIP).
October 2000.

- [11] Khalid Sayood, "Introduce to Data Compression
Second Edition", Morgan Kaufmann publishers 2000.
- [12] S. Gringeri, R. Egorov, K. Shuaib, A. Lewis, and B.
Basch. "Robust compression and transmission of
MPEG-4 video", In Proc. ACM multimedia, 1999.
- [13] N. Feamster and H. Balakrishnan, "Packet Loss
Recovery for Streaming Video", 12th International
Packet Video Workshop , Pitsburgh, PA, April 2002.