ICICS-PCM 2003
15-18 December 2003
Singapore

Unequal Packet Loss Protected Transmission for FGS Video

Lianji Cheng; Li Song; Songyu Yu

Institute of Image communication and Information Processing
Shanghai Jiao Tong University

Abstract

Video communication with Quality of Service (QoS) is an important and challenging task. The transmitted video stream must be able to afford the bandwidth variance and unavoidable packet loss in the Internet. In particular, fine-granular-scalability (FGS) video coding has been adopted by the MPEG-4 standard as the core video-compression method for streaming applications. In this paper, we use unequal loss protection (ULP) to protect FGS compressed video, and under the restriction of the network bandwidth, joint source-channel rate-distortion based optimization is performed in bit allocation to minimize the end-to-end distortion. Simulation results demonstrate effectiveness of our approach.

Keywords: FGS, rate-distortion optimization, unequal loss protection

1. Introduction

Transporting real-time video over the Internet while guaranteeing a required Quality of Service (QoS) level is a challenging problem. The difficulty can be attributed to the facts that digital video requires a huge transmission bandwidth even after compression and is very sensitive to variable transmission delays and packet losses. To provide reasonable video quality at high loss rates, it is important to make the source coders error-resilient and network-adaptive.

Bandwidth variation is one of the primary characteristics of "best-effort" networks, and the Internet is a prime example of such networks [1]. Consequently, for streaming applications, the selected video scheme should be capable of adapting to the unpredictable variation in bandwidth over the Internet [2]. Fine-granular-scalability (FGS) has been recently developed to meet these requirements [3] [4], and has also been adopted by MPEG-4 as the video-coding tool for streaming applications.

As we know, packet loss is unavoidable in the Internet and may have a significant impact on perceptual quality. Thus, some mechanisms must be in place to maximize video presentation quality in the presence of packet loss. The normally used approaches to recover from packet loss include retransmission-based scheme such as automatic repeat request (ARQ) and forward error correction (FEC). While ARQ techniques are effective in providing reliability, they can result in significant and unpredictable delay, making ARQ unsuitable for applications that have stringent real-time constraints. Therefore, FEC is a more appropriate error control method for real-time video applications.

It should be noted that not all of the transmitted bits have the same importance in FGS encoded video streams. The data dependency in the enhancement layer (EL) has to be taken into consideration such that if packets containing more significant bit-plane get lost, packets containing the less significant bit-plane in the same region should be discarded anyway. Consequently, more protection should be assigned to the important data. In [5]-[7], unequal packet-loss protection strategy using FEC has been used for transmission of layered scalable coding video stream and has already been proven to result in gracefully degrading transmission schemes. However, in these papers, they only added additional redundancy FEC data and did not consider the restriction of the transmission bandwidth.

In this paper, packet-loss robustness of the MPEG-4 FGS video coding technique in conjunction with unequal loss protection (ULP) is analyzed to provide graceful degradation of video quality in the packet erasure environment. Considering of the restriction of transmission bandwidth, joint source-channel rate-distortion (R-D) based optimization is performed to allocate the bit rate. Specially, in our scheme the optimal bit allocation is dynamically adjusted according to varying video characteristics and network conditions.

2. Unequal Loss Protection for FGS

2.1 System Overview

Fig. 1 shows the end-to-end architecture of the system that uses our proposed ULP for FGS streaming. The whole system is composed of the stream sender and receiver. In the sender, there are FGS Source Encoder model, ULP Assignment model and the Network Status Collection model. The receiver consists of ULP Decoder model and the Source Decoder model.

On the sender side, FGS Source Encoder encodes the raw video into BL and EL. The embedded R-D Extraction unit collects the rate-distortion information of the encoded

0-7803-8185-8/03/$17.00 © 2003 IEEE

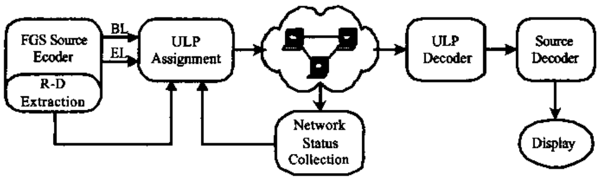

Fig. 1. End-to-end architecture of the system that uses ULP to protect scalable video streaming

stream in real-time. The Network Status Collection model periodically monitors the network status and some related parameters, such as packet loss ratio and the available bandwidth, are calculated. ULP Assignment model is the most important one in this system. In this model, according to the network condition, joint source-channel rate-distortion optimization is performed to achieve the minimal expected distortion, and the final optimized results are used to regulate the source code transmission rate and allocate the ULP channel code bits. In the receiver, the input stream is first processed by the ULP Decoder, and then the reconstructed bitstream is directed to the source decoder and display.

During the ULP assignment, the difficulty encountered in joint bit allocation between source and Internet channel is how to add FEC so that the decoder can recover the lost data correctly. Obviously, under a given channel rate, the additional FEC data reduce the available transmission rate for source coding, thus resulting in a trade-off between source coding and FEC.

In our proposed ULP for FGS streaming, joint source-channel R-D optimization is performed to allocate bits between source and channel code. Specially, in our scheme the optimal bit allocation is dynamically adjusted according to varying video characteristics and network conditions.

For the joint source-channel coding, the end-to-end distortion, $D(R)$, is composed of the source distortion and the channel distortion. The source distortion is caused by media rate control; while the channel distortion results from network packet loss. So, let $R(t)$ denotes the network bandwidth available for transmission at time t. Let $R_S(t)$ and $R_{FEC}(t)$ denote the FGS source rate and the FEC channel rate, respectively. Further more, let $D_S(t)$ and $D_{FEC}(t)$ represent FGS source distortion and distortion of FEC code, respectively. Then in terms of mathematics, the problem can be written as:

$$\begin{aligned} &\text{minimize } D = D_S(t) + D_{FEC}(t), \\ &\text{subject to} R(t) \geq R_S(t) + R_{FEC}(t). \end{aligned} \tag{1}$$

In order to calculate the end-to-end distortion, we must get the source distortion estimation and the channel distortion estimation, so we can optimize the bit allocation according to (1).

2.2 Source Distortion Estimation

The performance of the model based methods for rate allocation depends heavily on the accuracy of the selected model. And the commonly used exponential mo del is not suitable to accurately model the R-D properties of FGS enhancement layer data at low bit-rate. To overcome this problem, Zhang *et al.* proposed an R-D labeling scheme to characterize the R-D relationship of the source coding process in [8]. A set of actual R-D points is sampled in the encoding process and linear interpolation is used to estimate the actual R-D curve.

In our system, an R-D extractor in DCT domain is embedded in the FGS encoder to obtain the R-D information in real-time. In the ULP Assignment model, piecewise interpolation between data points is used to approximate the complete R-D curve of the source coding. Consider any two neighbor sampling points $D(R_m)$ and $D(R_n)$ along the R-D curve, where $R_m<R_n$. Let $\Delta R = R_n - R_m$ denote the difference between the two rates. The piecewise linear model is given below:

$$D_S(R_S) = D(R_m) - \frac{D(R_m) - D(R_n)}{\Delta R}(R - R_m) \tag{2}$$

where $D_S(R_S)$ is the distortion estimation when the FGS source rate is R_S, and $R_m=R_S=R_n$.

2.3 Channel Distortion Estimation

In this paper, we use Reed-Solomon (RS) codes to provide the ULP for the FGS coded video. Fig. 2 shows the detail of the unequal loss protection for one frame's FGS EL data. As shown in Fig. 2, the FGS EL data are assigned in layer 1 to L with some FEC bytes. We let f_i equal the number of FEC bytes assigned to layer i, so we can construct an L-dimensional FEC vector whose entries are the length of FEC assigned to each layer: $\overline{f} = (f_1, f_2, \cdots, f_L)$.

We define the total one frame's FGS EL data bytes as M. For a given $\overline{f}$, M is divided into fragments $M_i(\overline{f})$, which is the data bytes number of each layer, and we can get $M = \sum_{i=1}^{L} M_i(\overline{f})$. We denote a data set containing the data of the first i layers for the given FEC vector $\overline{f}$

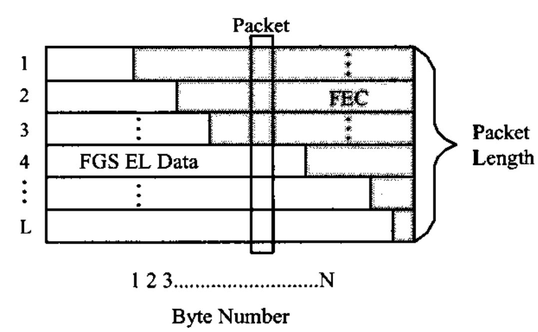

Fig. 2. Unequal loss protection for FGS enhancement layer data

as $M(\overline{f},i)=M_1(\overline{f})M_2(\overline{f})...M_i(\overline{f})$. We also define $d(M(\overline{f},i))$ as the distortion at the receiver when data set $M(\overline{f},i)$ is correctly decoded. So we can get $g_i(\overline{f})=d(M(\overline{f},i-1))-d(M(\overline{f},i))$, which is the amount of distortion reduction when the receiver decodes additional $M_i(\overline{f})$, given that all layers prior to i have already been decoded.

Because the embedded characteristic of the FGS EL data, we require that $f_i \ge f_{i+1}$; i=1, 2, 3, ..., L-1. With this requirement, if $M_i(\overline{f})$ can be decoded, then $M_1(\overline{f})$, $M_2(\overline{f})$, ..., $M_{i-1}(\overline{f})$ can also be decoded.

To evaluate the performance of an *RS(n, k)* (it contains k information packets and n-k redundancy packets) code in the Internet, we need to know the probability that more than *n-k* packets are lost since then the missing video packets cannot be reconstructed. We can compute this probability if we know the probability *P(m, n)* of *m* lost packets within the block of *n* packets. We use the 2-state Markov model [9] to estimate the network status. The Markov model is a renewal model, which is determined by the distribution of error-free intervals (gaps). Let a gap of length *v* be the event that after a lost packet *v*-1 packets are received and then again a packet is lost. The gap density function *g(v)* and gap distribution function *G(v)* can be derived. Let *R(m, n)* be the probability of *m*-1 packet losses within the next *n*-1 packets following a lost packet. Then the probability of m lost packets within n packets is:

$$P(m,n)=\begin{cases}\sum_{v=1}^{n-m+1} P_B G(v) R(m,n-v+1) & \text{for } 1\le m\le n\\ 1-\sum_{m=1}^{n} P(m,n) & \text{for } m=0\end{cases} \tag{3}$$

where P_B is the average packet-loss probability. The detailed derivation can be found in [6].

For a given FEC vector $\overline{f}$, using *P(m, n)*, we can calculate the probability that there are k or fewer packets lost in total N packets, $c(k)=\sum_{m=0}^{k} P(m,N)$. So the probability that layer i can be decoded by receiver is $c(f_i)$. Now, we can estimate the expected distortion of the *M* data after transmission for the given FEC vector $\overline{f}$:

$$D(\overline{f})=\sum_{i=1}^{L} c(f_i) g_i(\overline{f}). \tag{4}$$

To minimize the expected distortion $D(\overline{f})$, we use the hill-climbing algorithm described in [5] to search the appropriate FEC vector $\overline{f}$.

Now, for a given FGS source rate R_S (data bytes M of one frame can be calculated from R_S), we can get the channel distortion estimation $D_{FEC}(R_S)=\min D(\overline{f})$, for all $\overline{f}$, and the corresponding FEC bytes $R_{FEC}(R_S)=\sum_{i=1}^{L} f_i$.

2.4 Joint Source-Channel R-D Optimization

After getting the estimations of the source distortion and the channel distortion, we must optimize the bit rate allocation of the source coding and the FEC coding to satisfy (1). To reach the optimal solution, a search method is proposed. That is,

- Step 1: Set $R_S=R$ (R is the current transmission bandwidth), calculate $D_S(R_S)$ and set $D_{pre}=D_S(R_S)$;
- Step 2: Set $R_S=R_S-\Delta R$, calculate $D_S(R_S)$, $D_{FEC}(R_S)$ and $R_{FEC}(R_S)$;
- Step 3: If $R_S+R_{FEC}(R_S)>R$, then return to step 2; else, continue with step 4;
- Step 4: If $D_{current}=D_S(R_S)+D_{FEC}(R_S)<D_{pre}$, then set $D_{pre}=D_{current}$, record the current bit allocation solution, and return to step 2; else, finish the search and use the previous bit rate allocation solution as the final result.

To reduce the iterations, we can choose the search step size ΔR according to the channel condition. If the average packet loss ratio is small, it indicates that the network condition is good, and there are few FEC codewords needed, so we can select small step size to provide good accuracy. If the average packet loss ratio is large, it indicates that there are many packet losses in the network, and it needs more FEC codewords to protect source data, so we can choose the large step size to speed up the iteration.

3. Simulation Results

In this section, we present some of our experimental results obtained using simulations. We present comparisons of our ULP system with baseline FGS no protection transmission to show that our system can provide good graceful degradation in the presence of packet loss. Given an estimation of the network loss model, we compare our ULP system to baseline FGS no protection transmission by looking at frame PSNR values and average PSNR values corresponding to different packet loss ratios.

For our experiments, the Microsoft MPEG-4 FGS encoder/decoder is used. We simulate rather than use an actual network to allow greater control over the packet loss conditions. A 2-state Markov model [9] proposed by Gilbert is used to simulate packet loss in Internet channel.

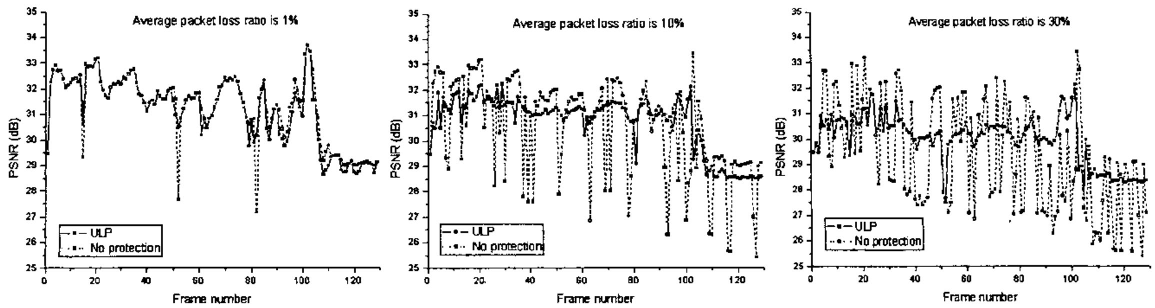

Fig. 3. Frame PSNR values for the Foreman sequence when the packet loss ration is 1%(left), 10%(center) and 30%(right).

TABLE I
COMPARISON OF AVERAGE PSNR

Average packet loss ratio	Average PSNR of ULP	Average PSNR of no protection
1%	31.2032 dB	31.2759 dB
10%	30.7338 dB	30.3830 dB
30%	29.9680 dB	29.1336 dB

We present results using the 400 frames QCIF Foreman sequence as the test sequence. The first frame is intra-coded and the remaining frames are coded as P frames. The testing sequence is coded at a temporal resolution of 10fps, and the base layer is encoded at a constant bit rate of 24 kbps using TM5 rate control. The results presented in this section use a target transmission bit rate of 128kbps. In our simulation, content protection is used to prevent packet loss of base layer, and we suppose that there are no packet losses in base layer stream.

Fig. 3 shows comparison results of frame PSNR for the Foreman sequence using our ULP system and no protection transmission scheme with different packet loss ratio. Table I depicts comparison results of average PSNR. Notice that the total available bandwidth is the same in all cases.

From Fig. 3 and Table I, it can be seen that our proposed ULP system obtains better results than the no protection scheme under packet loss network. And our system can adaptively regulate the FEC code rate according to the network conditions. Note that in Fig. 3 for some frames, the PSNR values of no protection scheme are higher than ours. This is because we spend some bits for error protection according to the source stream and the network condition. While for the no protection scheme all the bits are all allocated to the source. It can be seen from Fig. 3, when packet loss occurs, we achieve significant better PSNR compared to the no protection scheme. For FGS, since the prediction is always based on the BL, any data loss in one frame's EL would not affect the subsequent frames. So in Fig. 3, the frame PSNR values with no protection scheme fluctuate acutely. By contraries, our proposed ULP has smooth variation.

We also experimented with other video sequences and various settings of parameters (target transmission bit rate and network packet loss ratio) and found those results to be consistent with those presented in this section.

4. Conclusion

In this paper, we presented a joint source-channel R-D optimized unequal packet loss protection scheme for scalable video communication. Because of the different importance in FGS encoded video streams, we give more protection to the more important data. Considering of the restriction of the transmission bandwidth, R-D based bit allocation is used to minimize the end-to-end distortion. The simulation results demonstrated that our approach can adaptively regulate the FEC code rate according to the network conditions, and can provide graceful degradation of video quality in the presence of packet loss.

References

[1] V. Paxson, " End-to-end Internet packet dynamics," Proc. ACM SIGGCOM, Vol. 27, pp. 13-52, October 1997.

[2] J. Lu, "Signal processing for Internet video streaming: A review," Proc. IVCP, Vol. 2974, Proc. SPIE, pp. 246-259, Jan. 2000.

[3] H. Radha, M. van der Schaar, and Y. Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP," IEEE Trans. Multimedia, Vol. 3, No. 1, pp. 53-68, March 2001.

[4] W. Li, " Overview of fine granularity scalability in MPEG-4 video standard," IEEE Trans. Circuits Syst. Video Technol., Vol. 11, No. 3, pp. 301-317, March 2001.

[5] A. E. Mohr, E. A. Riskin, and R. E. Ladner, " Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," IEEE J. Select. Areas Commun., Vol. 18, No. 6, pp. 819-828, June 2000

[6] K. Stuhlmü ller, M. Link, and B. Girod, " Robust Internet video transmission based on scalable coding and unequal

error protection," Signal Process. Image Commun., September 1999.

[7] M. van der Schaar and H. Radha, " Unequal packet loss resilience for fine-granular-scalability video," IEEE Trans. Multimedia, Vol. 3, No. 4, pp. 381-394, December 2001.

[8] X. M. Zhang, A. Vetro, Y. Q. Shi, and H. Sun, " Constant quality constrained rate allocation for FGS video coded bitstreams," SPIE Conference on Visual Communications and Image Processing (VCIP), Vol. 4671, pp. 817-827, Jan. 2002.

[9] J. R. Yee, and E. J. Weldon, "Evaluation of the performance of error correcting codes on a Gilbert channel," IEEE trans. Comm., Vol. 43, no. 8, pp. 2316-2323, 1995.