FEC-Based Video Streaming over Packet Loss Networks with Pre-Interleaving

Jianfei Cai  Chang Wen Chen
Department of Electrical Engineering, University of Missouri-Columbia
{cai, cchen}@ee.missouri.edu

Abstract
In this paper, we present a FEC-based end-to-end error control scheme for video streaming over packet loss networks. We propose a novel robust video streaming system in which an interleaving is applied to the compressed bitstream before channel coding. The application of such pre-interleaving is able to improve the performance of video streaming over packet loss networks because pre-interleaving can simultaneously satisfy different requirements arising from both channel coding and source coding. We demonstrate with simulations that the proposed approach can achieve a significant improvement for video streaming over heavy packet loss networks.

1. Introduction

With rapidly growing demand for video streaming services, video streaming has received much attention for the last few years. Typically, a video streaming system consists of four major elements: the compressed video bitstream, the sender (or the server), the receiver (or the client) and the network [1]. It is well-known that there exists the problem of packet loss due to network congestion and delay in the packet-switching networks such as Internet and ATM networks. The problem of packet loss remains as one of the most challenging problems in video streaming applications. In order to achieve a high-quality video streaming, error correction schemes are usually combined with error concealment techniques to combat the packet loss. This is particularly necessary for moderate-to-heavy packet loss cases. Because of the stringent time constraint for video streaming, it is considered more beneficial to use forward error control (FEC) coding schemes than to apply more commonly used retransmission schemes.

In general, FEC schemes need to be integrated with appropriate packetization in order to efficiently make use of FEC channel coding capabilities. The packetization is expected to re-distribute the errors or erasures caused by packet loss to many channel blocks. However, the conventional design of FEC-based video transmission would result in significant performance degradation if there are residual errors remaining after channel decoding. This is because current video coding standards are all block-based and the redistribution of the errors by packetization may affect many source blocks if the channel coding cannot generate error-free video bitstream.

In this research, we propose a novel robust video streaming system in which an interleaving is applied to the compressed bitstream before the channel coding. The application of interleaving before channel coding will be able to greatly improve the performance of video streaming over packet loss networks. Section 2 will describe the characteristics of video streaming over packet loss networks and the different requirements for channel coding and source coding. Section 3 will describe the proposed system in detail. Section 4 will demonstrate the improved video streaming performance based on the simulations of the proposed approach. Section 5 concludes this paper with some discussions.

2. Characteristics of Video Streaming over Network

![Fig. 1. A conventional end system-based video streaming system at the transmitting end.](image)

A conventional end system-based video streaming system is shown in Fig. 1. First, channel coding is applied to the compressed video bitstream. In the case of block channel coding, such as the Reed-Solomon codes, every $K$ source symbols will generate a channel coded block of $N$ symbols. When we transmit these channel blocks over the network with simple direct packetization, packet loss will cause problem at the receiving end. This is because, in the case of direct packetization, the loss of a packet may severely damage a channel coding block so that the error...
recovery capability of the channel coding may be exceeded. On the other hand, the channel coding would be wasted when there is no loss of packet within a transmitted channel block [2].

2.1. Packetization and Channel Coding

To overcome this problem, orthogonal packetization is often employed to re-distribute the errors due to each lost packet to many channel blocks so that the bursty loss of each packet will become random errors in many channel blocks. Such packetization is equivalent to the symbol-level interleaving in digital communication and will lead to better performance in channel decoding at the receiver because it breaks up the bursty pattern of the packet loss.

For a simple symbol-level block interleaver, channel blocks of N symbols are loaded into a rectangular matrix row by row. After M rows are collected, in which M is the size of a network packet, symbols are read out column by column [3]. Each column constitutes a network packet. In this way, packetization is orthogonal to the channel coding direction and is able to re-distribute the packet loss to all the channel blocks within an interleaving matrix. Such packetization has been successfully applied to several video transmission systems [2, 4]. Furthermore, network packet loss is usually bursty, i.e., if one packet is lost, it is very likely that the consecutive packets may also be lost [2]. Packet-level interleaving can be further added after packetization to reduce the correlations among lost packets. However, this additional enhancement is obtained at the cost of much longer delay. Therefore, we consider only the symbol-level interleaving in this video streaming research.

2.2. Source Coding and Channel Coding

The design of packetization or interleaving as discussed above is to improve channel coding performance. For video streaming, the channel decoded bitstream needs to be decoded by a source decoder for final user application. It is true that, with orthogonal packetization, channel decoder is more effective in handling random symbol errors than the bursty symbol errors. However, many video communication studies based on MPEG standards [5, 6] show that the source decoder is more effective in handling bursty errors than the random errors. This is generally accomplished by embedding some error-resilient techniques, such as resynchronization words, in the compressed video bitstream. For the same average bit error rate, the source decoder can handle the bursty errors better than random errors since the standard video coding schemes are all block-based. It is easier for the source decoders to handle bursty errors within one or few video blocks than for them to handle random errors in many blocks. Therefore, conventional sequential application of channel coding followed by packetization would generate a bitstream that is incompatible with the block-based source decoder. It is expected that the packetization for the purpose of enhanced channel decoding may not result in an improved video bitstream reception because the residual random errors after channel decoding may lead to significant video quality degradation after source decoding.

3. Description of the Proposed System

Fig. 2 shows the diagram of the proposed video streaming system. At the transmitting end, the process is similar to a conventional video streaming system except that interleaving is applied before the channel coding. At the receiving end, it is essentially an inverse process as the transmitting end except that a postprocessing block is inserted before video source decoding. The function of postprocessing is to identify and delete those smallest independent units, such as slices in MPEG-2, that contain uncorrected erased symbols indicated by channel decoding.

The major innovation of the proposed approach is the application of the pre-interleaving. It is this pre-interleaving that results in bursty error patterns in the channel decoded video bitstream and allows the source decoder to handle bursty errors better than random errors since the standard video coding schemes are all block-based. It is easier for the source decoders to handle bursty errors within one or few video blocks than for them to handle random errors in many blocks. Therefore, conventional sequential application of channel coding followed by packetization would generate a bitstream that is incompatible with the block-based source decoder. It is expected that the packetization for the purpose of enhanced channel decoding may not result in an improved video bitstream reception because the residual random errors after channel decoding may lead to significant video quality degradation after source decoding.
Fig. 3. The comparison between the system with and without pre-interleaving. Top: with pre-interleaving. Bottom: without pre-interleaving.

coding and packetization are implemented within this matrix. To limit the error propagation, an EC frame contains only a multiple of the smallest independent unit. Therefore, any smallest unit can only exist in a single EC frame. The channel block size \( N \) and the packet size \( L_p \) are fixed. Based on current network conditions, i.e., the average packet loss rate \( P_L \) and the average bursty length \( L_B \), we can calculate the number of information symbols, \( K \), which we can put in a channel block. Then, the target total number of source symbols in an EC frame can be calculated as

\[
Target = K(L_p - 1) + \text{SizeOfHeadInfo},
\]

where \( \text{SizeOfHeadInfo} \) is a constant. Notice that the first row of the rectangular matrix contains some header information as well as source symbols. \( \text{SizeOfHeadInfo} \) represents how many source symbols are contained in the first row. The first row is well protected to ensure that the header information is transmitted error-free. After we obtain \( Target \), we can fetch a certain multiple of the smallest unit into the EC frame, where the total size of those units, \( T_s \), is closest to \( Target \). The actual \( K \) for each row can be determined by

\[
K = \begin{cases} 
\lfloor \frac{T_s}{L_p - 1} \rfloor + 1, & 0 \leq j < T_s \ mod \ (L_p - 1) \\
\lfloor \frac{T_s}{L_p - 1} \rfloor, & T_s \ mod \ (L_p - 1) \leq j < L_p
\end{cases}
\]

where \( T_s = T_a - \text{SizeOfHeadInfo} \).

The scheme without pre-interleaving is also shown in Fig. 3 for comparison purpose. It is clear that, without pre-interleaving, the packetization will re-distribute the channel error into all blocks of source bitstream within an EC frame and may result in more degradation after source decoding as we have discussed early.

4. Experimental Results

In this section, we present some experimental results to demonstrate that the proposed scheme can achieve superior performance comparing with the conventional scheme. In this research, we choose MPEG-4 as the source codec because the MPEG-4 standard provides several error-resilient tools: resynchronization markers, data partitioning, reversible variable length coding (RVLC) and header extension codes (HEC). However, the proposed video streaming system can be applied to any compressed bitstream provided that certain number of resynchronization words can be equally or nearly equally inserted in the compressed bitstream. We choose the network packet size \( L_p = 47 \) byte, the same as ATM payload with one byte for sequence number. The proposed video streaming system can be adopted for Internet video streaming if the Internet packet size is adopted.

We use 150 frames of "Foreman" in QCIF format as the test sequence. This sequence is coded by simple profile MPEG-4 coder with a frame rate of 15 fps and a source bit rate of 100 kbps. In MPEG-4, the smallest independent unit is called video packet (VP). The designed VP length is 93 bytes, and there is one I-frame every 50 frames. We adopt the GEC model [7] to generate bursty characteristics of packet-loss channels. The testing average packet loss rate \( (P_L) \) is from 0 to 10%, and the average burst length is fixed to 5 packets. We use the Reed-Solomon codes with the channel block size \( N = 128 \) and the allowed redundancy for error protection is fixed to 10% of the total bandwidth. We assume the header of the video object layer and the first I-frame are transmitted error-free. These assumptions can be justified because in practice we can always find a good timing to start, or we may use ARQ to guarantee that the short time interval at the beginning of the transmission is error-free. The rest of the MPEG-4 bitstream are then fed into the proposed error control system.

Fig. 4 shows the average PSNR results of Y component and the average throughput over 30 simulations under dif-
different packet loss rates. It is clear that the performance of the system with pre-interleaving is much better than that of the system without pre-interleaving. In the case of $P_l = 0.1$, the system with pre-interleaving achieved more than 5 dB gain in PSNR and about 30% improvement in throughput comparing with the system without pre-interleaving. As $P_l$ decreases, the difference between the systems with and without pre-interleaving becomes smaller. This is because more lost packets can be recovered via channel coding in the case of low packet loss rate. Therefore, the proposed pre-interleaving is more efficient when the network experiences moderate and high packet loss. Fig. 5 shows the average PSNR results of Y component for all frames over 30 simulations with $P_l = 0.1$ and $L_B = 5$. It is clear that the proposed video streaming system can obtain greatly improved video quality for all frames.

5. Conclusion

In this paper, a novel FEC-based pre-interleaving error control scheme is proposed for video streaming over packet loss networks. This system takes into consideration the characteristics of packet loss, channel coding, and source coding. We demonstrated that the pre-interleaving is able to generate the desired error patterns for video source coding while still preserving the features of the conventional combination of channel coding and packetization. We showed a greatly improved performance in video streaming with simulations of the network packet loss based on ATM cell payload size.

This innovation can be easily extended to wireless video as the wireless channels are often experiencing bursty errors due to fading. The extension to wireless video is currently under investigation. We are also studying the design of rate control scheme so that the proposed system can be made adaptive to the changing of the network bandwidth, for both wired and wireless links.

6. References


