OPTIMAL INTERLEAVING FOR 3-D ZEROTREE WAVELET VIDEO PACKETS OVER BURST LOSSY CHANNELS

Yi Zhao, Stanley C. Ahalt

The Ohio State University
Dept. of Elec. and Comp. Engineering
Columbus, OH 43201

Jianyu Dong

California State University, Los Angeles
Dept. of Elec. and Comp. Engineering
Los Angeles, CA 90032

ABSTRACT

Burst packet loss imposes significant quality degradation in wireless video streaming applications. This paper proposes a near-optimal packet interleaving method that maximizes the quality of 3-D wavelet video streamed over burst lossy channels with error concealment. The proposed method consists of two steps: 1) spatial interleaving is conducted during packetization so that the damage of the packet loss is dispersed and 2) temporal interleaving is applied to minimize the expected distortion of lost packets at the receiver under given delay constraints. To allow real-time implementation, an iterative exchange-based algorithm is developed to reduce the computational complexity to $O(N^2)$. In addition, to simplify estimating the characteristics of time-varying channels, a novel measurement, "temporal neighbor packet distance", is proposed as an alternative optimization criteria of the expected distortion. Experimental results show that the proposed interleaving method can improve the effect of error concealment by up to 50%.

1. INTRODUCTION

Many researchers have found that the burst characteristics of packet losses occurring on lossy networks leads to a more serious quality degradation than independent packet losses [1]. Interleaving, a traditional approach in channel coding to mitigate the effect of burst bit errors, has recently been investigated as a possible solution for streaming applications [2, 3, 4]. Having the advantage of no extra bit cost, interleaving also has its own limitations. First, interleaving causes extra delays. Second, since interleaving only disperses the damage caused by burst losses and facilitates error concealment, its performance is highly dependent on channel loss behavior and the error concealment technique employed by the receiver.

By far, most existing research on interleaving techniques for 3-D wavelet video streaming is focused on spatial interleaving in packetization [5, 6, 7]. Although effective in combating isolated errors, the aforementioned approaches cannot optimize the video quality under burst lossy channels since no effort has been made to joint optimize the effect of interleaving and error concealment. In this paper, we investigate the characteristics of the packetization of 3-D zerotree wavelet video, the behavior of burst lossy channels, and the distortion reduction of error concealment. We propose a novel approach to determine the best real-time packet interleaver to maximize the quality of 3-D wavelet video streaming. The proposed method consists of two steps: 1) spatial interleaving is conducted during packetization so that the damage of the packet loss is dispersed and 2) temporal interleaving is applied to minimize the expected distortion of lost packets at the receiver under given delay constraints. To allow real-time implementation, an iterative exchange-based algorithm is developed to reduce the computational complexity to $O(N^2)$. In addition, to simplify estimating the characteristics of time-varying channels, a novel measurement, "temporal neighbor packet distance," is proposed to serve as an alternative optimization criteria of the conventional approach, expected distortion. Experimental results demonstrated that the optimization based on temporal neighbor packet distance has almost the same performance as the one based on distortion.

This paper is organized as follows. In Section 2, we formulate the optimization problem of packet interleaving of the 3-D zerotree wavelet video. A real-time algorithm to determine the near-optimal interleaver based on the channel model and the error concealment technique is described in Section 3. Section 4 presents the simulation results needed to evaluate the proposed approach. Section 5 summarizes the paper.

2. PROBLEM FORMULATION

2.1. Markov Channel Model

In this paper, a first-order Markov model of burst packet-erasure channel is assumed, as shown in Fig. 1. This model has two states: the good state, S_G, with no packet loss, and the bad state, S_B, with all packets being lost. The transition probabilities are P_{GB} (from S_G to S_B) and P_{BG} (from S_B to S_G), respectively.

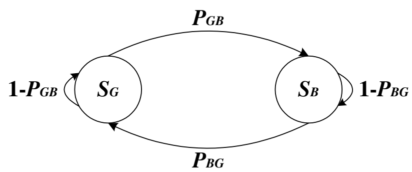


Fig. 1. Markov model for burst packet-erasure channel.

2.2. Optimization Goal

An optimal interleaver should maximize the received video quality under the given error concealment method. If the received video quality is measured by distortion and the optimal interleaving policy (transmission schedule) is denoted by Π, then the optimal interleaving problem can be formulated as:

$$\Pi = \arg\min_{\pi} E\left(D\left(\pi\right)\right), \tag{1}$$

where $E\left(D\left(\pi\right)\right)$ is the expected overall distortion for interleaving policy π.

Suppose the number of packets in one Group of Frames (GOF) is N, then due to the independent decodable packetization being employed, this expected overall distortion can be expressed as

$$E\left(D\left(\pi\right)\right) = \sum_{i=1}^{N} E\left(D_i\left(\pi\right)\right), \tag{2}$$

where $E\left(D_i\left(\pi\right)\right)$ is the expected distortion of the i^{th} packet under π. When the i^{th} packet is lost, the error concealment method is employed to approximate the content in this packet with its neighbor packet set $N\left(i\right)$. To describe the receipt status of the neighbor packets, a status vector, $\mathbf{S}_{N(i)}$, is defined as

$$\mathbf{S}_{N(i)}\left(j\right) = \begin{cases} 0, & j^{th} \text{ neighbor packet is lost} \\ 1, & j^{th} \text{ neighbor packet is received} \end{cases}. \tag{3}$$

The expected *error concealment distortion* of the i^{th} packet becomes

$$\overline{D_i^{EC}\left(\pi\right)} = \sum_{\mathbf{S}_{N(i)}} D_i^{EC}\left(\mathbf{S}_{N(i)}\right) P\left(\mathbf{S}_{N(i)} \mid \pi,\ i \text{ lost}\right), \tag{4}$$

where $D_i^{EC}\left(\mathbf{S}_{N(i)}\right)$ and $P\left(\mathbf{S}_{N(i)} \mid \pi,\ i \text{ lost}\right)$ are the error concealment distortion and conditional probability for $N\left(i\right)$ being status $\mathbf{S}_{N(i)}$, respectively.

Therefore, the optimization goal of temporal packet interleaving can be rewritten as

$$\Pi = \arg\min_{\pi} \sum_{i=1}^{N} \sum_{\mathbf{S}_{N(i)}} D_i^{EC}\left(\mathbf{S}_{N(i)}\right) P\left(\mathbf{S}_{N(i)} \mid \pi, i \text{ lost}\right). \tag{5}$$

3. OPTIMAL PACKET INTERLEAVING

3.1. Error Concealment Distortion

The conditional probability $P\left(\mathbf{S}_{N(i)} \mid \pi,\ i \text{ lost}\right)$ can be obtained from the assumed Markov channel model. Due to strict limits on algorithm delays in real-time video streaming, the error concealment distortion of one packet p_i is estimated by its error concealment error (ECE) in wavelet domain

$$D_i^{EC} = \sum_{xyz \in \Omega_i} \left(c_{x,y,z} - \hat{c}_{x,y,z}\right)^2, \tag{6}$$

where Ω_i is the set of coefficients contained in p_i and $\hat{c}_{x,y,z}$ are the estimated coefficients. Based on our experimental results as shown in Fig. 2, Equ. (6) provides an excellent approximation of error concealment distortion.

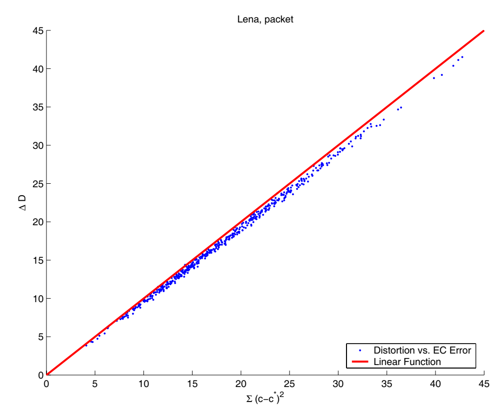


Fig. 2. Error concealment distortion is linear proportional to error concealment error.

3.2. Iterative Exchange-Based Search

Since the optimization presented by Equ. (5) is an *NP* problem, the straightforward *exhaustive search* has a computational complexity of $N!$ and its implementation is impractical for real-time video streaming applications. To significantly reduce the complexity, we propose an *iterative exchange-based search* algorithm to find a sub-optimal solution to Equ. (5). This fast searching algorithm checks each packet pair and exchanges their positions in the transmission schedule when such an exchange leads to a decrement of the expected distortion at the receiver:

1. Set the initial schedule so that $p_i = i$ for $i = 1, \cdots, N$. Compute the expected distortion, D^{cur}, via Equ. (2), (4), and (6) for this initial schedule.

2. Let $i = 1$.

   (a) Repeat Steps (b) and (c) while $i < N$.

   (b) Let $j = i + 1$.

      i. Repeat Steps ii and iii while $j \leq N$.

      ii. Exchange packets at the i^{th} and the j^{th} positions and compute the expected distortion D^{new} for the new schedule. If $D^{new} < D^{cur}$, keep this exchange. Otherwise, discard it.

      iii. Increment j by 1.

   (c) Increment i by 1.

3. Repeat Step 2 for P times.

The computational complexity of this fast sub-optimal searching algorithm is significantly reduced to only N^2P and makes implementation practical.

3.3. Temporal Neighbor Packet Distance

Using distortion as optimization criteria, calculating error concealment distortion and conditional probability is essential, but involves a lot of computation. To further reduce computational complexity, a new optimization criteria is proposed.

Due to the first-order Markov channel model, increasing the temporal distance between the packets will reduce the probability of losing them simultaneously. Therefore, we introduce two quantitative measurements of the temporal relationship among the packets, namely *temporal packet distance* and *temporal neighbor packet distance*. They are defined as follows:

Definition 1: Assuming that the packets appeared in the i^{th} and j^{th} positions of the transmission schedule are denoted by p_i and p_j, respectively, the temporal packet distance between the two packets is defined as

$$d_T\left(p_i, p_j\right) = \left|i - j\right|. \tag{7}$$

Definition 2: For a packet p_i, its temporal neighbor packet distance is defined as the minimum temporal packet distance between p_i and any packet in its neighbor packet set $N\left(p_i\right)$

$$d_T^{NB}\left(p_i\right) = \min_{p_j \in N(p_i)} d_T\left(p_i, p_j\right). \tag{8}$$

The newly proposed optimization criteria is based on the aforementioned definitions. Assuming that each packet's content is of equal importance, the interleaving optimization can be achieved by maximizing the average temporal neighbor packet distance for all packets from one GOF:

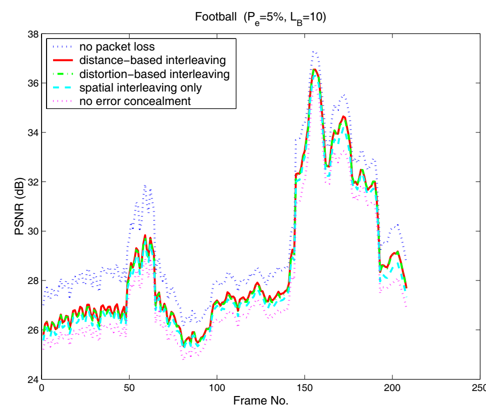
$$\Pi_d = \arg\max_{\pi} \overline{d_T^{NB}\left(p_i\right)}. \tag{9}$$

From Equ. (9), the optimal interleaving policy Π_d does not depend on specific network characteristics. This alleviates the challenges posed in estimating the channel parameters for a time-varying burst lossy network.

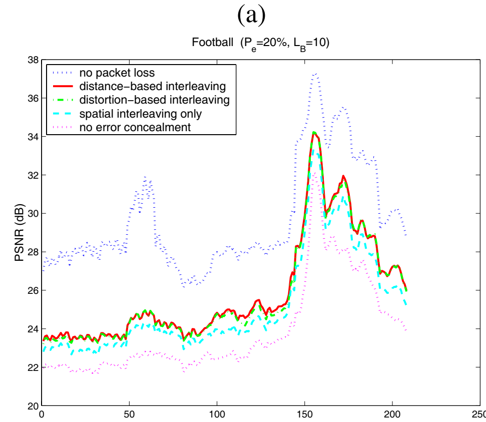
4. EXPERIMENTS AND DISCUSSIONS

To evaluate the performance of our proposed packet interleaving scheme, we have conducted a series of simulations to test four different error control approaches: 1) spatial interleave only, 2) spatial interleave with error concealment, 3) distortion-based optimal interleaving, and 4) neighbor packet distance-based optimal interleaving.

The frame-by-frame PSNR comparisons of *Football* sequences are shown in Fig. 3. As can be observed, both



(a)



(b)

Fig. 3. Frame-by-frame PSNR comparison of *Football* sequence at packet loss rates of (a) 5% and (b) 20%.

the distortion-based and neighbor packet distance-based approaches achieve much better reconstructed video quality under both low- and high-packet loss rates. For *Football* sequences, error concealment can improve the average quality by $0.52\mathrm{dB}$ and $1.33\mathrm{dB}$ for a packet loss rate of 5% and 20%, respectively. The proposed interleaving approach can further improve the error concealment effect by $0.27\mathrm{dB}$ and $0.71\mathrm{dB}$, respectively. Overall, the interleaving approaches improve the performance of error concealment by about 50%. In addition, the neighbor packet distance-based approach achieves almost the same performance as the distortion-based approach, but with much lower computational complexity. Therefore, the distance-based approach is very suitable for real-time streaming applications.

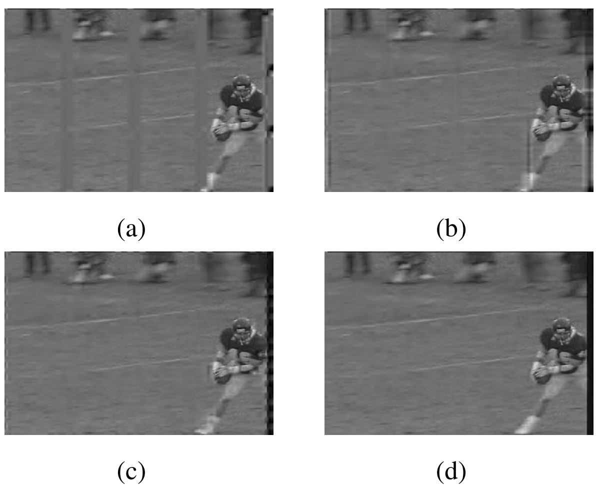


Fig. 4. Reconstructed images of frame 167 in *Football* sequence with different schemes: (a) spatial interleaving only, (b) spatial interleaving with error concealment, (c) distance-optimized interleaving, and (d) no packet loss.

5. CONCLUSIONS

In this paper, we have presented a near-optimized interleaving method to mitigate the burst packet loss in wireless video streaming. Special effort has been made to reduce the optimization algorithm's complexity to allow real-time implementation. The proposed iterative exchange-based search reduces the complexity from $O(N!)$ to $O\left(N^2\right)$. Furthermore, a new optimization criteria, "temporal neighbor packet distance," is proposed to save the computational cost of the expected distortion. A 3-D wavelet video streaming system was developed to evaluate the performance of the proposed approaches. The experimental results on multiple video sequences demonstrate that the effect of error concealment techniques can be consistently improved by the proposed approaches.

6. ACKNOWLEDGMENT

This publication was partially funded through the support provided by DoD HPCMP PET activities through Mississippi State University under the terms of Agreement No. # GS04T01BFC0060. The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of DoD or Mississippi State University.

7. REFERENCES

[1] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, June 2000.

[2] J. Cai and C. W. Chen, "Use of pre-interleaving for video streaming over wireless access networks," in *Proc. ICIP 2001*, October 2001, pp. 934–937.

[3] S. Aravith, C.-W. Lin, S. Roy, and M.-T. Sun, "Wireless video transport using conditional retransmission and low-delay interleaving," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 558–565, June 2002.

[4] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Model-based delay-distortion optimization for video streaming using packet interleaving," in *Conference Record of the 36th Asilomar Conference on Signals, Systems and Computers*, November 2002, pp. 1315–1319.

[5] J. K. Rogers and P. C. Cosman, "Wavelet zerotree image compression with packetization," *IEEE Signal Processing Letters*, vol. 5, no. 5, pp. 105–107, May 1998.

[6] S. Cho and W. A. Pearlman, "A full-featured, error-resilient, scalable wavelet video codec based on the set partitioning in hierarchical tree (SPIHT) algorithm," *IEEE Trans. Circuits Systems for Video Technology*, vol. 12, no. 3, pp. 157–171, March 2002.

[7] I. V. Bajic and J. W. Woods, "Domain-based multiple description coding of images and video," *IEEE Trans. Image Processing*, vol. 12, no. 10, pp. 1211–1225, October 2003.