Error Concealment Protection for Loss Resilient Bitplane-Coded Video Communications

*Chih-Ming Fu, *Wen-Liang Hwang , and Chung-Lin Huang*
Electrical Engineering Department, National Tsing-Hua University
*Institute of Information Science, Academia Sinica, Taiwan

Abstract—In this paper, we propose an error control scheme for video communications over lossy channels. The proposed algorithm uses error concealment protection (ECP) approach to coordinate with multiple encoded streams which is capable of handling error concealment to achieve robust transmissions. Unlike previous methods, our algorithm focuses on joint optimizing multiple substreams distortion with concealment over error prone channels. The algorithm minimizes the expected rate-distortion function to achieve the optimal FEC result. In the experiments, we demonstrate the effectiveness of our method by using a 3-D SPIHT algorithm. Simulation results show that the proposed protection strategy achieves about 2 dB higher peak signal-to-noise ratio compared to conventional method.

I. INTRODUCTION

Dealing with packet losses in video communications is an active research area. Average coding and error concealment result is taken as the distortion introduced by the channel. The set partitioning in hierarchical trees (SPIHT) [1] algorithm generates an embedded bitstream for video for rate adaptation under time-varying network. The rate changes of bitstream can be accommodated instantaneously by dropping appropriate packets. However, embedded bitstreams are very sensitive to data losses. Retransmission request for loss packets may not be acceptable for some video streaming applications because of the delay constraints of continuous playback at the receiver.

Forward error correction (FEC) is an appropriate error control method. Furthermore, progressive bitstreams provide a natural basis for unequal error protection (UEP), by which more important parts of the bitstream are assigned a greater level of error protection. As a result, many researchers use FEC-based error protection methods as promising error control solutions for video streaming over lossy channels [2]–[5]. Recently, several methods that use multiple correlated substreams transmitted independently have been proposed for wavelet-based coders to get additional error-resilience at high loss rates [6]–[11]. The advantage of this multiple-substream approach is that errors in one substream do not affect the others because each substream is independently decodable. This provides some robustness against packet losses. Unlike previous methods, our

algorithm focuses on joint optimizing concealment distortion over error prone channels. The algorithm minimizes expected rate-distortion function to achieve the optimal FEC result.

Our approach combines error concealment and distortion function for UEP and performs the optimal source protection bits assignment for the input video under different channel status. To solve the video coding and transmission problem, we introduce a probabilistic "expected distortion" point of view for achieving complete optimization. In the next section, we introduce the ECP and how it can be implemented on a 3-D SPIHT video codec. In Section 3, we present the experimental setup and results.

II. MODELING DISTORTION OF ERROR-CONCEALMENT AND ERROR-PROTECTION

The multiple substreams are generated for effective error-concealment at the receiver. We consider temporal-domain partitioning schemes to generate the sub-streams, and the even frames and odd frames in a GOP are used to form the two subsequences, as depicted in Figure 1. Each subsequence is independently transformed and encoded with a 3-D SPIHT algorithm.
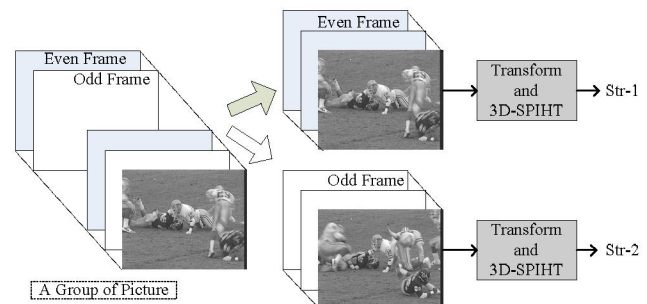

Figure 1. Temporal-domain partitioning. scheme.

A. Error Concealmen for Bit-plane-coded Bitstream

We assume that one of the two substreams has a smaller number of decodable bit planes because of excessive packets loss. We can estimate the missing bit-plane values of *str-s* (where *s* is stream index), using the correlation between the two substreams; we can enhance the reconstructed picture

quality at the receiver. The philosophy behind the error-concealment algorithm is to refine one bit plane at a time, and each missing coefficient value in a particular bit plane is either 0 or 1.

The correct choice (between 0 and 1) for the coefficients in the missing bit plane will result in a smoother reconstructed GOP most of the time. Assuming that the two substreams have a strong correlation, the correct choice of the missing bit-plane values will minimize a distortion metric. We define a distortion metric that measures the smoothness between the two substreams in the wavelet transform domain, similar to the one proposed in [6]. A simple and fast algorithm that estimates the missing bit-plane values of the low-quality substream using the correlation between the two substreams has been proposed in [6]. They use an additive distortion metric which approximates the mean-squared error (MSE), as follows:

$$D = \sum_x c_h(d,x) - \hat{c}_l(d,x) \qquad (1)$$

Let x denote the pixel of frames in a GOP. In addition, $c_h(d, x)$ is a wavelet coefficient after decoding up to the d-th bit-plane for the high-quality substream, and $\hat{c}_l(d, x)$ is an estimated wavelet coefficient for the low-quality substream. The optimal estimates of the missing bit-plane values for the low-quality substream are the ones that minimize the distortion metric defined in (1). To find the optimal solution for this problem is to enumerate all possible reconstruction values and choose the one that minimizes the distortion metric.
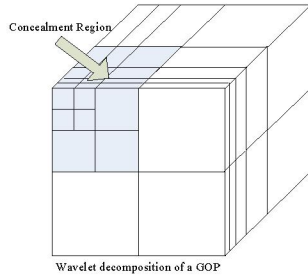

Figure 2. Shaded area specifies the wavelet coefficients that are refined using error-concealment algorithm.

We modify the algorithm proposed in [6] and select the values of the refined decoding map for the coefficients in a larger concealment region (shown in Figure 2) as the ones that minimize the distortion metric among the possible choices. Since the number of possible states in a refined decoding map for a given state in a decoding map in a bit-plane is limited, the computational complexity of the concealment method is very low. Our algorithm considers the case where more than one bit plane has been lost by estimating the missing bit planes one bit plane at a time, starting from the highest missing bit plane. We define the concealment ratio of a bit-plane as follows

$$\beta(i) = \frac{MSE_c(i+1) - MSE_l(i)}{MSE_l(i+1) - MSE_l(i)}$$

where $MSE_c(i)$ is the MSE of the lower quality bitstream with the missing bit-plane-i had been concealed and $MSE_l(i)$ is the MSE of the low-quality substream after decoding up to the ith bit-plane. Figure 3 shows the improvement of the concealment technique of Football and Foreman sequence where a GOP contains 32 frames.
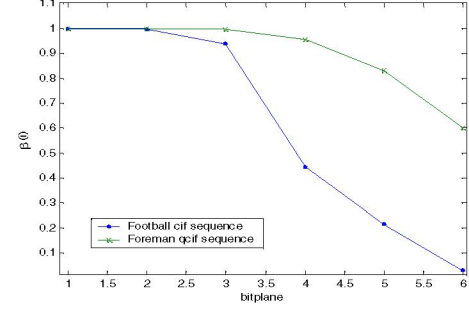

Figure 3. The performance improvement of Football sequence by using our concealment technique. Solid: Proposed. Dashed: Proposed in [6].

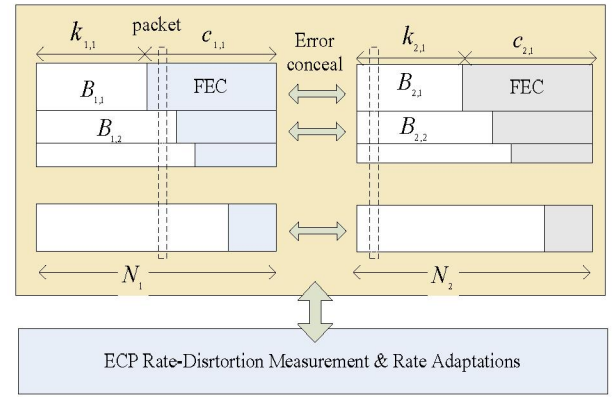## B. Error Concealment Protection Algorithm

Figure 4. Framework of error concealment protection for two substreams that have different minimum threshold values.

We define $b_{s,i}$ as the size of bit-plane i of bit-stream s in bytes. Since the number of source data bytes of bit-plane i, $b_{s,i}$, is sometimes not divisible by $k_{s,i}$, the number of source data bytes protected by an $(N_s, k_{s,i})$ code is

$$B_{s,i} = \left\lfloor \frac{b_{s,i-1} - B_{s,i-1} + b_{s,i}}{k_{s,i}} \right\rfloor \times k_{s,i}$$

As shown in Figure 4, the total number of transmitted packets is $N = N_1 + N_2$ and $c_{s,i} = N_s - k_{s,i}$. We also define the length that bit-plane in each packet as

$$l_{s,i} = \frac{B_{s,i}}{k_{s,i}}$$

The proposed ECP algorithm finds a redundancy assignment vector

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,I-1} & c_{1,I} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,I-1} & c_{2,I} \end{bmatrix}.$$

Let us define

$$B_s(i,C) = \sum_{j=1}^{i} B_{s,j}$$

as the accumulated number of source data bytes up to and including bit-plane i and $MSE_s(i,B(i,C))$ as the mean square error (MSE) at the receiver relative to the original video sequence when the receiver decodes source-s $B_s(i,C)$ data bytes, given the redundancy assignment vector C.

$$\Delta D_{s,i} = MSE_s(i,B(i,C),N_1) - MSE_s(i-1,B(i-1,C),N_2) \ , \ s = 1,2$$

There are two independent substreams and each one is protected by $(N_s, k_{s,i})$ RS-code, where s is stream index, and i is the i-th bitplane. Although using PSNR in [13] is easier, we use mean square error (MSE) to measure the performance accuracy of each substream The expected distortion reduction of bitplane-i is

$$\Delta D_i(C,N_1,N_2) = P(N_1,k_{1,i})P(N_2,k_{2,i})(\Delta D_{1,i} + \Delta D_{2,i})$$
$$+ (1 - P(N_1,k_{1,i}))P(N_2,k_{2,i})(\beta(i)\Delta D_{1,i} + \Delta D_{2,i})$$
$$+ P(N_1,k_{1,i})(1 - P(N_2,k_{2,i}))(\Delta D_{1,i} + \beta(i)\Delta D_{2,i})$$

where the $P(N_s,k_{s,i})$ is the probability of stream-s bit-plane-i recovered correctly and $\beta(i)$ is the performance improvement of the video sequence for the loss bit-plane-i.. The performance measurement joint consider the error correction code and error concealment of both substream. We assume that each substream is transmitted independently; therefore the correct probability is separable. The additional distortion reduction cause by concealment technique is $\Delta D_{s,i}^c$, which is occur when substream-s layer-i is miss and which can be concealed by the other substream. The rate allocation algorithm is based on expected distortion reduction function. We can now calculate the expected MSE reduction at the receiver as

$$EMSE(C,N_1,N_2) = \sum_{i=1}^{I} \Delta D_i(C,N_1,N_2) \qquad (2)$$

The problem of finding a redundancy assignment vector that maximizes the expected quality in (2) under the constraint of total available bytes for a GOP can be expressed as

$$\max_{C} EMSC(C) \quad \text{subject to} \quad NL \le B \qquad (3)$$

where N is the number of packets to be transmitted for a GOP, L is the packet size in bytes, and B the is rate constraint. The optimal solution for this problem can be found by using an exhaustive search method. However, the use of an exhaustive search method is unrealistic for real-time video because of the excessive amount of computation required. Instead, the maximization problem can be solved by an approach based on dynamic programming. In [5], a similar problem has been addressed for the joint source-channel coding of images that are progressively transmitted over noisy channels. The principal idea of the solution is to write the expected quality as a sum of incremental rewards, which are accumulated as each bit-plane is successfully decoded by the receiver. For a redundancy assignment vector, let us define

$$EMSE(i,C) = \sum_{j=i}^{I} \Delta D_i(C,N_1,N_2) \qquad (4)$$

We can verify that satisfies the following recursion:

$$EMSE(i,C,N_1,N_2) = \Delta D_i(C,N_1,N_2) + EMSE(i+1,C,N_1,N_2)$$

Note that, for a vector C, $EMSE(i,C)$ do not depend on $\{c_{s,1},c_{s,2},...,c_{s,i-1}\}$. Hence, the solution to the maximization problem in (3) can be specified only over a sub-sequence $\{c_{s,i},c_{s,i+1},...,c_{s,I}\}$. Based on the recursive nature of specified in (4), we can apply a dynamic programming approach to solve the maximization problem in (3). for notation convenient, let $EMSE^*(i,B)$ denote the optimal value of the objective function in (3) and satisfy the following dynamic programming equation:

$$EMSE^*(i,B) = \max_{c_{1,i} > c_{1,i+1}, c_{2,i} > c_{2,i+1}, N_1 + N_2 = N} \left(\Delta D_i + EMSE^*(i+1,B-B_i) \right) \quad (5)$$

The sub-sequence $\{c_{s,i+1}^*, c_{s,i+2}^*, ..., c_{s,I}^*\}$ will solve the maximization problem in (3) for a constraint of $B - B_i^*$ bytes. Here, we require $c_{s,i} \ge c_{s,i+1}$ for progressive bitstream transmission and the total number of transmitted packet N.

III. EXPERIMENTAL RESULTS

We present the simulation experiments to illustrate the performance of the proposed algorithm. First we compared the performance of the proposed algorithm with [6] at different rate constraint. Second, we studied the performance of the proposed algorithm under various channel condition. We have encoded 160 frames Football sequence at CIF resolution and Foreman sequence at QCIF resolution. The channel model is 2-state Markov model with mean burst length 5. We define that P_b is the mean packets loss rate, L_b is the mean burst length, and Bpp is bits per pixel. At each loss rate, we perform 20 simulations for each video sequence for the PSNR. We see that the proposed system is better than independent rate allocation algorithm. The performance of

error concealment is caused by the highly correlation between two substreams. In Figure 5, the proposed method is significant better then [6] for $N=30$, because our method can allocate source and channel efficiently. In Figure 6 we show the performance advantage of our method will decrease for $N=62$ because [6] method will generate more effective FEC bits allocation for error protection. However, larger N means more small-size packets will be generated. Since the overhead of packet transmission for small-size packet is bigger than the large-size packet, we prefer small N. Figure 7 shows the performance in unbalance channel and the performance is better than that in balance channel.
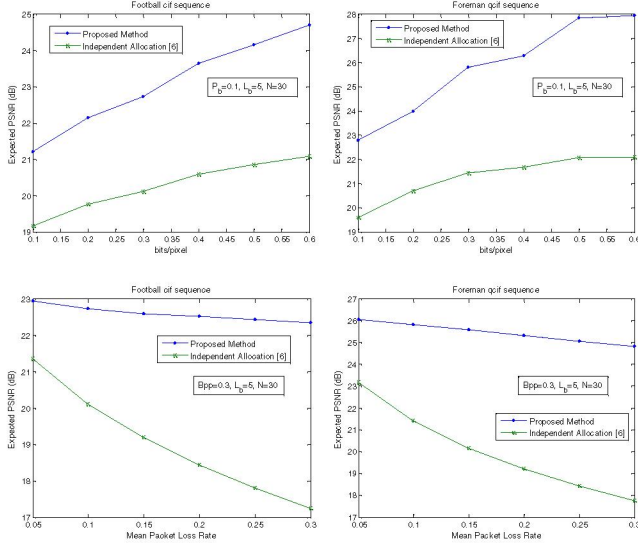

Figure 5.   Performance of the proposed algorithm applied to 3D-SPIHT bitstream
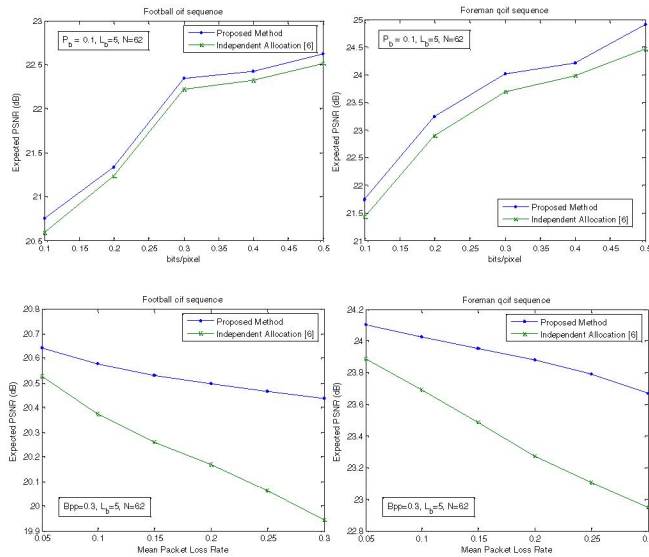

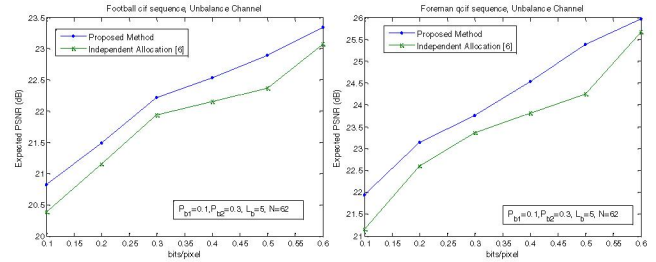Figure 6.   Performance comparision for $N=62$.

Figure 7.   Unbalance channel condition

IV.   CONCULSION

We have presented an error concealment protection method for video communications over lossy channels. Our method focuses on reliable estimation of decoder distortion and joint bits allocation for both error-concealment and error-protection. We have shown that the proposed method outperforms the popular methods under various channel condition. To exploit the benefits of error concealment of other video codec is worthy for feature work.

REFERENCES

[1]   A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 4, pp. 243–250, Jun. 1996.

[2]   A. Mohr, E. Riskin, and R. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channel through forward error correction," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 818–828, Jun. 2000.

[3]   "Approximately optimal assignment for unequal loss protection," presented at the IEEE Int. Conf. Image Processing, Oct. 2000.

[4]   R. Puri, K. W. Lee, K. Ramchandran, and V. Bharghavan, "An integrated source and congestion control framework video streaming in the internet," *IEEE Trans. Multimedia*, vol. 1, no. 1, pp. 18–32, Mar. 2001.

[5]   J. Kim, R. M. Mersereau, and Y. Altunbasak, "Error-resilient image and video transmission over the internet using unequal error protection," *IEEE Trans. Image Process.*, vol. 12, no. 2, pp. 121–131, Feb. 2003.

[6]   J. Kim, R. M. Mersereau, and Y. Altunbasak, "A Multiple-Substream Unequal Error-Protection and Error-Concealment Algorithm for SPIHT-Coded Video Bitstreams," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1547–1553, Dec. 2004.

[7]   C. Creusere, "A new method of robust image compression based on the embedded zerotree wavelet algorithm," *IEEE Trans. Image Process.*, vol. 6, no. 10, pp. 1436–1442, Oct. 1997.

[8]   S. Cho and W. Pearlman, "Error resilient compression and transmission of scalable video," *Proc. SPIE*, vol. 4115, pp. 396–405, 2000.

[9]   A. A. Alatan, M. Zhao, and A. N. Akansu, "Unequal error protection of SPIHT encoded image bit streams," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 814–818, Jun. 2000.

[10]   S. Cho and W. Pearlman, "A full-featured, error-resilient, scalable wavelet video codec based on the set partitioning in hierarchical trees algorithm," *IEEE Trans. CSVT*, vol. 12, no. 2, pp. 157–171, Mar. 2002.

[11]   J. Kim, R. M. Mersereau, and Y. Altunbasak, "Distributed video streaming using multiple descryiption coding and unequal error protectiton" *IEEE Signal Processing.*, vol. 14, no. 7, pp. 849–861, July. 2005.