DESIGN ERROR-RESILIENT MULTIPLE SUBSTREAMS 3D CODER INCLUDING RECEIVER POST-PROCESSING IN ANALYSIS

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

We propose an error control scheme for video communications over lossy channels. The proposed algorithm uses receiver post-processing in analysis to coordinate with multiple encoded streams which is capable of handling error concealment and error protection 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.

***Index Terms*—** Forward error correction, error concealment

1. INTRODUCTION

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 [1]–[3]. 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 [4]–[6]. 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. The multiple substreams approach not only provide transmission robustness but also provide receiver-end post-processing to conceal error artifact even that the transmitted bits is lost and can not be recovered.

Unlike previous methods, we include the receiver post-processing in analysis to design a joint optimized error-resilient video transmission framework. This will incorporate the joint source channel coding (JSCC) of encoder and the post-processing of decoder to construct a novel error-resilient framework. We deal with the coordination of total distortion of all video transmission process including source encoder, channel encoder, channel estimation, channel decoder, source decoder, and post-processing. To solve this problem, we use the expected rate-distortion-function to construct a unified measurement which combines the channel error probability and error concealment parameter

The rest of the paper is organized as follows. In the next section, we review the background knowledge of error-resilient multiple substreams 3D video coder and explain how it can be applied into a generalized scalable video codec. In Section 3, we formulate the constraint problem for a single layer case and introduce the property encoding transmission (PET) to deal with the multi-layer case problem and propose an algorithm to search the optimal solution. The simulation results are given in Section 4 and Section 5 concludes the paper.

2. BACKGROUND

For a scalable video coder, it is necessary to introduce the layer concept to advantage of unequal error protection. A scalable bitstream can be decoded progressively and the bitstream need to be protected unequally according their significance order. We show four layers case and assign unequal amount of Forward Error Correction (FEC) Code according to their priority. Because the higher layers need the lower layers data to decode the received data, the packetization format and the data recover probability should satisfy that the lower layers could be recovered as decoding the higher layers. We define the protection rate of layer j as $r_p^{(j)} = (n-k^{(j)})/n$ which is corresponding to the $(n,k^{(j)})$ RS error correction code and let $c^{(j)} = n - k^{(j)}$. The protection rate should satisfy the PET constraint

$$c^{(1)} > c^{(2)} > ... > c^{(L)},$$

where L is the total number of layers we used. We interlace the data of different layers into a single packet and transmit these packets in the order as shown in Fig. 1, where the dash line box represents a packet. This packetization structure ensures that each layer has the same number of loss packets and the lower layers protect more than higher ones.
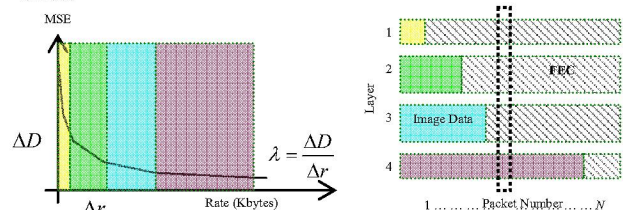
Fig.1. Multi-layer unequal error protection.

2.1. Receiver-end Post-processing and Error Concealment Concept

To avoid error propagation, we separate a single video into two independent encoded bitstream and if only one bitstream is corrupted, the others will maintain an acceptable decoded video quality. The multiple substreams will be generated for a receiver to apply the error-concealment technique. We consider the temporal-domain partitioning scheme to generate the sub-streams, and the even frames and odd frames in a GOP will generate two sub-sequences, as depicted in Fig.2. Each sub-sequence will be transformed and encoded independently using the 3-D SPIHT algorithm.
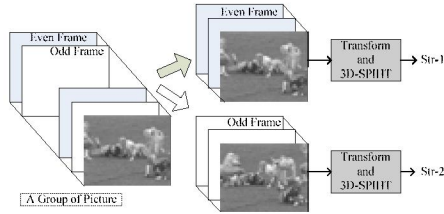

Fig.2. Temporal-domain partitioning. scheme.

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]. We define the concealment ratio of a bit-plane as follows

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

where $MSE_c(j)$ is the MSE of the lower quality bitstream with the missing bit-plane-j had been concealed and $MSE_l(j)$ is the MSE of the low-quality substream after decoding up to the jth bit-plane. Fig.4 shows the performance improvement of the concealment technique, where a GOP contains 32 frames.
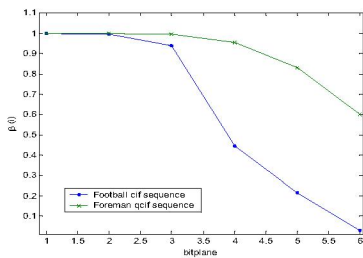

Fig.3. The performance improvement of Football sequence by using our concealment technique.

3. JONIT POST-PORCESSING ERROR PROTECTION FRAMEWORK (JPPEP)

We formulate the receiver post-processing into rate-distortion analysis to solve the optimization problem. For simplicity, we first discuss the single layer case without loss generality and introduce PET concept to deal with multilayer case.

3.1. Formulation of Single Layer Case

By taking advantage of the receiver-end post-processing, we can further consider each substream to be sent or not. For simplicity, we first discuss the single layer and two sub-streams case as shown in Fig.4.
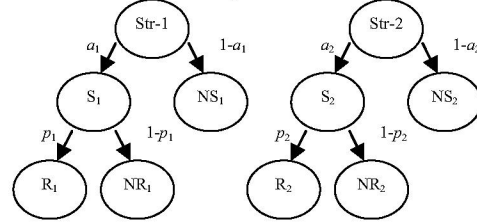

Fig.4. Single layer two substreams case.

The notation Str-s stands for the independent encoded bitstream s, and the state S_s stand for that Str-s will be send to the receiver with transmission probability a_s. Because the performance improvement of the error concealment technique is good enough, we probably have a chance to discard one substream. The state NS_i is that the bitstream Str-s is not sent to the receiver-end with probability $(1 - a_s)$. Once the bitstream is sent, the receiver may not recover the video successfully because the transmission errors has the error probability (1- p_s). In the state R_s, the Str-i can be decoded properly, and in the state NR_s, the error will be too serious to decode the video sequence. Let ΔD_s as the distortion reduction of Str-s to discuss the expected distortion of the proposed simplified transition structure. The expected distortion reduction of Str-1 for two substream case is

$$E[\Delta D_1] = \Delta D_1 (a_1 p_1 + \beta_1 a_2 p_2 ((1-a_1) + a_1 (1-p_1)) \cdot$$

where β_s is the performance improvement parameter conducted by error concealment technique. Therefore, the total expected distortion reduction of all substreams is

$$\Delta \overline{D} = \sum_s E[\Delta D_s]$$

Because each bitstream can be send or not, the total transmission rate will be changed according to the their probability. The total transmission rate is not only affect by the transmission probability, but also by error correction code. We can change the error probability of the transmitted packets by using error correction code, and therefore, the transmission rate will be changed. Using the RS (N, k_s) code as an example, the recover probability p_i will vary with the channel coding ratio (k_s / N). Therefore, the recover probability will dependent on the protection bits assignment. To denote the relationship, let $c_s = N - k_s$ and the recover probability will become $p_s(c_s)$. The approximation of total rate requirement for single layer case becomes

$$E[\Delta r_s] \approx \Delta r_s a_s \frac{N}{N - c_s}$$

where Δr_s is the source rate of substream Str-s. The total expected rate requirement for single layer is

$$\Delta \bar{r} = \sum_s E[\Delta r_s]$$

Our optimization algorithm will select the parameter a_s and c_s which maximize the expected distortion reduction $\Delta \bar{D}$ under the constraint $\Delta \bar{r} \le R$, where R is the rate constraint of the error-resilient video transmission.

3.2. Formulation of MultiLayer Case

We use PET concept to design a multilayer error-resilient 3D coder which include the receiver post-processing in analysis. PET structure ensure that the lower layer data can recovered correctly as decoding higher layers. Therefore, we have $c_s^{(1)} > c_s^{(2)} > ... > c_s^{(L)}$ as shown in Fig.5.
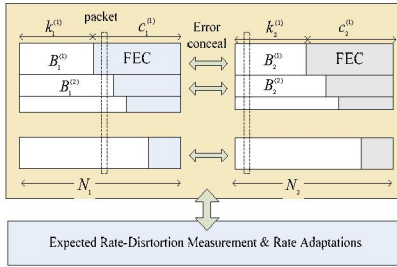

Fig.5. Multi- layer case.

We define $b_s^{(j)}$ as the size of layer-j of bit-stream s in bytes. Since the number of source data bytes of layer-j, $b_s^{(j)}$, is sometimes not divisible by $k_s^{(j)}$, the number of source data bytes protected by an $(N_s, k_s^{(j)})$ code is

$$B_s^{(j)} = \left\lfloor \frac{b_s^{(j-1)} - B_s^{(j-1)} + b_s^{(j)}}{k_s^{(j)}} \right\rfloor \times k_s^{(j)}$$

We also define the aligned source allocation vector as

$$\mathbf{B} = \begin{bmatrix} B_1^{(1)} & B_1^{(2)} & ... & B_1^{(L)} \\ B_2^{(1)} & B_2^{(2)} & ... & B_2^{(L)} \end{bmatrix}$$

As shown in Fig.5, the total number of transmitted packets is $N = N_1 + N_2$ and $c_s^{(j)} = N_s - k_s^{(j)}$. We also define the length that bit-plane in each packet as

$$l_s^l = \frac{B_s^{(j)}}{k_s^{(j)}}$$

The proposed method finds a redundancy assignment vector

$$\mathbf{C} = \begin{bmatrix} c_1^{(1)} & c_1^{(2)} & ... & c_1^{(L)} \\ c_2^{(1)} & c_2^{(2)} & ... & c_2^{(L)} \end{bmatrix}$$

Let us define

$$B_s(j, \mathbf{C}) = \sum_{q=1}^{j} B_s^{(q)}$$

as the accumulated number of source data bytes up to layer-j. $MSE_s(j, B_s(j, \mathbf{C}))$ is the mean square error (MSE) at the receiver relative to the original video sequence when the

receiver decodes source-s $B_s(j, \mathbf{C})$ data bytes, given the redundancy assignment vector $\mathbf{C}$.

$$\Delta D_s^{(j)} = MSE_s(j-1, B(j, \mathbf{C}), N_s) - MSE_s(j, B(j-1, \mathbf{C}), N_s) \quad , s = 1, 2$$

Next, we define $\Delta \bar{D}^{(j)}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ as the amount of expected distortion reduction when the receiver decodes the additional layer-j, given that all layer prior to the j-th layer have already been decoded. We can now calculate the expected distortion reduction at the receiver as

$$\Delta \hat{D}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \sum_{j=1}^{L} \Delta \bar{D}^{(j)}(\mathbf{A}, \mathbf{B}, \mathbf{C})$$

where

$$\mathbf{A} = \begin{bmatrix} a_1^{(1)} & a_1^{(2)} & ... & a_1^{(L)} \\ a_2^{(1)} & a_2^{(2)} & ... & a_2^{(L)} \end{bmatrix}$$

$a_s^{(j)}$ is the sent parameter which decide whether $B_s^{(j)}$ to be sent or not, and the total rate requirement is

$$\Delta \hat{r}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \sum_{j=1}^{L} \Delta \bar{r}^{(j)}(\mathbf{A}, \mathbf{B}, \mathbf{C}) < R$$

R is the total rate constraint. We will search the optimal parameter $\mathbf{A}$ and $\mathbf{C}$. The problem of finding a redundancy assignment vector that maximizes the expected quality under the constraint of total available bytes for a GOP can be expressed as

$$\max_{\mathbf{C}} \Delta \hat{D}(\mathbf{A}, \mathbf{B}, \mathbf{C}) \quad \text{subject to} \quad \Delta \hat{r}(\mathbf{A}, \mathbf{B}, \mathbf{C}) \le R$$

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 [3], a similar problem has been addressed for the joint source-channel coding of images that are progressively transmitted over noisy channels.

We describe an algorithm to find a good FEC assignment vector. Finding the globally optimal assignment of FEC data to each of the streams within the JPPEP framework appears to be computationally prohibitive for useful amount of data. We therefore developed a local search hill-climbing algorithm that makes limited assumptions about the data, but is computationally tractable. As mentioned before, we constrain $c_s^{(j)} \ge c_s^{(j+1)}$ and assume that a single byte missing from the progressive bit stream caused all later bytes to become useless.

We initialize each layer to contain only data bytes, such that $c_s^{(j)} = 0$. In each iteration, our algorithm examines a number of possible assignments equal to $2QL$, where Q is the search distance and L is the number of layers. We determine $\Delta \hat{D}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ after adding or subtracting 1 to Q bytes of FEC data to each layer, while satisfying our constraint $c^{(j)} \ge c^{(j+1)}$. We choose $\mathbf{C}$ corresponding to the highest $\Delta \hat{D}(\mathbf{A}, \mathbf{B}, \mathbf{C})$, update the allocation of FEC data to

all affected layers, and repeat the search process until none of the cases examined improves the expected distortion reduction. Our algorithm finds a local maximum that we believe is quite close to the global maximum and, in some cases, may be identical.

The search distance Q is a parameter of the algorithm that is chosen ahead of time. There is clearly a tradeoff: the larger Q is, the more likely the algorithm will find a global optimum, but the algorithm will require more time to run. When the channel model is well behaved, a small Q seems to yield excellent results.

4. 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 and Akyio sequence at CIF resolution. The channel model is the 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.

Fig.6 shows the performance comparison of proposed method, 2-D dynamic programming, and 1-D independent dynamic programming rate assignment algorithm. The test sequence Akyio is CIF format, and the mean packet loss rate is 5% with burst length 5. The proposed method has better performance than the other methods at higher bit rate, because the proposed method has different number of layers which allow the concealment effect to be significant. Fig.7 shows the performance of JPPEP method with different sent parameter **A**, where JPPEP-A have $a_2^{(1)} = 0$ for the first layer, and JPPEP-B have $a_2^{(j)} = 0$ for j =1,2. Because post-processing allow us sent only one bitstream to increase the EPSNR (Expected Peak Signal to Noise Ratio), the experimental result shows that JPPEP-B have better performance at rate lower then 15k. However, when the bit rate is increased, the transmitted sequence will be protect properly and the concealment technique become useless. Fig.8 shows the performance of difference mean packet loss rate with transmission 20 kbps where the performance of proposed method is similar to 2D-dynamic programming. The performance of the proposed method will become worse when the mean packet loss rate increase. This result is reasonable, because the post-processing may make video quality even worse when both bitstream can not be recover properly.
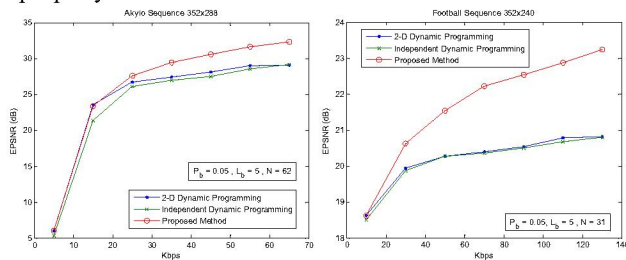

Fig.6. The performance comparison of proposed method, 2-D dynamic programming, and 1-D independent dynamic programming rate assignment.
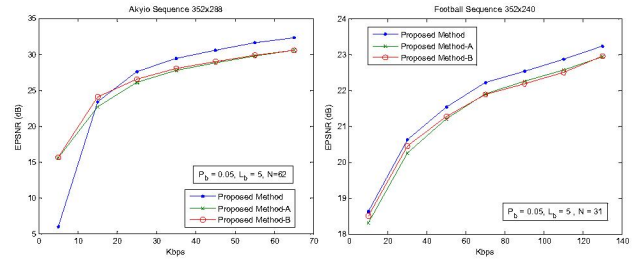

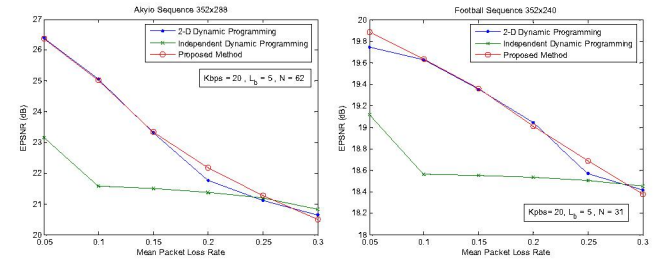Fig.7. The performance comparison of proposed method with different parameter **A**

Fig.8. The performance comparison of proposed method with different mean packet loss rate

5. CONCLUSION

We have presented an joint post-processing and protection framework 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. 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[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.