Adaptive Leaky Factor Determination for Robust Video Transmission over Error-Prone Networks

Jun-Hyung Kim, Min-Cheol Hwang, Tae-Shick Wang, Ju-Hun Nam, Sung-Jea Ko

Department of Electrical Engineering
Korea University
Anam-Dong, Sungbuk-ku, Seoul, Korea
E-mail: sjko@dali.korea.ac.kr

Abstract

When video packets are transmitted over error-prone networks, the leaky prediction can be used to mitigate the effect of error propagation. The leaky factor provides trade-off between coding efficiency and error resilience in the leaky prediction. In this paper, we propose an improved leaky prediction method where the leaky factor is adaptively determined for each frame by minimizing the estimated end-to-end distortion at the encoder. Experimental results show that the proposed method with the adaptive leaky factor shows the better performance of the error robustness as compared with the conventional method.

Keyword: Leaky prediction, error resilient video coding, video transmission, end-to-end distortion

1. Introduction

Standard video coding such as H.264/AVC [1] and MPEG-4 [2] achieves good coding gain through various techniques which reduce temporal and spatial redundancy in video sequences. However, when video signals are transmitted over unreliable channels like UMTS and WiMAX, the bitstream is vulnerable to channel errors because of the reduced redundancy. Even a single bit error can severely degrade the quality of the video frame. Moreover, the error can be propagated to successive frames.

Therefore, various error resilience (ER) techniques have been introduced to minimize the visual degradation caused by the transmission error [3]. Error concealment (EC) [4] [5] is an efficient method which attempts to recover the damaged data based on temporal and spatial correlation at the decoder side. However, in the case of the burst channel error, the EC technique cannot be utilized due to the lack of available information for retrieving erroneous data.

Error-resilient coding (ERC) is a different approach to alleviate the effect of the transmission error. Various redundant data are inserted at the bitstream level [2] or the source level [6] to make the encoded video data more resilient to potential errors. Other methods [7] [8] employ two or more channels with equal or unequal reliability for robust video transmission over error-prone networks. Although, ERC can successfully suppress the deterioration caused by the channel error, it suffers from loss of the coding efficiency and high implementation cost which make ERC unsuitable for many practical applications.

The leaky prediction is an alternative method which prevents the errors from being temporally propagated at the expense of the efficiency of the motion-compensated prediction [9]. In the leaky prediction, the leaky factor provides trade-off between coding efficiency and error resilience. Therefore, it is important to determine an appropriate leaky factor according to the channel condition.

Many studies have been performed to find the optimum leaky factor in the layered video coding [10]. However, most existing schemes use a constant value for the leaky factor in the single-layer video coding. In this paper, we propose an improved leaky prediction method which adaptively adjusts the leaky factor at a frame-level in the single-layer video coding. The proposed method utilizes the end-to-end distortion (EED) model [11], [12]. In general, EED consists of the quantization distortion induced in source encoding, and the error concealment and propagation distortions caused by channel errors. In the proposed method, the best leaky factor for each frame is determined by minimizing the estimated EED at the encoder. In addition, a finite set of candidate leaky factors is used to reduce computational complexity of the calculation of EED.

This paper is organized as follows. In Section 2, we describe the proposed algorithm in detail. Experimental results are given and discussed in Section 3. Finally, Section 4 concludes this paper.

2. Proposed Method

First, we briefly review the EED model and the leaky prediction. Then, the proposed method with the adaptive leaky factor is described in detail.

2.1 End-to-End Distortion Model

In [11], several EED models are introduced for various encoding conditions. For ease of explanation without loss of generality, we assume that the encoder utilizes subpel motion vectors, constrained-intra prediction, and no deblocking filter.

In general, the total distortion D_n for nth frame f_n can be decomposed into the encoder-induced error D_n^e and channel-induced error D_n^c as follows, $D_n = D_n^e + D_n^c$. Since D_n^e is caused by quantization, it can be easily calculated at the encoder. Let P and β_n denote the channel error rate and intra macroblock (MB) rate, respectively. The average channel distortion of f_n is formulated as

$$D_n^c = (1-P)((1-\beta_n)D_n^P + \beta_n D_n^I) + P \cdot D_n^L, \quad (1)$$

where D_n^I, D_n^P, and D_n^L represent the error-propagated distortions in the correctly received intra MBs and inter MBs, and the error concealment (EC) distortion of the corrupted f_n, respectively. The assumption of constrained-intra prediction produces $D_n^I = 0$ and D_n^P given by

$$D_n^P = E\{(f_p^e - f_p^d)^2\} = aD_{n-1}^c, \quad (2)$$

where a is the model parameter, f_p^e and f_p^d are the motion-compensated frames at the encoder and decoder, respectively. D_n^L in (1) is given by

$$D_n^L = D_n^{EC} + hD_{n-1}^c. \quad (3)$$

where D_n^{EC} represents the distortion caused by a specific EC method and h is the model parameter. Substituting (2) and (3) into (1), gives

$$D_n^c = (a(1-\beta_n)(1-P) + hP)D_{n-1}^c + P \cdot D_n^{EC}. \quad (4)$$

2.2 Leaky Prediction Method

In the leaky prediction [9], a new reference frame $\hat{f}_p$ is obtained by modifying the original motion compensated frame f_p as follows:

$$\hat{f}_p = \alpha f_p + (1-\alpha)C, \quad (5)$$

where α and C denote the leaky factor and a constant value, respectively. The leaky factor α ($0 \leq \alpha \leq 1$) controls trade-off between the coding efficiency and the error resilience. As α approaches to 0, channel distortions rapidaly decrease but more bits are needed to encode the current frame [9]. Note that the constant term C in (5) should be equal to $E[f_p]$ for preserving the energy of the reference frame.

2.3 Adaptive Leaky Factor Determination

By adaptively selecting the leaky factor α in (5), we can improve the performance of the leaky prediction method. In the proposed method, the EED model is utilized to determine the optimal leaky factor. The error-propagated distortion in inter MBs, D_n^P in (1) is modified by substituting (5) into (2).

$$\begin{aligned} D_n^p &= E\{[\alpha f_p^e + (1-\alpha)C - \alpha f_p^d - (1-\alpha)C]^2\} \\ &= \alpha^2 E\{(f_p^e - f_p^d)^2\} = \alpha^2 a D_{n-1}^c. \end{aligned} \quad (6)$$

Accordingly, D_n^c is derived as

$$D_n^c = (\alpha^2 a(1-\beta_n)(1-P) + hP)D_{n-1}^c + P \cdot D_n^{EC}. \quad (7)$$

In our simulation, we use the parameter values in [11] for a and h. From (7), we can confirm that D_n^c is reduced as α decreases.

Unlike the conventional method employing a constant α, the proposed method adaptively determines α for each frame. In the proposed method, a set of candidate values between 0 and 1 is selcted for α. For each candidate α_k in the set S, the number of coded bits $R_n^{\alpha_k}$ and total distortion $D_n^{\alpha_k}$ of f_n are calculated. Then, we evaluate the rate-distortion slope as follows

$$-\frac{D_n^{\alpha_k} - D_n^1}{R_n^{\alpha_k} - R_n^1} \geq T, \quad (8)$$

where R_n^1 and D_n^1 are the number of coded bits and total distortion, respectively, when $\alpha = 1$. T is a threshold controlling the bitrates and error resilience. Among the candidates which satisfy the inequality, the smallest one is chosen as the best α. If no candidate satisfies the inequality, α is set to 1. In the proposed method, α can be adjusted frame by frame.

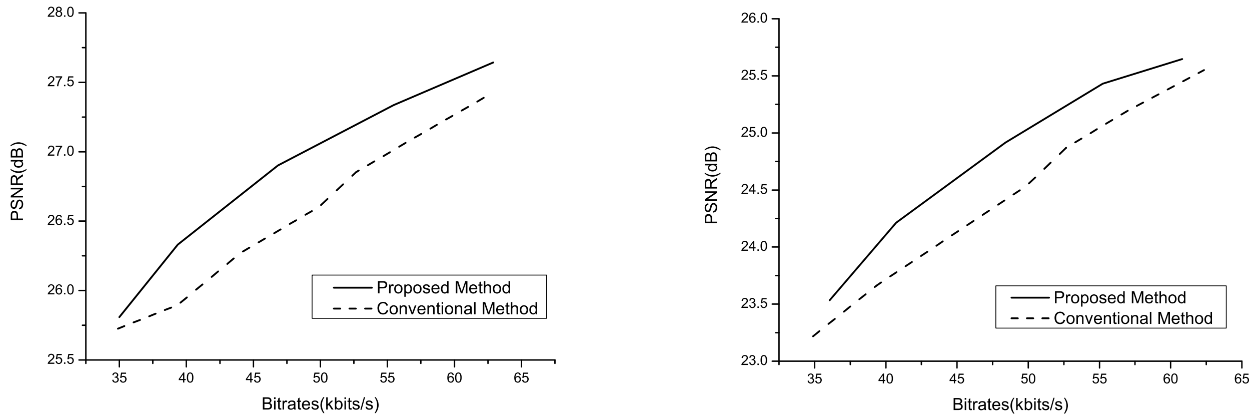
Using the weighted prediction tool in H.264/AVC [13], α and C can be incorporated into the coded bitstream. Therefore, the proposed method is standard compatible when α and C are stored as a weighting factor and an offset, respectively.

3. Simulation Results

To evaluate the proposed algorithm, we use JM12.2 reference software [14] and two test sequences, 'Foreman' and 'Football' with QCIF(176 × 144) resolution and 15Hz frame-rate. Each sequence consists of 60 frames with YUV 4:2:0 format. Only the first frame is an I-frame and remaining frames are P-frames, i.e., a GOP structure is IPPP. In our experiment, QP is set to 36 and frame copy supported by JM12.2 is selected for an EC method. We generate 100 error patterns with random frame loss of 5% and 10% which are known as the representative error rates of the wireless environment [11], [12]. The candidate sets for the leaky factor are set to {121/128, 122/128, ..., 127/128} and {25/32, 26/32, ..., 31/32} for 'Foreman' and 'Football' sequences, respectively.

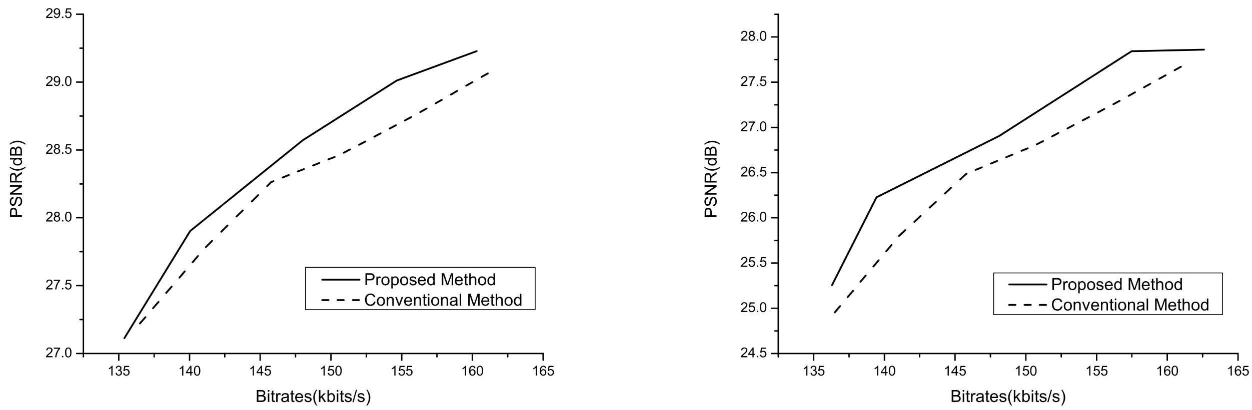
Figs. 1 and 2 show the comparison of the PSNR performance for various bitrates and channel error rates. Average PSNR value is calculated over all of the error patterns and sequences. In the conventional method, the same leaky factor is applied to all frames. To obtain various output bitrates, the conventional method changes the leaky factor from the largest to the smallest candidate in S one by one. In the proposed method, different thresholds are utilized to control the output bitrates and the adaptive α for each frame is determined using (8) with the predefined T. Figs. 1 and 2 show that the proposed method outperforms conventional method in the rate-distortion sense.

The proposed method exhibits not only higher PSNR performance but also better subjective visual quality than the conventional methods. Fig. 3 illustrates the comparison result of the subjective visual quality of the conventional and proposed methods on the Foreman se-



(a) (b)

Figure 1. Comparison of the PSNR performance for various bitrates and channel error rates on Foreman. (a) Error rate = 5%. (b) Error rate = 10%.



(a) (b)

Figure 2. Comparison of the PSNR performance for various bitrates and channel error rates on Football. (a) Error rate = 5%. (b) Error rate = 10%.

quence. Since the transmission error cannot be perfectly concealed by the frame copy method, the frame following the corrupted and recovered frame is degraded by the error-propagation. In the conventional method, the annoying artifacts can be seen around boundaries of a face, a nose, and eyes of a foreman as shown in Fig. 3(b). However, the proposed method significantly reduces the deterioration caused by the error-propagation as shown in Fig. 3(c).

4. Conclusion

In this paper, we have proposed an adaptive leaky factor determination method which improves the performance of the leaky prediction. The EED model is utilized to adaptively determine the leaky factor for each frame. Simulation results have shown that the proposed method outperforms the conventional method in the rate-distortion sense.

Acknowledgement

This research was supported by Seoul Future Contents Convergence (SFCC) Cluster established by Seoul R&D Program.

References

[1] T. Wiegand, G. J. Sullivan, G. Bjϕntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, July 2003.

[2] *Information technology – Coding of audio-visual objects – Part 2: Visual*, ISO/IEC 14496-2, 1999.

[3] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," Proc. IEEE, vol. 86, no. 5, pp. 974–997, May 1998.

(a) (b) (c)

Figure 3. Comparison of the subjective visual quality on Foreman with 5% error rate. (a) No error. (b) Conventional method. (c) Proposed method.

[4] M.-C. Hwang, J.-H. Kim, D. T. Duong, and S.-J. Ko, "Hybrid temporal error concealment methods for block-based compressed video transmission," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 198–207, June 2008.

[5] M.-C. Hwang, J.-H. Kim, H.-Y. Yang, S.-J. Ko, and A. W. Morales, "Frame error concealment technique using adaptive inter-mode estimation for H.264/AVC," *IEEE Consumer Electron.*, vol. 54, no. 1, pp. 163–170, Feb. 2008.

[6] J.-S. Lee, J.-W. Kim, K.-H. Lee, and S.-J. Ko, "Residual motion coding method supporting error concealment in low bitrate video coding," *Electron. Lett.*, vol. 41, no. 7, pp. 406–408, Mar. 2005.

[7] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[8] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proc. IEEE*, vol. 93, no. 1, pp. 57–70, Jan. 2005.

[9] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by ATM cell loss," Proc. IEEE ICASSP, San Francisco, CA, vol. 3, pp. 545–548, Mar. 1992.

[10] Z. Li and E. J. Delp, "Channel-aware rate-distortion optimized leaky motion prediction," Proc. IEEE ICIP, Singapore, pp.801–804, Oct. 2004.

[11] Y. Wang, Z. Wu, and J. M. Boyce, "Modeling of transmission-loss-induced distortion in decoded video" IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 6, pp. 716–732, June 2006.

[12] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 6, pp. 511–523, June 2002.

[13] J. M. Boyce, "Weighted prediction in the H.264/MPEG AVC video coding standard," Proc. IEEE ISCAS, Vancouver, vol. 3, pp. 789–792, May 2004.

[14] K. Suhring, Ed., JM 12.2 Reference Software. [Online] Available at: http://iphome.hhi.de/suehring/tml/download/