Adaptive Lossy Error Protection Architecture in H.264 Video Transmission

Zong-Yi Chen, Tien-Hsu Lee, and Pao-Chi Chang

Department of Communication Engineering National Central University, Jhongli, Taoyuan 32001, Taiwan E-mail: {zychen, thlee, pcchang}@ vaplab.ce.ncu.edu.tw

Abstract—Systematic Lossy Error Protection (SLEP) is a robust error resilient mechanism which uses Wyner-Ziv coding to protect the video bitstream. In this paper, we propose a low overhead adaptive lossy error protection (ALEP) mechanism that provides a good trade-off between the error resilience and decoded video quality. The proposed method can generate appropriate redundant slices to provide proper error correction capability for varying channel conditions.

The proposed method maintains good video quality at low packet loss rate compared to original SLEP and still provides sufficient error correction capability at high packet loss rate in our simulation results. It achieves 2-3 dB PSNR improvement at 5% packet loss rate for various video sequences in our simulations.

I. INTRODUCTION

H.264/AVC [1] has become the spotlight of all video coding standards. It provides significant improvement of coding efficiency in comparison to all past standards. H.264/AVC has attracted a lot of attention from industry and has been adopted by a variety of video applications. It is expected that H.264/AVC will be commonly used for video transmission in the future. When the video is delivered over packet-switched networks, there inevitably exists packet losses and leads to incorrect decoded video. Because H.264/AVC greatly utilizes the relationship in spatial and temporal domains to achieve the high coding efficiency, it generates more serious error propagation than past standards when packets are lost. Thus the video quality will degrade drastically if no error protection is applied.

In order to correct video transmission errors, forward error correction (FEC) code is generally applied to protect the source bitstream. The FEC scheme applies on source bitstream leads to the penalty of bit rate overhead. Besides, when the channel error rate exceeds the error correction capability of FEC code, the video quality degrades rapidly thus leads to the so-called "cliff" effect. Another common way to recover the lost portions of the video is to encode redundant slices and transmit them. However, transmitting these redundant slices is also rate consuming. Some conventional approaches [2], [3] employ layered video coding schemes combined with unequal error protection can achieve graceful degradation of the video quality in the presence of channel errors. However, most layered video coding schemes in past standards were not used in practice because of the substantial rate-distortion penalty relative to the single layer schemes. Although the H.264/SVC had been standardized in 2007 [4] and claimed to hold excellent rate-distortion performance, the error protection of single layer bitstream is still a vital issue. To overcome this problem, Systematic Lossy Error Protection (SLEP) architecture was proposed [5]-[7]. SLEP uses Wyner-Ziv coding [8][9] to protect the bitstream can also achieve graceful degradation of the decoded video quality without the requirement of layered coding.

In this paper we analyze the trade-off of SLEP and propose a simple and practical method to improve the SLEP scheme. The proposed adaptive lossy error protection (ALEP) can adjust the quality of redundant slices according to the network condition, so as to provide different degrees of error protection to protect videos.

The rest of the paper is organized as follows. We review the SLEP architecture and state the trade-off problem of SLEP in Section II. In Section III we present the proposed ALEP method. Experimental results are given in Section IV, and we conclude this work in Section V.

II. REVIEW AND ANALYSIS OF SYSTEMATIC LOSSY ERROR PROTECTION

In this section, we first describe the principle of SLEP briefly, then analyze the performance from simulation results.

A. Principle of SLEP

The block diagram of SLEP is shown in Fig. 1. Input videos are encoded with H.264/AVC encoder to generate the so-called primary slices. Then the primary slices are transmitted without any protection. The redundant slices are encoded with the same coding information, including motion vectors and block modes, as the corresponding primary slices but with larger quantization step size, i.e., larger QP. Using the same coding information in redundant slices prevents the time consuming motion estimation processes from being executed again. Although this strategy results in non-optimized rate-distortion performance in redundant slices encoding, the low complexity process makes this scheme practicable in real applications.

SLEP then applies Reed-Solomon (RS) codes across the redundant slices but only the generating parity bits are transmitted. If any slices are detected to be lost at the



Conventional H.264/AVC Codec and Transmission

Fig. 1. SLEP block diagram.

H.264/AVC decoder, the redundant slices with the same lost portions as the corresponding primary slices are generated. The erroneous redundant slices are combined with the received parity slices and Reed-Solomon decoding is performed. Finally, the lost parts of primary slices are replaced by the corrected redundant slices. For low motion videos such as "Akiyo" sequence, we just need to encode the foreground, i.e. often means the regions with moving objects, to generate the redundant slices. The portions that do not exist moving objects, e.g. the background, do not need protection because decoder-based error concealment could reconstruct them with even better quality than SLEP can provide. The saved redundant bit rate can provide the foreground more protection.

By using this method, SLEP can provide more protection for redundant slices than FEC codes applied directly to primary slices under the constraint that the total transmission rate is fixed. Thus SLEP can provide graceful degradation of video quality in the presence of increasing packet loss rate.

B. Analysis of SLEP

We first show some experimental results for discussion. We use the JM 11.0 H.264 reference software in our simulations. MBLineIntraUpdate is enabled. Three kinds of QCIF video sequences belong to Class A, Class B, and Class C (Akiyo, Foreman, and Stefan) respectively are used for test. In our simulations, one slice with 11 macroblocks are packetized as one packet, and we simulate video transmission with random packet losses. For the slices that can not be recovered by FEC, the JM 11.0 built-in error concealment mechanism will conceal the lost portions in all cases. For each sequence the parity bit rate is fixed to stand for constrained transmission rates. Fig. 2 shows the experimental results for Foreman video sequence. The curve of EC indicates only JM 11.0 built-in error concealment is performed when packets are lost. FEC case indicates that we use FEC to protect the primary bitstream directly. SLEP_QPxx represents the use of SLEP to protect the video and the redundant slices are encoded with QP = xx. Table I shows the RS codes and the generated parity and redundant slices bit rates of Foreman. Other two videos have similar results.

The total transmission rate of this structure is the sum of primary bit rate and parity bit rate. For fair comparison, the video bit rate of EC case will be a little higher than others because no parity is needed. From Fig. 2 and Table I, we can observe that the FEC case indeed suffers from the cliff effect, and the error resilience of SLEP increases when the redundant slices are coarser. We also notice that the coarser the QP is, the smaller the redundant bit rate is, thus under the same constrained parity bit rate, we can provide stronger RS code to protect the redundant slices.



Fig. 2. Foreman@1205kbps (150kbps parity)

Foreman	Primary_QP	Primary bit rate (kbps)	Parity bit rate (kbps)	Redundant bit rate (kbps)	RS(n,k)
EC	Rate Control Enabled	1205.5	х	х	х
FEC	16	1055.5	150	х	RS(255,223)
SLEP_QP24	16	1055.5	150	387.31	RS(255,183)
SLEP_QP32	16	1055.5	150	151.29	RS(255,127)

TABLE I RS CODES USED BY FOREMAN

From the observation of experimental results, SLEP exists a trade-off between error robustness and overall coding efficiency. The higher error resilience, the worse the redundant slices quality, and the lost portions will be replaced with coarser versions. The situation is obvious when the loss rate becomes lower. At a specific packet loss rate that does not exceed the FEC correction capability in all SLEP cases, we certainly want to use the better redundant slices for protection. Thus, we have to adaptively adjust the encoding of redundant slices to achieve a better trade-off between the error robustness and final decoded video quality.

III. PROPOSED ADAPTIVE LOSSY ERROR PROTECTION

Most unequal error protection (UEP) schemes are designed on the basis of knowledge of importance of the multimedia data (known a-priori) [10]. Some methods that consider error propagation and try to optimize the rates or modes of primary and redundant slices have been proposed, too [7][11]. These approaches could result in optimized rate-distortion performances. However, these approaches have to estimate the average per-pixel end-to-end distortion in advance. The requirement of an a-priori study is time consuming and increases substantial computational complexity. An easier method with little rate-distortion loss may be desirable. In this work, we propose a method from the statistical results that introduces very low computational overhead but maintains good rate-distortion performance.

A Reed-Solomon code over GF(2⁸) across redundant slices is applied in this work. An 8-bit Reed-Solomon code with 2*t* parity symbols has the capability of correcting up to *t* symbol errors within a code word of up to 255 symbols. For a RS(*n*, *k*) code, the parity length is 2t = n - k. In our simulations the parity length is fixed, so we can determine *k* for RS(*n*, *k*) according to the length of redundant slice as (1).

$$\frac{k}{n-k} = \frac{Redundant \ bit \ rate}{Parity \ bit \ rate} \tag{1}$$

During the SLEP encoding procedure, the RS encoding is followed by redundant slice encoding, hence we can use (1) to determine the RS code. However, the exact bit rate of the redundant slices for the selected RS code to recover all errors is not trivial to determine. From our experimental results, we

TABLE II PACKET LOSS RATE VS. REQUIREMENT OF REDUNDANT SLICES QUALITY FOR ERROR-FREE

Packet loss rate (PLR)	Foreman redundant QP (Primary QP=16)	Akiyo redundant QP (Primary QP=24)	Stefan redundant QP (Primary QP=16)
3%	QP20(Pri+4)	QP28(Pri+4)	QP20(Pri+4)
5%	QP20(Pri+4)	QP28(Pri+4)	QP24(Pri+8)
10%	QP24(Pri+8)	QP32(Pri+8)	QP28(Pri+12)
12.50%	QP28(Pri+12)	QP36(Pri+12)	QP32(Pri+16)
16.67%	QP28(Pri+12)	QP36(Pri+12)	QP32(Pri+16)

can get the approximate relations between the packet loss rate (PLR) and the QP used to encode redundant slices so that the selected RS code can correct almost all errors, as shown in Table II. For example, when Foreman primary slices are encoded with QP=16, channel packet loss rate is 3%, if the redundant slices are encoded with QP=20, then the selected FEC code can correct all errors.

We assume the decoder can provide the network or channel condition for encoder by any means, then we can adaptively choose the QP of redundant slices to be suitable for the current packet loss rate at the encoder. Based on the heuristic results and to achieve our original objective, the proposed algorithm works as the following steps:

Step a. Encoder gets the packet loss rate from the decoder.

Step b. Choose proper QP to encode redundant slices by (2), where QP_{red} is the QP of redundant slices and QP_{pri} is the QP of primary slices.

$$QP_{red} = QP_{pri} + \lfloor 100 \times \text{PLR} \rfloor$$
(2)

Step c. After redundant slices are encoded, use (1) to determine the proper RS codes.

At *step b*, coarser redundant slices will be encoded against higher PLR. We can observe that the proper QPs of redundant slices for different type of video sequences are also different from Table II. We use the same criterion to choose the QP of redundant slices for simplicity, although this may not be the best way in some situations, it just has little effect to the decoded video quality as shown in Section IV. This method is practical because we do not have to know what type of video is transmitted in advance and it costs low computation.

IV. EXPERIMENTAL RESULTS

Fig. 3 shows the experimental result for Stefan sequence. The experimental environment setting is the same as in Section II. The curve ALEP represents the proposed scheme. Fig. 4 and Fig. 5 show the experimental results for "Carphone" and "Coastguard" sequences respectively those are not used in the algorithm development phase to demonstrate the usability of the proposed method. Fig. 6



Fig. 3. Stefan@2550kbps (320kbps parity) with ALEP



Fig. 4. Carphone@1015kbps (127kbps parity) with ALEP



Fig. 5. Coastguard@1491kbps (187kbps parity) with ALEP



Fig. 6. Video quality variation of Foreman

further shows the quality variation of Foreman sequence under increasing packet loss rate environment. We can observe that ALEP provides more stable video quality than others. From the simulation results we can find the proposed ALEP just suffers a little quality reduction compared with other cases at certain packet loss rates. However the proposed method clearly outperforms other cases on average. ALEP not only provides high error resilience but also maintains high decoded video quality at low packet loss rate.

V. CONCLUSIONS

This work proposes an efficient adaptive lossy error protection method which can adaptively adjust the generation of redundant slices in SLEP. The proposed ALEP makes the video quality gracefully degraded when the network condition becomes worse and worse. Besides, it does not have penalty at low packet loss rate as original SLEP. ALEP makes a good trade-off between video transmission robustness and decoded video quality. It only needs very simple calculations and we do not have to know what kind of video is transmitted in advance.

In the future, we intend to derive the exact QP of redundant slices encoding at any packet loss rates. We should also concern the network delay when the channel condition information reported from the decoder to the encoder.

REFERENCES

- ISO/IEC ITU-T Rec. H264: Advanced Video Coding for Generic Audiovisual Services, Joint Video Team (JVT) of ISO-IEC MPEG & ITU-T VCEG, Int. Standard, May 2003.
- [2] M. Gallant and F. Kossentini, "Rate-distortion optimized layered coding with unequal error protection for robust internet video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 357-372, Mar. 2001.
- [3] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 819-828, Jun. 2000.
- [4] 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.
- [5] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93,no. 1, pp. 71-83, Jan. 2005.
- [6] M. Ramon, F. X. Coudoux, and M. G. Gazalet, "Efficient priority encoding transmission scheme based Wyner-Ziv coding and spatial scalability," in *Proc. of SPIE*, July. 2005, vol. 5960, pp. 1903-1910.
- [7] P. Baccichet, S. Rane, A. Chimienti, and B. Girod, "Robust Low-Delay Video Transmission using H.264/AVC Redundant Slices and Flexible Macroblock Ordering," in *IEEE Int. Conf. Image Process.*, 2007, vol. 4, pp. 93-96.
- [8] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 471–480, Jul. 1973.
- [9] A. Wyner and J. Ziv, "The rate-distortion function for source coding wilh side information at the decoder," *IEEE Trans. Inf. Theory*, vol. IT-22. no. 1, pp. 1-10, Jan. 1976.
- [10] V. K. M. Vadakital, M. M. Hannuksela, M. Rezaei, and M. Gabbouj, "Method for Unequal Error Protection in DVB-H FOR Mobile Television," in *IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2006.
- [11] J. C. Schmidt and K. Rose, "Jointly Optimized mode decisions in redundant video streaming," in *IEEE ICASSP*, Apr. 2009, pp. 797-800.