A Significant Motion Vector Protection-based Error-Resilient Scheme in H.264

Jan-Ru Chen[1,2], Chun-Shien Lu[2], and Kuo-Chin Fan[1]

[1]Department of CSIE, National Central University, Chung-Li, Taiwan, ROC

[2]Institute of Information Science, Academia Sinica, Taipei, Taiwan, ROC

Abstract— **This paper proposes a significant motion vector protection (SMVP) scheme for error-resilient transmission of videos. In terms of a rate-distortion optimization model, we show how to determine the significant motion vectors (SMVs) and how much rate space should be preserved to store SMVs. The idea behind our method is to give more protection to significant data. As a result, our method can be regarded as a kind of unequal error protection mechanisms. In comparison with the conventional forward error correction (FEC) and error concealment methods, experimental results demonstrate the effectiveness of the proposed method. In particularly, our method shows its superiority under the situation of higher packet loss rates.**

I. Introduction

With the advancement of the Internet and multimedia technologies, transmission of compressed media data in an error prone environment becomes frequent. Due to the inherent packet loss property of the Internet, robust transmission of video via error control mechanism has attracted much attention for a long time. In the literature, a number of error control schemes have been proposed [1], [2], [5], [6], [7], [8], [9] for robust video transmission. Wang and Lin [8] provided a MDC approach, which can stop errors from propagating into subsequent frames. Puri *et al.* [3] proposed a FEC-based multiple description coding (MDC) scheme for layer coding that can recover the first important k layers when any k out of n ($k \leq n$) descriptions can be successfully received.

Data hiding technologies have also been employed for correcting errors. In [2], Lu proposed a fragile watermarking method for error detection and recovery of image transmission. Shanableh and Ghanbari [6] proposed an error concealment scheme, which can be regarded as a data hiding-like method. Since motion vectors play a crucial role in the video decoding process, they exploited the inherent B-pictures property in that the concealment motion vectors can be restored if it is forced to be derived from the relationship between forward motion vector and backward motion vector. It was efficient for error concealment at the decoder end at the expense of sacrificing the video coding efficiency. In [7], Song and Liu proposed a MV protection scheme by embedding the MV parity bits of the current frame into its subsequent frame. The main drawback is that the number of lost slices within a frame was restricted to 1. Yin *et al.* [9] improved Song and Liu's work by embedding frame-wise parity bits that were generated from motion vectors for resisting burst errors.

Corresponding Author: C. S. Lu (email: lcs@iis.sinica.edu.tw).

Recently, side information concept [1], [5] has been addressed for robust video transmission. Sehgal *et al.* [5] proposed the idea of peg frame in H.264, which was used to play the role of a reference frame in order to prohibit from error propagation. However, video frames were restricted to refer peg frame only in motion estimation, thus the coding efficiency would be sacrificed. Aaron *et al.* [1] proposed frame hash to help decoder in estimating motion vectors.

On the other hand, unequal error protection (UEP) is usually adopted for video transmission in that significant data is given more protections than insignificant data. In this situation, different MVs should be endowed with different degrees of importance in order to give proper protection. In this paper, we propose a significant motion vector protection-based error-resilient scheme for H.264. We investigate how to identify significant motion vectors and how they can be determined from a rate-distortion model.

The remainder of this paper is organized as follows. In Sec. II, we discuss how to generally determine important side information according to a distortion model. In Sec. III, the proposed method based on using motion vectors as side infromation is described. Experimental results are given in Sec. IV and conlusions are drawn in Sec. V, respectively.

II. Distortion-based Significant Side Information Determination

In this section, we discuss where SMVs can be extracted according to distortion measurement. More specifically, SMVs are extracted from those macroblocks (MBs) decoded (using side information) with larger distortions. The side information may be motion vector, residual errors, and so on that are used at decoder. In this paper, if a corrupted MB is corrected using error concealment only, then it is said that no side information is used.

Let $D(i)$ be the distortion of ith MB decoded without using side information, and let $D^s(i)$ be the distortion of ith MB decoded using side information. Distortion here means the mean-square-error (MSE) measured between the corrected MB and its original version. Let $Y(i)$ be the difference between $D(i)$ and $D^s(i)$ as

$$Y(i) = |D(i) - D^s(i)|. \qquad (1)$$

$Y(i)$ is employed to determine what MBs can be efficiently corrected using side information.

Eq. (1) can be interpreted in two ways. First, when $Y(i)$ is small, it means that corrupted macroblock could be properly

corrected using default error concealment technique only and error correction using side information would not be helpful. In this case, the side information associated with this kind of MBs is defined to be insignificant and are not further stored in our method. Second, when $Y(i)$ is large, it means that corrupted macroblock cannot be properly corrected using default error concealment technique only and error correction using side information is more helpful. In this case, the side information associated with this kind of MBs is defined to be significant and are further stored in our method for error correction.

So far, we have described what side information should be stored for further error correction. How many significant motion vectors need to be stored will be discussed in the next section based on a rate-distortion model.

III. PROPOSED METHOD

As mentioned in the previous section, "side information" is helpful in correcting transmission errors and in identifying significant motion vectors. In this section, we will adopt motion vectors, which are available during encoding, as the side information. In addition, two major issues will be addressed. First, we shall define what are significant motion vectors (SMVs). Second, we have to determine how much rate space should be preserved for SMVs.

A. Significant Motion Vectors

Suppose there are, in total, N macroblocks in a video frame. As mentioned in Eq. (1), the SMVs are defined based on the error function $E(.)$ as follows:

$$E(i) = |P_{MC}(i) - P_{CE}(i)|, \ \forall \ 1 \le i \le N, \qquad (2)$$

where $P_{MC}(i)$ is the PSNR value of the macroblock i (denoted as MB_i) decoded using motion vector only and $P_{CE}(i)$ is the PSNR value of MB_i obtained using error concealment. Based on $E(i)$'s, the significance of a motion vector is determined according to the following rules:

- If $E(i)$ is small, it means that error concealment is feasible to decode MB_i. The MV associated with MB_i is determined to be insignificant.
- If the $E(i)$ is large, it represents the MV of MB_i can not be predicted accurately. Thus, this kind of MVs is determined to be significant.

According to Eq. (2), all $E(j)$'s can be sorted in determining the order of the significance of MVs. Let $\Pi = (\pi_1, \pi_2, ..., \pi_N)$ be a policy vector for all macroblocks in a frame, where π_i is the policy of the ith MV in determining it importance. If the ith MV is determined to be significant, we set $\pi_i = 1$; otherwise we set $\pi_i = 0$.

B. Rate-Distortion Model

How many MVs should be determined to be significant is based on a rate-distortion model. First, we consider three cases of distortions that may be caused by decoding macroblocks.

- If the ith macroblock, MB_i, can be decoded, the distortion, $P_e^q(i, Q_p)$, of the macroblock is purely caused by quantization.
- If MB_i can not be decoded and its corresponding SMV exists, the distortion, $P_e^r(i, Q_p)$, of the macroblock is the residual error.
- If MB_i can not be decoded and its corresponding SMV does not exist, the distortion, $P_e^c(i)$, of the macroblock is caused by error concealment at decoder.

Given a macroblock loss rate p_l and a macroblock policy π_i, the quality obtained by decoding MB_i can be statistically defined as follows:

$$
\begin{aligned}
P(i, Q_p) \ = \ & (1 - p_l) \cdot P_e^q(i, Q_p) \\
& + \ p_l \cdot [\pi_i \cdot P_e^r(i, Q_p) + (1 - \pi_i) \cdot P_e^c(i)], \ (3)
\end{aligned}
$$

where Q_p is a quantization parameter. In addition, the frame quality $J(\Pi, Q_p)$ is defined to be the concatenation of $P(i, Q_p)$'s, i.e., $J(\Pi, Q_p) = \oplus_i P(i, Q_p)$.

For a given video frame, its objective quality, $J(\Pi, Q_p)$, depends on the quality parameter Q_p and the policy vector Π. In order to store the significant motion vectors that are defined in the previous section, the original video data rate needs to be reduced. Then, the reduced bit rate is used to store significant motion vectors. How many space should be reduced depends on the quantization parameter Q_p, as defined in Eq. (3). Once Q_p is determined, the number of significant motion vectors can be determined as well. Suppose R^* is the target frame bit-rate, and $R(\Pi, Q_p)$ is the total bit-rate that is the sum of the frame bit-rate obtained by encoding with Q_p, together with the bit-rate obtained by encoding SMVs. Thus, the policy vector Π could be determined by maximizing $J(\Pi, Q_p)$ subjected to the bandwidth limitation $[R(\Pi, Q_p) \le R^*]$.

For a specific Q_p, a policy vector Π can be generated from $J(\Pi, Q_p)$. Therefore, the optimal policy vector Π can be found by selecting the maximal $J(\Pi, Q_p)$ from among all Q_p's.

C. Embedding and Detection of SMVs

H.264 provides multiple-reference frame prediction for selecting the reference frame. In general, the time distance between the current frame and its reference frame is usually small. Due to burst errors, both the current frame and its reference frame may be simultaneously corrupted. In order to reduce the probability that both SMVs and MVs are corrupted simultaneously, we choose the least referenced frame to embed SMVs. To simplify implementation, we embed SMVs directly into the video bitstream using the same strategy of MV encoding.

The SMVs extraction process is basically the inverse operation of the MV decoding process. When a specific frame is decoded, its SMVs, in fact, has already been extracted in the previous frames. When the SMVs are available, the decoder can simply copy the corresponding MBs indexed by SMVs to recovery the corrupted MBs. On the contrary, if SMVs are not

available, the decoder simply uses default error concealment algorithm to deal with the corrupted MBs.

IV. EXPERIMENTAL RESULTS

Two commonly used videos, "Mother&Daughter" and "Foreman", were adopted in our experiments. The bit-rate of "Mother&Daughter" was 66.71 kbit/s at 30 frames/sec and the bit-rate of "Foreman" was 150.7 kbit/s at 30 frames/sec. The GOP structure was "IPP...P" and our method was developed based on the H.264 codec [10]. Since a slice is defined as a row of macroblocks in a frame and each slice is packed as a packet, therefore, packet loss rate is equal to macroblock lost rate p_l, as previously shown in Eq. (3).

In our method, SMVs are directly stored into a bitstream. Thus, it can be regarded as a FEC-based unequal error protection mechanism. In order to fairly compare the performace of our method with the existing methods, we adopted two widely used methods. One is the inherent error concealment mechanism adopted in H.264. The other one is a FEC-based approach, which is EREC [4] + Reed-Solomon (RS) code.

Fig. 1 plots the PSNR curves of "Mother&Daughter" averaged from 10 runs when packet loss rate was 0.1. It can be observed that SMVP method shows the best result and FEC-based method shows the worst result. The reason is that "Mother&Daughter" is a low bit-rate video that is very sensitive to video rate variations. Fig. 2 plots the PSNR curves of "Foreman" averaged from 10 runs when packet loss rate was 0.1. It can be observed that the result obtained using SMVP method is lower than that obtained using FEC with RS(15,9,3) up to 2 dB. However, it should be noted that it is impossible to determine in advance what kind of FEC parameters need to be used. As a result, the merit of FEC-based error control method is not always guaranteed.

Under the circumstance of higher packet loss rates, our SMVP scheme reveals its superior ability. Fig. 3 plots the PSNR curves of "Foreman" averaged from 10 runs when packet loss rate was 0.4. It can be observed that even thought the packet loss rate is as high as 0.4, our method shows the best visual quality. The PSNR values of the SMVP scheme is higher than FEC-based scheme up to 2dB. The reason is that the important motion vectors have been properly protected.

V. CONCLUSION

In this paper, we have presented a significant motion vector protection-based error-resilient scheme in H.264. The so-called SMVs are determined based on a rate-distortion model. The idea behind our method is that UEP-like scheme should be considered from decoder's perspective by exploiting the side information at the encoder side. In addition, we also describe that what kind of MBs need the sidef-information to correct the errors. Experimental results show that SMVs offer promising ability in recovering packet loss errors. Future work will focus on the protection of residual errors since we have observed that it is still not enough to merely protect motion vectors.
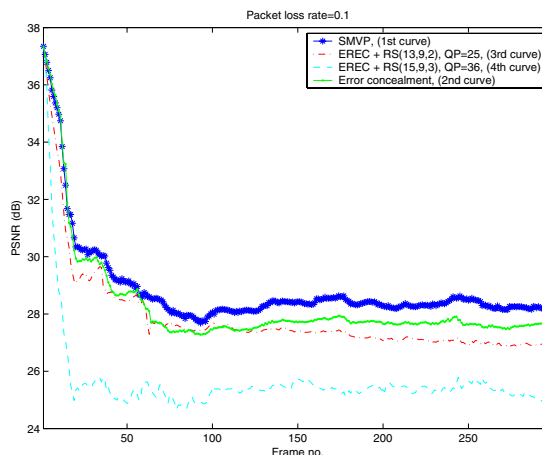

Fig. 1. Comparisons between different error control schemes for the Mother&Daughter video with packet loss rate 0.1.
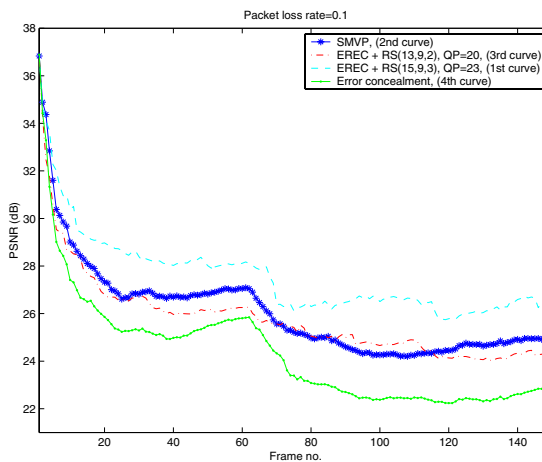

Fig. 2. Comparisons between different error control schemes for the Foreman video with packet loss rate 0.1.
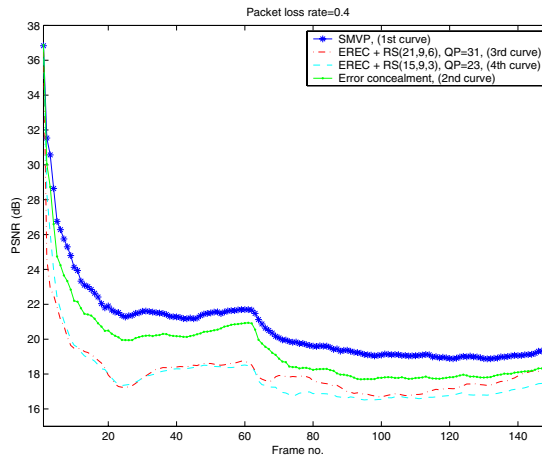

Fig. 3. Comparisons between different error control schemes for the Foreman video with packet loss rate 0.4.

REFERENCES

[1] A. Aaron, S. Rane and B. Girod, "Wyner-Ziv Video Coding with Hash-based Motion compensation at the Receiver," *Proc. IEEE Int. Conf. on Image Processing*, 2004.

[2] C. S. Lu, "Wireless Multimedia Error Resilience via A Data Hiding Technique," *Proc. 5th IEEE Int. Workshop on Multimedia Signal Processing*, US Virgin Islands, USA, Dec. 9-11, 2002.

[3] R. Puri, K. Ramchandran, K. W. Lee and V. Bharghavan, "Forward error correction (FEC) codes based multiple description coding for internet video streaming and multicast," *Signal Processing: Image Communication*, Vol. 16, pp. 745-762, May 2001.

[4] D. W. Redmill and N. G. Kingsbury, "The EREC: an eerror-resilient technique for coding variable-length blocks of data," *IEEE Trans. on Image Processing*, Vol. 5, No. 4, Apr. 1996.

[5] A. Sehgal, A. Jagmohan and N. Ahuja, "Wyner-Ziv Coding of Video: An Error-Resilient Compression Framework," *IEEE Trans. on Multimedia*, Vol. 6, No. 2, Apr. 2004.

[6] T. Shanableh and M. Ghanbari, "Loss Concealment Using B-Pictures Motion Information," *IEEE Trans. on Multimedia*, Vol. 5, No. 2, 2003.

[7] J. Song and K. J. R. Liu, "A Data Embedded Video Coding Scheme for Error-Prone Channels," *IEEE Trans. on Multimedia*, Vol. 3, No. 4, 2001.

[8] Y. Wang and S. Lin, "Error-Resilient Video Coding Using Multiple Description Motion Compensation," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 12, No. 6, Jun. 2002.

[9] P. Yin, M. Wu and B. Liu, "A Robust Error Resilient Approach for MPEG Video Transmission over Internet," *Proc. of Inter. Conference on Visual Comm. and Image Processing (VCIP'02)*, San Jose, CA, Jan. 2002.

[10] *http://bs.hhi.de/ suehring/tml/download/jm20.zip*.