2012 Fourth International Conference on Multimedia Information Networking and Security

Adaptive Video Error Concealment Using Reversible Data Hiding

Yuanzhi Yao[1], Weiming Zhang[1,2], and Nenghai Yu[1]
[1]Department of Electronic Engineering and Information Science
University of Science and Technology of China, Hefei 230027, China
[2]Department of Information Research
Zhengzhou Information Science and Technology Institute, Zhengzhou 450002, China
e-mail: yaoyz@mail.ustc.edu.cn, zwmshu@gmail.com, ynh@ustc.edu.cn

***Abstract*—Video transmission in error-prone networks suffers bit errors and packet losses that degrade the decoded video quality. In this paper, a video error concealment scheme based on H.264/AVC is proposed to enhance the error robustness using reversible data hiding. At the encoder, motion vectors of macroblocks belonging to the region of interest (ROI) are embedded into the background region within the same frame. The histogram modification technique based on quantized DCT coefficients of coded macroblocks is used to achieve lossless recovery after data extraction. If a macroblock in the ROI is found missing at the decoder, the embedded motion vector can be extracted from the corresponding host macroblock for the recovery of the missing macroblock. This proposed scheme makes a tradeoff between error concealment performance and compression efficiency. Experimental results show the better reconstructed video quality can be achieved compared with the related error concealment methods.**

Keywords-Error concealment; reversible data hiding; motion vector; FMO; video transmission

I. INTRODUCTION

With the rapid development of multimedia technologies and communication networks, the higher demand of various video applications is needed urgently. However, video transmission in error-prone environment, such as the wireless network, is always a tough work and has gained a lot of attention. The main problem is how to make a tradeoff between error robustness and video compression efficiency. Much related work has been done focusing on this problem.

H.264/AVC is the recent international video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group [1]. It builds on the concepts of earlier standards such as MPEG-2 and MPEG-4 and can obtain excellent compression efficiency and network-friendly characteristic which has become the most popular video coding standard. In order to enhance the error robustness of video transmission in error-prone environment, a number of error resilience and error concealment mechanisms are standard-compliant in H.264/AVC.

Flexible macroblock ordering (FMO) is an important error resilience tool defined in H.264/AVC which allows the encoder to assign MBs to slices in an explicit order other than the original scan order. In FMO, each macroblock (MB) is assigned to a slice group by using a macroblock allocation map. Prediction mechanisms are only allowed if the spatially neighboring macroblocks belong to the same slice group [2]. FMO benefits the error concealment a lot because the lost macroblock has several spatially neighboring macroblocks which belong to the other slice. Depending on this wonderful characteristic, FMO is usually combined with other effective techniques to enhance the error robustness.

Error concealment techniques are effectively used to recover the missing macroblocks at the decoder to improve the reconstructed video quality which make use of the high inherent spatial and/or temporal correlations of video sequences [3]. Spatial error concealment (SEC) recovers pixel values by spatially interpolating available pixels in neighboring macroblocks [4]. The method with weighted averages of the boundary pixels has become the spatial error concealment method implemented in the H.26L test model [5]. In the temporal error concealment (TEC), the motion vector (MV) of the missing macroblock is estimated according to the temporal correlation between consecutive frames and the decoder can copy the macroblock from the previous decoded frame to recover the missing one through motion compensation [6]. The classic boundary matching algorithm (BMA) is proposed in [7] to recover the lost MV from the candidate MVs which minimizes the total matching error between the internal boundary and the external boundary of the reconstructed block. Recently, some hybrid and effective spatial-temporal error concealment algorithms have been proposed. For example, Wu *et al*. [8] proposed to use the mode selection algorithm for SEC and the prediction-based MV estimation algorithm for TEC respectively. In [9], a cost function which exploits both spatial and temporal smoothness properties is introduced to recover the lost MV.

Error concealment combining with data hiding is a novel and effective technique, which is proposed for error concealment of intra-frames in [10]. By motion estimation between two intra-frames, the motion vector of a MB in the intra-frame is embedded into the quantized DCT coefficients of the other 16×16 MB in the current frame. If a MB is found missing at the decoder, the embedded motion vector can be extracted from the corresponding MB for the recovery of the lost MB. This method is developed by using reversible data hiding technique to achieve no quality degradation [11]. However, in [10] and [11], motion estimation between two intra-frames demand that the intra-frame period should be very short. If the distance of two consecutive intra-frames is too long, the accuracy of prediction may decrease severely. Besides, the insertion of intra-frames with a high frequency causes serious bit rate increment. In [12], Luo *et al*. proposed

978-0-7695-4852-4/12 $26.00 © 2012 IEEE
DOI 10.1109/MINES.2012.56

to embed motion vectors of the ROI to the background using the odd-even embedding method. Inspired by this method, we design an error concealment scheme using reversible data hiding to solve the quality degradation problem.

The remainder of this paper is organized as follows. Section II describes the implement of the proposed error concealment scheme. Section III shows the experimental results which confirm the superiority of our proposed scheme. Finally, Section IV summarizes the paper.

II. PROPOSED ERROR CONCEALMENT SCHEME

In this section, we describe the proposed adaptive error concealment scheme using reversible data hiding in detail. We first introduce the MBs allocation method combining with FMO. Then we present the histogram modification-based motion vectors embedding algorithm. The source and channel end-to-end distortion analysis is conducted at last to demonstrate the performance of the proposed scheme.

A. Adaptive ROI Determination and MBs Allocation

According to the human visual characteristic, people usually pay more attention to the region of interest (ROI) such as the center of the picture and the moving objects. Besides, in many practical video applications such as video telephony and video conferencing system, some regions draw more interest than the other regions within the same frame. In general, it is more necessary to purposely protect the ROI than the rest of the frame in the bandwidth-limited and error-prone transmission network.

Therefore, we compute the motion intensity to determine the relatively more important MBs which constitute the ROI. We obtain the motion intensity of a MB by calculating the frame difference between a current frame and a previous frame. It can be shown as

$$d(i,j,t) = \left| f(i,j,t) - f(i,j,t-1) \right| \tag{1}$$

in which $f(i,j,t)$ is the luma pixel value at position (i,j) and t is the frame index.

Then, the motion intensity of the MB with size M×N is defined as

$$I(x,y,t) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} d(i_0+i, j_0+j, t) \tag{2}$$

(i_0, j_0) represents the left-top pixel position of the MB in equation (2). The procedure of ROI determination can be shown as

$$ROI(x,y) = \begin{cases} 1, if \ I(x,y,t) \geq I_{th} \\ 0, if \ I(x,y,t) < I_{th} \end{cases} \tag{3}$$

(x,y) is the coordinate of the MB and I_{th} is the predefined threshold according to the content of the video. The MB with $ROI(x,y)$ which is equal to 1 belongs to the ROI. This above detection method provides a simple and effective compromise between the extra computational cost and the effectiveness in determining the MBs which constitute the ROI. After adaptive ROI determination, we use FMO to allocate the MBs in the ROI and other MBs in the background region into different slice groups. In order to prevent error drift and isolate the erroneous MBs, the MBs in the background region should be allocated into two slice groups using FMO Type 1, which is called dispersed type. The motion intensity can be computed at the first frame in the GOP and the other frames within the same GOP just use the same MBs allocation result as the first frame to reduce the encoding complexity. Fig. 1 illustrates the MBs allocation method based on the motion intensity.

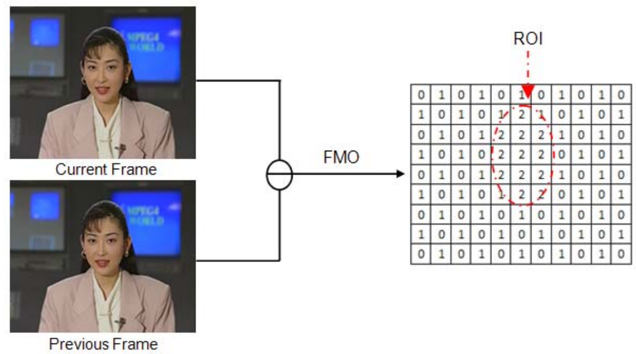

Figure 1. Illustration of ROI determination and MBs allocation.

B. Histogram Modification-based Motion Vectors Embedding Algorithm for Error Concealment

Due to the high temporal correlation between consecutive frames, the motion vector is undoubtedly the key information of a coded MB for motion compensation and error recovery. In [11], the histograms of quantized DCT coefficients are found to be very similar to the Laplace distributions. As has been pointed out in [13], a peak point corresponds to the grayscale value which the maximum number of pixels in the given image assumes and the number of bits that can be embedded into an image equals to the number of pixels which are associated with the peak point. In a similar way, in the histograms of quantized DCT coefficients, zero can be obviously the peak point. Therefore, no matter it is an I-frame or a P-frame, there is enough capacity for host MBs to contain the motion vectors.

In the proposed scheme, the motion vectors of 16×16 MBs in the ROI are obtained through motion estimation at half-pixel accuracy and the search range is set to ±15. The reference frame is the previous coded frame. The motion estimation is also used for I-frames for obtaining MVs. So the bit number for representing the motion vector can be denoted as

$$L = 2\times(\lceil \log_2(2\times 15+1) \rceil + 1) = 12 \tag{4}$$

We first convert the MV of each MB in the ROI into a binary sequence $b(n), 0 \leq n \leq 11$. Then, we embed these MVs of MBs in slice group 2 into the low-frequency quantized AC DCT coefficients of MBs in slice group 0 and slice group 1 which belong to the background region within the same frame respectively. We process the zero quantized DCT coefficients $Q_z(n)$ and the non-zero quantized DCT coefficients $Q_{NZ}(n)$ in zig-zag order using the histogram modification technique referring to the method in [11].

$$\overline{Q_{NZ}(n)} = \begin{cases} Q_{NZ}(n)+1, if\ Q_{NZ}(n) \ge 1 \\ Q_{NZ}(n), otherwise \end{cases} \quad (5)$$

$$\overline{Q_Z(n)} = \begin{cases} 1, if\ b(n)=1 \\ 0, if\ b(n)=0 \end{cases} \quad (6)$$

The data hiding is reversible, so the extraction of motion vectors can be conducted in the inverse process for temporal error concealment in equation (7), in which mv_x and mv_y represent the embedded motion vector components. The quantized DCT coefficients are same as those before embedding, so the reconstructed video quality will not degrade if the MBs are received in right at the decoder.

$$f(i,j,t) = f(i+mv_x, j+mv_y, t-1) \quad (7)$$

It should be noted that there is an obvious difference on quantized DCT coefficients between MBs in I-frames and MBs in P-frames. Therefore, we use the selective embedding method for P-frames to reduce the coded bit rate increment. Through inter prediction, there are more zero values of quantized DCT coefficients in inter-coded MBs and not each MB in the P-frame can be used as the host MB. If a MB in the P-frame is intra-coded or its coded mode is P_Skip, it is excluded from the host MBs. For host MBs in P-frames, we should modify their coded block pattern (CBP) after embedding. Therefore, the actual number of MVs that can be embedded in the P-frame is denoted as

$$n_{mv} = \mathbf{min}\{n_h, n_{ROI}\} \quad (8)$$

where n_h is the number of host MBs and n_{ROI} is the number of MBs in the ROI within a P-frame.

C. Error Concealment Performance Analysis

In [14] and [15], the effective rate-distortion optimization models in the packet loss video transmission environment have been proposed to analyze the overall performance of the video transmission system. For analysis simplicity, we think the overall distortion of one MB can be represented by

$$D_d(x,y) = (1-p)[D_s(x,y) + D_{ep\_ref}(x,y)] + pD_c(x,y) \quad (9)$$

where p denotes the packet loss rate; the D_s denotes the quantization distortion, D_{ep_ref} denotes the referenced potential error propagation distortion; and D_c denotes the error concealment distortion if the current MB is lost. (x,y) is the coordinate of the current MB.

In our proposed error concealment scheme, we divide all MBs into two classes and transport them in separate packets using FMO. The MBs in one of the classes belong to the ROI and the other MBs do not. The total distortion of one frame can be denoted as the sum of distortion of MBs in the two classes in equation (10).

$$D_f = \sum_{(x,y)\in non\_ROI} D_d(x,y) + \sum_{(x,y)\in ROI} D_d(x,y) \quad (10)$$

It is known that MBs in the ROI usually have more intense motion and more complex texture, so their effect on the distortion is relatively larger. We use reversible data hiding to purposely protect the MBs in the ROI, which does not impair the reconstructed quality of MBs in the background region. So in the bandwidth-limited network, the end-to-end distortion can be effectively decreased. Due to the selective embedding method, the total bits used for encoding video cannot increase greatly also.

III. EXPERIMENTAL RESULTS

In this section, the proposed error concealment scheme has been implemented in the Joint Model (JM) version 10.2 [16] of H.264/AVC. We use the packet loss program called RTP loss to simulate the error-prone channel. Then the lossy bit stream is decoded. We compare the error concealment performance of the proposed method with the BMA method in the JM 10.2, the DHEC method in [10] and the RDHEC method in [11]. The proposed method is called ROI-RDH.

For test sequence *Foreman* QCIF and *Carphone* QCIF, we set the period of I-frames to 15 and the number of reference frames to 1. For all error concealment methods, the total number of slice groups is set to 3. The GOP structure is set to IPPPP and the quantization parameter (QP) 28 is used to encode the two sequences for 100 frames respectively.

TABLE I. COMPARISON OF AVERAGE LUMA PSNR VALUES USING DIFFERENT METHODS AT 5% PACKET LOSS RATE

PSNR(dB)	BMA	DHEC	RDHEC	ROI-RDH
Foreman	32.58	32.00	33.23	**34.23**
Carphone	33.87	32.57	34.66	**35.35**
Average	33.23	32.29	33.95	**34.79**

TABLE II. COMPARISON OF AVERAGE LUMA PSNR VALUES USING DIFFERENT METHODS AT 10% PACKET LOSS RATE

PSNR(dB)	BMA	DHEC	RDHEC	ROI-RDH
Foreman	30.49	30.13	30.95	**31.14**
Carphone	32.19	31.67	33.26	**34.03**
Average	31.34	30.90	32.11	**32.59**

In the experiments, the 5% and 10% packet loss rates are randomly generated. Table I and Table II show the average luma PSNR values using different methods at 5% and 10% packet loss rates. It is observed that our proposed ROI-RDH method achieves the best PSNR performance in these four methods. The proposed ROI-RDH method outperforms BMA, DHEC and RDHEC about 1.41 dB, 2.10 dB and 0.66 dB respectively on average luma PSNR values. Due to the distortion caused by data hiding in DHEC, the average luma PSNR values are lower than those in BMA.

TABLE III. TOTAL BIT RATE INCREMENT COMPARED WITH BMA

%	BMA	DHEC	RDHEC	ROI-RDH
Foreman	-	1.37%	1.57%	4.97%
Carphone	-	1.30%	1.52%	3.27%

Table III shows the total coded bit rate increment using three data hiding-based methods compared with the MBA method. Due to the characteristic of the adaptive ROI determination and MVs embedding for I-frames and P-frames, the bit rate in the ROI-RDH method is relatively larger. But the coded bit rate in our proposed method can adaptively change according to the content of the video to achieve better error concealment performance. Totally speaking, the bit rate increment in our proposed method can maintain in an acceptable range.

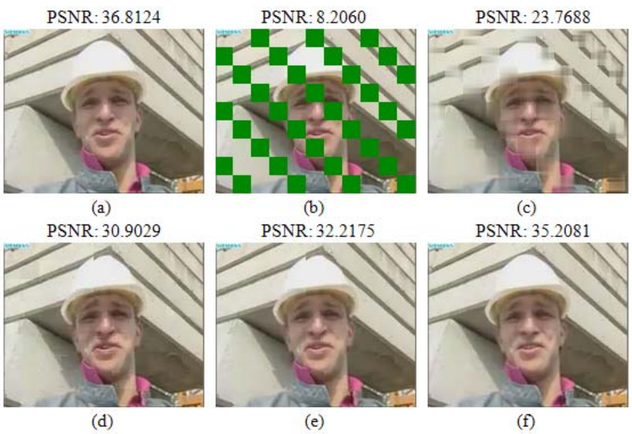

Figure 2. Visual quality comparison for *Foreman* sequence. (a) Error-free frame; (b) damaged frame with missing blocks; concealed using (c) JM 10.2; (d) DHEC; (e) RDHEC; (f) ROI-RDH.

For comparing the visual quality of the reconstructed video using different methods after the packet loss experiments, Fig. 2 shows the decoded subjective and objective quality of the 60th frame using different error concealment methods. The PSNR (dB) of luma component is labeled above each picture. The reconstructed picture quality using data hiding-based methods is obviously better than that using the spatial pixels interpolation method in JM10.2. It is observed that DHEC method and RDHEC method result in the similar visual quality. Because our proposed ROI-RDH method can improve the recovery performance for I-frames and P-frames, the best visual quality can be achieved.

IV. CONCLUSION

In this paper, we have proposed an adaptive video error concealment scheme to enhance the error robustness in error-prone transmission environment. We embed MVs of MBs belonging to the region of interest (ROI) into the background region within the same frame using reversible data hiding. And the extracted MVs can be used to recover the missing MBs at the decoder. Experimental results demonstrate that the proposed scheme improves the decoded video quality in error-prone transmission significantly and the performance analysis is also conducted.

It should be noted that the error concealment method based on reversible data hiding may increase the coded bit rate compared with other methods. As future work, we seek for possible extensions of our proposed scheme using efficient coding techniques to achieve high embedding efficiency and suppress the bit rate increment.

ACKNOWLEDGMENT

This work is partially supported by the Natural Science Foundation of China under Grant 61170234 and Grant 60803155, the Strategic and Piloted Project of CAS under Grant XDA06030601, and the National Science and Technology Major Project under Grant 2010ZX03004-003.

REFERENCES

[1] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, 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, Jul. 2003.

[2] S. Wenger, “H.264/AVC over IP,” IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 645-656, Jul. 2003.

[3] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, “Error resiliency schemes in H.264/AVC standard,” Journal of Visual Communication and Image Representation., vol. 17, no. 2, pp. 425-450, Apr. 2006.

[4] Z. Rongfu, Z. Yuanhua, and H. Xiaodong, “Content-adaptive spatial error concealment for video communication,” IEEE Trans. Consumer Electron., vol. 50, no. 1, pp. 335-341, Feb. 2004.

[5] Y.-K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, “The error concealment feature in the H.26L test model,” in Proc. Int. Conf. Image Processing., 2002, vol. 2, pp. 729-732.

[6] J. Wu, X. Liu, and K.-Y. Yoo, “A temporal error concealment method for H.264/AVC using motion vector recovery,” IEEE Trans. Consumer Electron., vol. 54, no. 4, pp. 1880-1885, Nov. 2008.

[7] W. M. Lam, A. R. Reibman, and B. Liu, “Recovery of lost or erroneously received motion vectors,” in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process., 1993, vol. 5, pp. 417-420.

[8] G.-L. Wu, C.-Y. Chen, T.-H. Wu, and S.-Y. Chien, “Efficient Spatial-Temporal Error Concealment Algorithm and Hardware Architecture Design for H.264/AVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 11, pp. 1409-1422, Nov. 2010.

[9] Y. Chen, Y. Hu, O. C. Au, H. Li, and C. W. Chen, “Video Error Concealment Using Spatio-Temporal Boundary Matching and Partial Differential Equation”, IEEE Trans. Circuits Syst. Video Technol., vol. 10, no. 1, pp. 2-15, Jan. 2008.

[10] S. Chen and H. Leung, “A Temporal Approach for Improving Intra-Frame Concealment Performance in H.264/AVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 3, pp. 422-426, Mar. 2009.

[11] K.-L. Chung, Y.-H. Huang, P.-C. Chang, and H.-Y. M. Liao, “Reversible Data Hiding-Based Approach for Intra-Frame Error Concealment in H.264/AVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 11, pp. 1643-1647, Nov. 2010.

[12] Z. Luo, L. Song, S. Zheng, Y. Xu, and X. Yang, “Improved error concealment of region of interest based on the H.264/AVC standard,” Optical Engineering., vol. 49, no. 4, Apr. 2010.

[13] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, “Reversible data hiding,” IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 3, pp. 354-362, Mar. 2006.

[14] Y. Zhang, W. Gao, H. Sun, Q. Huang, and Y. Lu, “Error resilience video coding in H.264 encoder with potential distortion tracking,” in Proc. Int. Conf. Image Process., 2004, vol. 1, pp. 163-166.

[15] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao , “Joint Source-Channel Rate-Distortion Optimization for H.264 Video Coding Over Error-Prone Networks,” IEEE Trans. Circuits Syst. Video Technol., vol. 9, no. 3, pp. 445-454, Apr. 2007.

[16] The H.264/AVC Joint Model , ver. 10.2 [Online]. Available: http://iphome.hhi.de/suehring/tml/download/old_jm/