ERROR CONCEALMENT BY REGION-FILLING FOR INTRA-FRAME LOSSES

Mengyao Ma[†], *Oscar C. Au*[‡], *Liwei Guo*[‡], *Yan Chen*[‡] *and S.-H. Gary Chan*[†]

[†] Dept. of Computer Science and Engineering, [‡] Dept. of Electronic and Computer Engineering
Hong Kong University of Science and Technology
{myma, eeau, eeglw, eeychan and gchan}@ust.hk

ABSTRACT

In this paper, we propose an *Error Concealment* algorithm for INTRA-frame losses over packet loss channels. The novelty is that not only the INTRA-frame but also the subsequent INTER-frames are refined using the received INTRA-MBs. Simulation results are given to demonstrate the performance of the proposed algorithm.

Index Terms— Error Concealment, Error Propagation, Motion Compensation, Region Filling

1. INTRODUCTION

Error Resilience (ER) and *Error Concealment* (EC) techniques are very important for video transmission today, due to the use of predictive coding and *Variable Length Coding* (VLC) in video compression [1]. Compared with ER, EC requires no change to the encoder and does not increase the bit rate, so it is more preferable for low bit-rate applications [2]. Lots of EC algorithms have been developed, such as spatial interpolation using some smoothness measure and temporal compensation based on inter-frame correlation [3][4]. Most of the current EC methods assume that only a few MBs or slices in a video frame are lost. However, in low bit-rate applications, one frame is usually carried in one data packet in order to save transmission overhead. As a result, the loss of one packet will lead to the loss of one entire frame [2]. When frame loss occurs, we can copy the previous received frames to reconstruct the lose ones. More sophisticated methods recover the motion vectors (MVs) in pixel or block level based on the assumption of translational motion [2][5][6].

As in most of the block-based video coding systems all the INTER-frames are encoded based on the preceding INTRA-frame, the protection and restoration of INTRA-frames is especially important for the decoding of subsequent frames. However, as far as we know, most of the EC algorithms in the literature focus on the restoration of INTER-frames, and only a few works deal with the EC of INTRA-frames. In addition, almost all these algorithms assume that only part of the INTRA-frame is corrupted so that the lost MBs can be reconstructed using the information from the neighbors [7][8]. Since in low bit-rate video transmissions the loss of a packet usually results in the loss of a whole frame, an EC algorithm for INTRA-frame losses is necessary in reality. In this paper, we will focus on this problem and propose an algorithm to improve the reconstructed video quality when INTRA-frame loss occurs. The novelty is that not only the INTRA-frame but also the subsequent INTER-frames are refined using the received INTRA-MBs.

Random INTRA Refresh (RIR) scheme has been used in both earlier and current standards such as MPEG-4, H.263 and H.264, where INTRA-coded MBs are randomly inserted into the bitstream to remove artifacts caused by error and INTER-prediction drift. Although coding efficiency is reduced a little, RIR with a low INTRA-rate is more practical than inserting periodic INTRA-frames due to the bit-rate constraint [9]. As the RIR scheme is implemented in the encoder and does not introduce any decoding overhead, it is often jointly used with other ER or EC schemes. In our algorithm we assume that the received bitstream contains such INTRA-MBs. When an INTRA-frame is lost, the received INTRA-MBs in the subsequent frames can be used to refine their INTER-neighbors (MBs) based on the strong correlation between adjacent pixel values. We use the idea of region-filling algorithm proposed in [10] to fill the target pixels, where higher priority of synthesis is given to the regions along strong edges. As it is a patch-based filling approach, the overhead caused by EC can be reduced compared to the pixel-based ones. In addition, *Motion Compensation* (MC) is also used to refine the INTER-pixel, which has an INTRA-pixel in its motion trajectory.

The rest of this paper is organized as follows. In Section 2, we describe the proposed EC algorithm. Its performance is demonstrated in Section 3 by simulation results. Section 4 is conclusion.

2. THE PROPOSED EC ALGORITHM

In conventional EC algorithms, only the corrupted (lost) frames are error-concealed. Although the subsequent frames can be decoded as usual, some annoying artifacts will exist due to the drifting errors and the video quality can be even worse in the case of INTRA-frame loss. In this work, we propose to use three ways to reconstruct the subsequent INTER-MBs after a lost INTRA-frame:

- Decoding directly as in the conventional codec;
- Error concealment by motion compensation (MC);
- Error concealment by region-filling (RF).

After decoding the INTER-frame, each MB inside is refined using the algorithm illustrated in Figure 1. We will describe the two EC ways (MC and RF) in the following subsections. In addition, as the INTRA-MBs coded by *Random INTRA Refresh* (RIR) can help to stop the propagated error, for each pixel we use one mark to represent whether it is error-free (*refreshed*) or not. For a lost frame, all the pixels are set to be *non-refreshed*. And when an INTRA-MB is received later, the corresponding pixels are marked *refreshed*. So a map needs to be maintained for each frame in the frame buffers, one bit for one pixel.

2.1. EC by MC

In this step, the given INTER-MB (MB_c) is refined pixel by pixel. Suppose there are L frames in the reference frame buffer. For each pixel p in MB_c, we have its motion vector MV_0 and the corresponding reference frame index k_0, $k_0 \in \{1, 2, \ldots, L\}$.

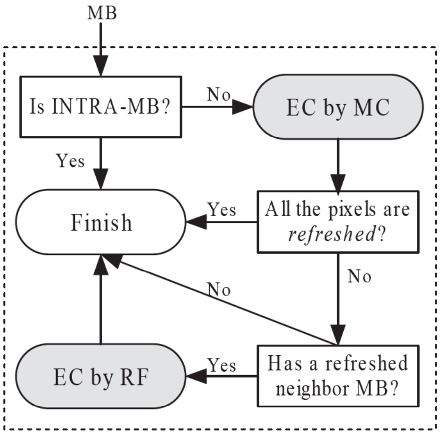


Fig. 1. The flow chart of the proposed EC algorithm.

Then p can be refined by motion compensation (MC) if there is a *refreshed* pixel in its motion trajectory. In detail,

1) Use MV_0 to find the reference pixel of p, i.e. q_0. If q_0 lies at an integer-pixel position marked as *refreshed*, or if q_0 lies at a sub-pixel position surrounded by *refreshed* pixels, mark p as *refreshed* and stop. Otherwise, set $k = 0$ and go to 2).
2) Increase k by 1. If k is great than L, i.e. all the reference frames have been checked, stop. Otherwise, go to 3).
3) If k equals k_0, go to 2). Otherwise, estimate the MV of p to the kth reference frame based on the constant velocity model, i.e. $MV_k = MV_0 \times k/k_0$. Then Use MV_k to find the corresponding pixel q_k in the kth reference frame. If q_k lies at an integer-pixel position marked as *refreshed*, or if q_k lies at a sub-pixel position surrounded by *refreshed* pixels, replace p by the pixel value of q_k and stop. Otherwise, go to 2).

After EC by MC, we check the status of each pixel in MB_c. If all the pixels are marked as *refreshed*, we have finished the EC of MB_c and can go to the next MB.

2.2. EC by RF

For a given INTER-MB MB_c, we first check whether its neighboring MB has been fully refreshed, i.e. all the pixels are marked as *refreshed*. Four neighbors are checked: MB_u, MB_b, MB_l and MB_r, which correspond to the upper, the bottom, the left and the right MB, respectively. So MB_c can be error-concealed by region-filling (RF) from at most four directions. In detail, suppose MB_u has been fully refreshed. Then the current MB will be filled from top to bottom using the pixel values extracted from MB_u. Suppose the obtained MB is MB_c^u. The region-filling steps are listed as follows:

1) Mark all the pixels of MB_c^u as unfilled, e.g. -1. Initialize the row index of MB_c^u as -1 and go to 2).
2) Increase the row index by 1. If the row index exceeds 15, i.e. all the pixels have been filled, stop. Otherwise, go to 3).
3) If all the pixels in the current row have been filled, go to 2). Otherwise, for each unfilled pixel in this row, compute its horizontal gradient G_x. The gradient is estimated by applying *Sobel* filter on the surrounding filled pixels. Find the pixel with maximal G_x, i.e. $\hat{p}$, and go to 4).
4) Define patch $\psi_{\hat{p}}$ to be the $S \times S$ window centered at $\hat{p}$. Search in the source region (MB_u) for the patch which is most similar to $\psi_{\hat{p}}$. Formally,

$$\psi_{\hat{q}} = \arg \min_{\psi_q \in MB_u} d(\psi_{\hat{p}}, \psi_q), \tag{1}$$

where the distance between the two patches, i.e. $d(\psi_{\hat{p}}, \psi_q)$, is defined as the sum of square difference (SSD) of the already filled pixels in the two patches [10]. The Luma-components of the pixel values are used in the calculation. Then after copying the corresponding pixel values from $\psi_{\hat{q}}$ into the unfilled region of $\psi_{\hat{p}}$, go to 3).

Similarly, we can reconstruct the current MB by RF from other directions. Suppose the obtained MB by extrapolating MB_i is MB_c^i and $MB_c^i(x,y)$ is its pixel value at position (x,y), $i \in \{u,b,l,r\}$ and $x,y \in [0,15]$. The error-concealed MB using RF (MB_c^{rf}) is a weighted summation of these four MBs:

$$MB_c^{rf}(x,y) = \frac{\sum_{i\in\{u,b,l,r\}} w_i(x,y) \times MB_c^i(x,y)}{\sum_{i\in\{u,b,l,r\}} w_i(x,y)}, \tag{2}$$

where $w_i(x,y)$ is a weighting factor. Define $D_i(x,y)$ to be the distance from position (x,y) to the nearest boundary of MB_i, $i \in \{u,b,l,r\}$. The weighting factors are calculated as

$$w_i(x,y) = \begin{cases} \frac{1}{\|D_i(x,y)\|^2}, & \text{if } MB_i \text{ is fully refreshed,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Suppose the obtained MB in the previous step (*EC by MC*) is MB_c^{mc}. Then for each pixel in MB_c marked as *non-refreshed*, its final reconstructed value after RF is

$$MB_c(x,y) = w_{rf} \times MB_c^{rf}(x,y) + (1 - w_{rf}) \times MB_c^{mc}(x,y), \tag{4}$$

where weight w_{rf} is used to control the strength of the region-filling effect.

2.3. Summary for the EC algorithm

If an INTRA-frame (I_0) is lost, all the pixels inside are filled by grey color, i.e. 128 for all the YUV components. Each of the subsequent N frames is decoded and then error-concealed as follows until an INTER-frame is lost. Here N is an integer to control the number of frames for EC.

- For the first INTER-frame (P_1), compute the DC of the INTRA-MBs within this frame first, i.e. DC_{intra}. Fill the reference frame of P_1 (the buffer for I_0) and the *non-refreshed* pixels of P_1 by DC_{intra}. Then refine P_1 using the EC algorithm in Figure 1.
- For the subsequent frames, the INTER-MBs are error concealed as in Figure 1.

If an INTER-frame is lost, it is reconstructed by copying the previous frame (*copy-previous*).

2.4. The parameters used in the algorithm

Note that we have three parameters in the proposed EC algorithm: the patch size $S \times S$ in region-filling, the weighting parameter w_{rf} in Eqn. (4), and the frame number N to apply the EC algorithm. To determine their values, we have some observations listed as follows.

- Patch size $S \times S$: If S is very small, the samples in a patch is not sufficient and thus the patch found by Eqn. (1) may not be trustworthy. On the other hand, as the patch distance needs to be calculated in the patch-searching process, the complexity of RF increases with S. To make a balance between the performance and the complexity, we use $S = 5$ in our simulation.

- Weighting parameter w_{rf}: We can use some traditional approach, e.g. the one based on the minimum mean square error (MMSE) criterion, to adaptively determine the value of w_{rf} in Eqn. (4). However, as it is difficult to estimate the error variance from the received data, the calculated w_{rf} may not be accurate enough to reconstruct MB_c with a good quality. So we will just use a constant w_{rf} in this paper. Its value is trained by the experimental data. We find that when w_{rf} increases from 0.1 to 0.9, the reconstructed video quality (PSNR) increases at first and then decreases. This means a moderate value of w_{rf}, e.g. $w_{rf} = 0.5$, is more suitable.
- Frame number N: As RIR is jointly used to stop the propagated errors, we can roughly estimate the value of N from the INTRA-rate. For example, if the INTRA-rate is 3%, the error caused by a frame loss can be totally removed, on average, after 33.3 frames. As most of the pixels have been *refreshed* (error-free) after 30 frames, there is no need to apply error concealment on the subsequent frames. By doing this, the complexity of the algorithm is reduced.

3. SIMULATION RESULTS

We use the JVT reference software version 11.0 (baseline profile) for the simulations [11]. The first 300 frames of video sequences *Foreman* and *Akiyo* (QCIF) are encoded at 7.5fps, and only the first frame is I frame. Two reference frames are used for INTER-prediction. Parameter *UseConstrainedIntraPred* is set to be 1 in the reference software, i.e. INTER pixels are not used for the prediction of INTRA-MB. And the INTRA-rate for RIR is 3%. The compressed video is supposed to be transmitted though a packet loss channel, and one packet contains the information of one frame. The simulated packet loss patterns are obtained from [12], with loss rate $P = 3\%$, 5%, 10%, or 20%. Decoder PSNR is used as the objective measurement, which is computed using the original uncompressed video as reference. Given a packet loss rate P, the video sequence is transmitted 40 times, and the average PSNR for the 40 transmissions is calculated at the decoder side. Three EC algorithms are evaluated, which will act as follows in the case of INTRA-frame loss:

- *EC_F0_128*: The lost INTRA-frame is error-concealed by filling 128 for all the YUV components.
- *EC_F01_DC*: The lost INTRA-frame and the first INTER-frame are error-concealed, similarly as in *EC_MC_RF*. However, the neighbors of an INTRA-MB are filled by its DC instead of by region-filling.
- *EC_MC_RF*: The proposed algorithm in section 2.3, with parameters $S = 5$, $w_{rf} = 0.5$ and $N = 30$.

The lost INTER-frame is error-concealed by *copy-previous* for all these three algorithms.

We first simulate the case of INTRA-frame loss, and all the subsequent frames are assumed to be received. Constant QP (QP=30) is used to encode both *Foreman* and *Akiyo*. The decoder PSNR is computed and plotted in Figure 2. As shown in the figure, the video quality can be improved by just error-concealing the first two frames, i.e. filling with the DC of received INTRA-MBs. However, with the proposed *EC_MC_RF* algorithm, much more improvement can be obtained.

Table I. Average decoder PSNR for different loss rate P (QP=30).

P	3%	5%	10%	20%
Foreman	Decoder PSNR			
EC_F0_128	29.41	26.52	23.82	20.24
EC_F01_DC	29.47	26.59	23.94	20.51
EC_MC_RF	29.59	26.72	24.03	20.59
	Delta-PSNR			
EC_F01_DC	0.06	0.07	0.12	0.27
EC_MC_RF	0.18	0.20	0.21	0.35
Akiyo	Decoder PSNR			
EC_F0_128	34.78	33.92	31.90	28.27
EC_F01_DC	34.82	34.00	32.02	28.63
EC_MC_RF	34.94	34.23	32.18	28.84
	Delta-PSNR			
EC_F01_DC	0.04	0.08	0.12	0.36
EC_MC_RF	0.16	0.31	0.28	0.57

The performances of the EC algorithms under random packet loss conditions are given in Table I and Figure 3. Table I shows the average decoder PSNRs for video transmission under different packet loss rate P. To give a clearer illustration, we also present the difference between *EC_F01_DC*/*EC_MC_RF* and *EC_F0_128* for the same loss rate, as shown in the column named Delta-PSNR. From the table we can see that both *EC_F01_DC* and *EC_MC_RF* can obtain a higher PSNR than *EC_F0_128*, and the difference will increase with larger P. Figure 3 compares the RD curves of the three EC algorithms for a given packet loss rate P. From the figure we can see that by using *EC_MC_RF*, we can gain about 0.2dB for *Foreman* ($P = 5\%$) and about 0.6dB for *Akiyo* ($P = 20\%$), compared to using *EC_F0_128*.

Not that in Table I and Figure 3, the gap between *EC_MC_RF* and *EC_F0_128* is smaller than that in Figure 2. As in the case of INTER-frames losses, the two algorithms have the same action, i.e. *copy-previous*, and the advantage of *EC_MC_RF* over *EC_F0_128* is not obvious. Actually in such conditions, the received INTRA-MBs can also be used to refine the subsequent INTER-frames. We will take this as a future work.

4. CONCLUSION

In this paper, we propose an EC algorithm for INTRA-frame losses over packet loss channels. Both motion compensation and region-filling algorithm are used to refine the INTER-MBs in the subsequent frames. As a result, the propagated error can decrease much faster than just error-concealing the lost INTRA-frame.

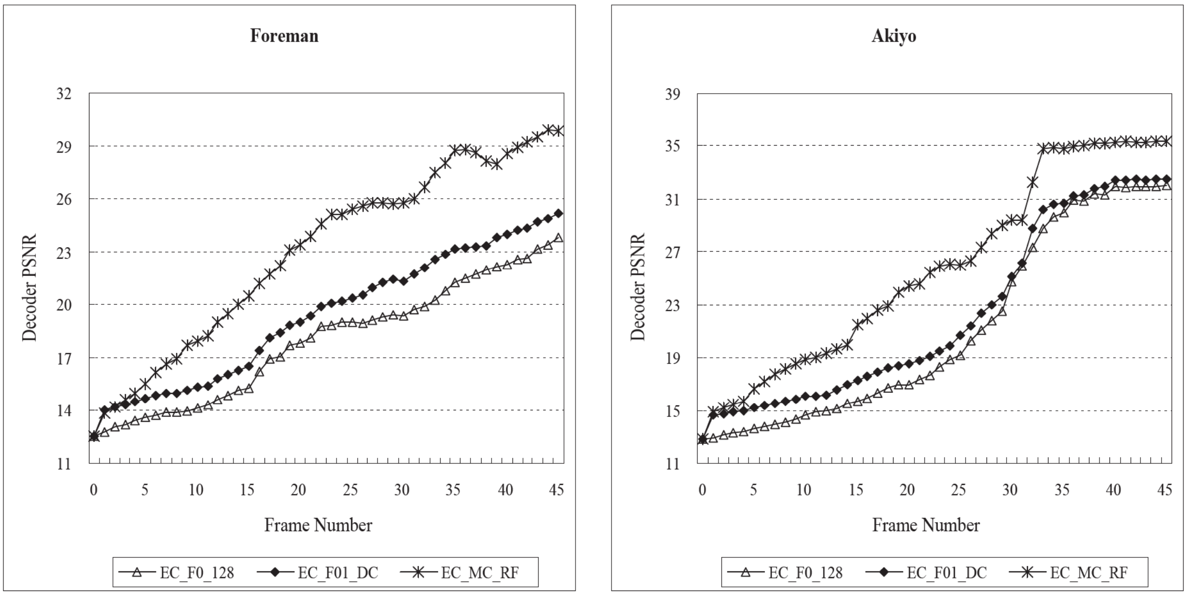
ACKNOWLEDGMENT

This work has been supported by the Innovation and Technology Commission of the Hong Kong Special Administrative Region, China (project no. GHP/033/05).

REFERENCES

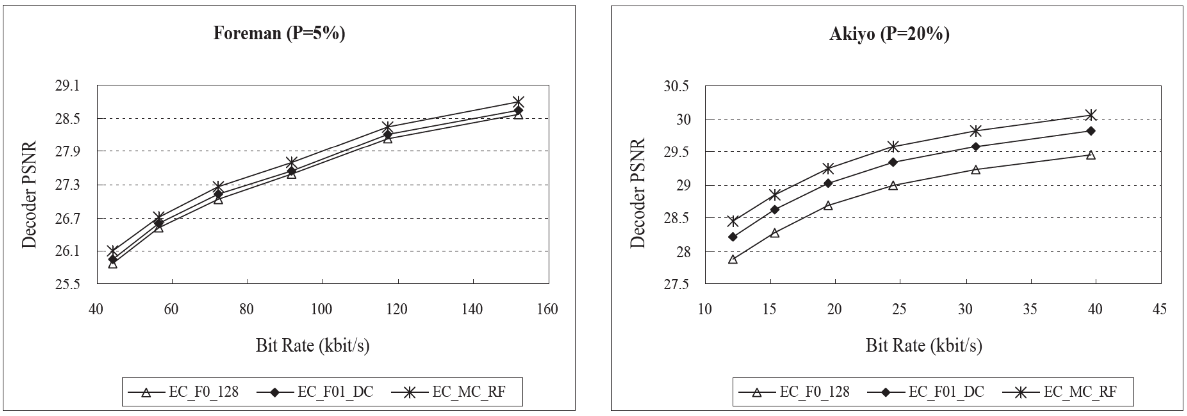
[1] Y. Wang and Q. F. Zhu, “Error control and concealment for video communication: a review,” in *Proc. IEEE*, May 1998, pp. 974–997.

[2] Y. Chen, K. Yu, J. Li, and S. Li, “An error concealment algorithm for entire frame loss in video transmission,” in *Proc. PCS*, Dec. 2004.



(a) (b)

Fig. 2. The decoder PSNR of different EC algorithms for INTRA-frame loss.



(a) (b)

Fig. 3. The RD curves of different EC algorithms. (a) *Foreman* with loss rate $P = 5\%$; (b) *Akiyo* with loss rate $P = 20\%$.

[3] W. Zhu, Y. Wang, and Q.-F. Zhu, “Second-order derivative-based smoothness measure for error concealment in DCT-based codecs,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 713 –718, Oct. 1998.

[4] S.-C. Hsia, S.-C. Cheng, and S.-W. Chou, “Efficient adaptive error concealment technique for video decoding system,” *IEEE Trans. Multimedia*, vol. 7, pp. 860–868, Oct. 2005.

[5] S. Belfiore, M. Grangetto, E. Magli, and G. Olmo, “Concealment of whole-frame losses for wireless low bit-rate video based on multiframe optical flow estimation,” *IEEE Trans. Multimedia*, vol. 7, pp. 316–329, Apr. 2005.

[6] Z. Wu and J. M. Boyce, “An error concealment scheme for entire frame losses based on H.264/AVC,” in *Proc. IEEE ISCAS*, May 2006, pp. 4463–4466.

[7] P. Nasiopoulos, L. Coria-Mendozal, H. Mansour, and A. Golikeri, “An improved error concealment algorithm for intra-frames in H.264/AVC,” in *Proc. IEEE ISCAS*, May 2005, pp. 320–323.

[8] H. Wang and J. Lv, “A novel error concealment scheme for intra frames of H.264 video,” in *Proc. IEEE Int. Workshop VLSI Design & Video Tech.*, May 2005, pp. 300–303.

[9] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, “Error resiliency schemes in H.264/AVC standard,” *Elsevier J. Vis. Commun. Image Represent.*, vol. 17(2), pp. 425–450, Apr. 2006.

[10] A. Criminisi, P. Prez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE Trans. Image Processing*, vol. 13, pp. 1200–1212, Sept. 2004.

[11] Jvt reference software, version 11.0. [Online]. Available: http://iphome.hhi.de/suehring/tml/download/

[12] S. Wenger, “Error patterns for internet experiments,” in *ITU-T SG16 Doc. Q15-I-16r1*, Oct. 1999.