Bilateral Motion Estimation Based Frame Error Concealment for Video Transmission over Wireless Networks

Dinh Trieu Duong, Jun-Hyung Kim, Sung-Hyun Cheon, Yeo-Jin Yoon, and Sung-Jea Ko
Department of Electronic Engineering, Korea University, Seoul, Korea
Anam-Dong, Sungbuk-ku, Seoul 136-701, Korea
E-mail : sjko@korea.ac.kr

Abstract: In this paper, we propose a new error concealment algorithm that can effectively reconstruct the lost frame caused by transmission errors in the wireless environment. The proposed algorithm employs the bilateral motion estimation scheme where the received motion vectors (MVs) in the neighboring frames are weighted and utilized to construct the MV field for the concealed frame. Experimental results show that the proposed algorithm outperforms other techniques in term of both peak signal-to-noise ratio (PSNR) performance and subjective visual quality.

1. Introduction

For robust video transmission over wireless networks, it is necessary to cope with several problems including transmission error, channel bandwidth variation, and fading effects [1][2]. Among them, the transmission error is the most critical problem since it may cause the loss of synchronization required between the encoder and decoder for the correct decoding. Compressed video streams are also extremely vulnerable to the transmission error due to the use of temporal predictive coding and variable length coding (VLC). Thus, a single erroneously recovered sample can lead to error propagations in the following frames.

For video transmissions over low bit-rate wireless networks, the compressed video frames may have small size, so that the payload of a single packet can often contain a whole coded frame [3]. In this case, the loss of a single packet can leads to the loss of a whole video frame. Without modifying standard compression codecs and increasing bit-rate, error concealment (EC) at the decoder is an effective method to reconstruct the lost frame and protect the image quality against the transmission errors [4].

There have been several algorithms proposed for the whole frame EC. Bandyopadhyay *et al.* [5] introduced frame copy (FC) and motion vector copy (MVC) methods for concealing the lost frame. In the FC method, the lost frame is simply replaced by the previous frame without considering the motion between the previous and current frames. In the MVC method, the MVs of the co-located blocks in the previously decoded reference frame are copied to the current lost frame, and then the lost frame is reconstructed using these copied MVs. The performance of the MVC method is generally better than that of the FC method. However, for the sequences with complicated motions, the MVC method may yields visible blocking artifacts at the reconstructed block boundaries. Belfiore *et al.* [6] exploited the hypothesis of the optical flow estimation to improve the accuracy of motion estimations for the missing frame. This algorithm provides good EC performance in term of PSNR and visual quality, but it requires pixel operations with high computational complexity. In order to reduce the computational complexity, Baccichet *et al.* [7] utilized the block based instead of pixel based concealment algorithm. The algorithm is also based on the optical flow estimation and conventional motion compensation to reconstruct the lost frame.

In this paper, we propose a novel error concealment algorithm that can effectively reconstruct the lost frame. The proposed algorithm employs the bilateral motion estimation scheme where the received MVs in the neighboring frames are weighted and utilized to construct the MV field for the concealed frame. Based on the constructed MV field, the lost frame is motion-compensated from the previous frame in a blockwise fashion. Unlike the conventional algorithms, the proposed scheme does not produce any overlapped pixel and hole region in the reconstructed frame.

The rest of the paper is organized as follows. Section 2 describes the proposed algorithm in detail. Experimental results are discussed in Section 3. Finally, Section 4 concludes this paper.

2. Proposed Frame Error Concealment

In order to reconstruct the lost frame, we need to re-estimate MVs of all blocks which are lost due to the transmission error. Based on the estimated MVs, the lost frame is motion-compensated from the reference frames. In the proposed algorithm, the bilateral motion estimation scheme is performed to construct the MV field which includes MVs of all missing blocks in the lost frame. We assume that the pixels and MVs of the current frame are lost, while the MVs of the previous and next frames are correctly received at the decoder. In the proposed algorithm, the motion estimation is performed on each separated missing block which is non-overlapped with the others in the lost frame. Therefore, the problem of hole and overlapping regions which degrade the image quality of the concealed frame can also be solved.

Let F_n denote the nth frame, F_{n-1} and F_{n+1} be the previous and next frames of F_n, respectively. F_{n-1} is the last correctly decoded frame and F_n is the lost frame due to the transmission errors. In Fig. 1, B_n^k is the kth block in F_n and $\widehat{V}$ is the one of candidate MVs estimated for B_n^k. For each $\widehat{V}$, we can define two predicted blocks $\widehat{B}_{n-1}^k$ and $\widehat{B}_{n+1}^k$ of B_n^k in F_{n-1} and F_{n+1}, respectively as shown in Fig. 1. Generally, since blocks $\widehat{B}_{n-1}^k$ and $\widehat{B}_{n+1}^k$ can be located at the non-grid block position in F_{n-1} and F_{n+1}, the MVs of these blocks do not exist. In the proposed algorithm, we use the weighted average of the received MVs in F_{n-1} and F_{n+1} to compute

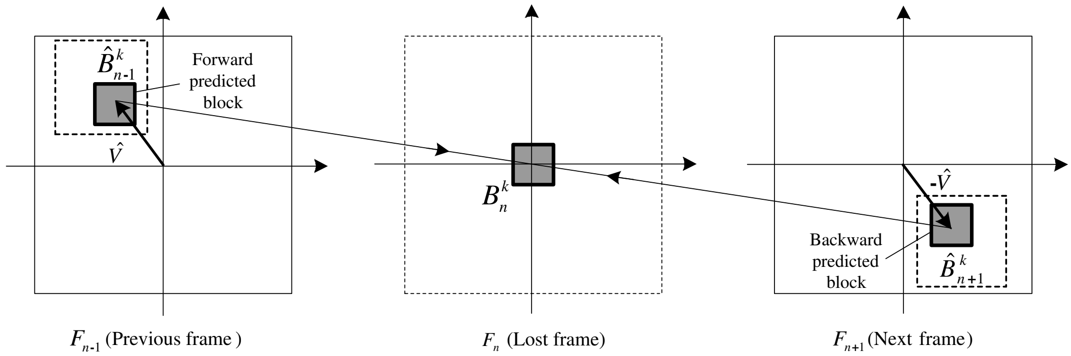


Figure 1. Bilateral motion estimation for the missing block.

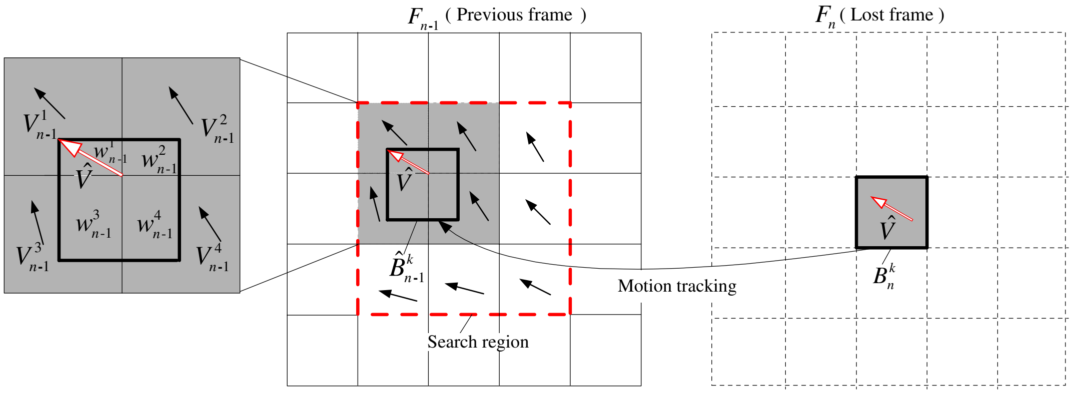


Figure 2. Motion vector computation for the predicted blocks in the previous frame.

the MVs for $\widehat{B}^k_{n-1}$ and $\widehat{B}^k_{n+1}$. Fig. 2 shows an example of the MV computation for $\widehat{B}^k_{n-1}$ in F_{n-1}. Four blocks with the received MVs V^i_{n-1} $(i = 1, 2, 3, 4)$ are overlapped with $\widehat{B}^k_{n-1}$, then the MV of $\widehat{B}^k_{n-1}$ is determined by

$$\widetilde{V}_{n-1} = \frac{\sum_{i=1}^{N} w^i_{n-1}.V^i_{n-1}}{\sum_{i=1}^{N} w^i_{n-1}} \quad (1)$$

where the weighting factor w^i_{n-1} is equal to the number of pixels in the ith overlapped region and N is the number of overlapped regions.

Similarly, in the next frame F_{n+1}, the MV of $\widehat{B}^k_{n+1}$ is given by

$$\widetilde{V}_{n+1} = \frac{\sum_{i=1}^{N} w^i_{n+1}.V^i_{n+1}}{\sum_{i=1}^{N} w^i_{n+1}} \quad (2)$$

For each vector $\widehat{V}$ in a given search range $\Re$ as shown in Fig. 2, we can obtain the corresponding MVs $\widetilde{V}_{n-1}$ and $\widetilde{V}_{n+1}$ using (1) and (2). Among different candidate MVs in the search range, to estimate an optimal one for the missing block B^k_n, we use a novel measure, namely the absolute motion difference (AMD) defined by

$$\text{AMD}(\widehat{V}) = \alpha(|\widehat{V} - \widetilde{V}_{n-1}|) + (1-\alpha)(|\widehat{V} + \widetilde{V}_{n+1}|) \quad (3)$$

where α is a weighting coefficient of the AMD.

In (3), the first term represents the relation of $\widehat{V}$ and $\widetilde{V}_{n-1}$ in the forward prediction direction, while the second one represents the relation of $\widehat{V}$ and $\widetilde{V}_{n+1}$ in the backward direction. It is generally assumed that the MVs on the consecutive frames are analogous to each other because the movements of objects in real video sequences tend to be smoothly changed [6]. Thus, along the motion trajectory passing through B^k_n from $\widehat{B}^k_{n-1}$ to $\widehat{B}^k_{n+1}$, it is seen that $\widetilde{V}_{n-1}$ and $\widetilde{V}_{n+1}$ generally have high correlations with the true MV estimated for $\widehat{V}$. Therefore, in the proposed algorithm, to estimate the optimal MV for the missing block B^k_n, we search the best motion trajectory passing through B^k_n from $\widehat{B}^k_{n-1}$ to $\widehat{B}^k_{n+1}$ by minimizing the AMD as follows:

$$V_{opt} = \arg\min_{\widehat{V} \in \Re} \text{AMD}(\widehat{V}) \quad (4)$$

3. Experimental Results

Several experiments have been performed to illustrate the effectiveness of the proposed method. The experiment results are reported for several video sequences using JM12.2 reference software of H.264/AVC standard [8]. The test sequences including *Table tennis*, *Foreman*, and *Football* are in 4:2:0

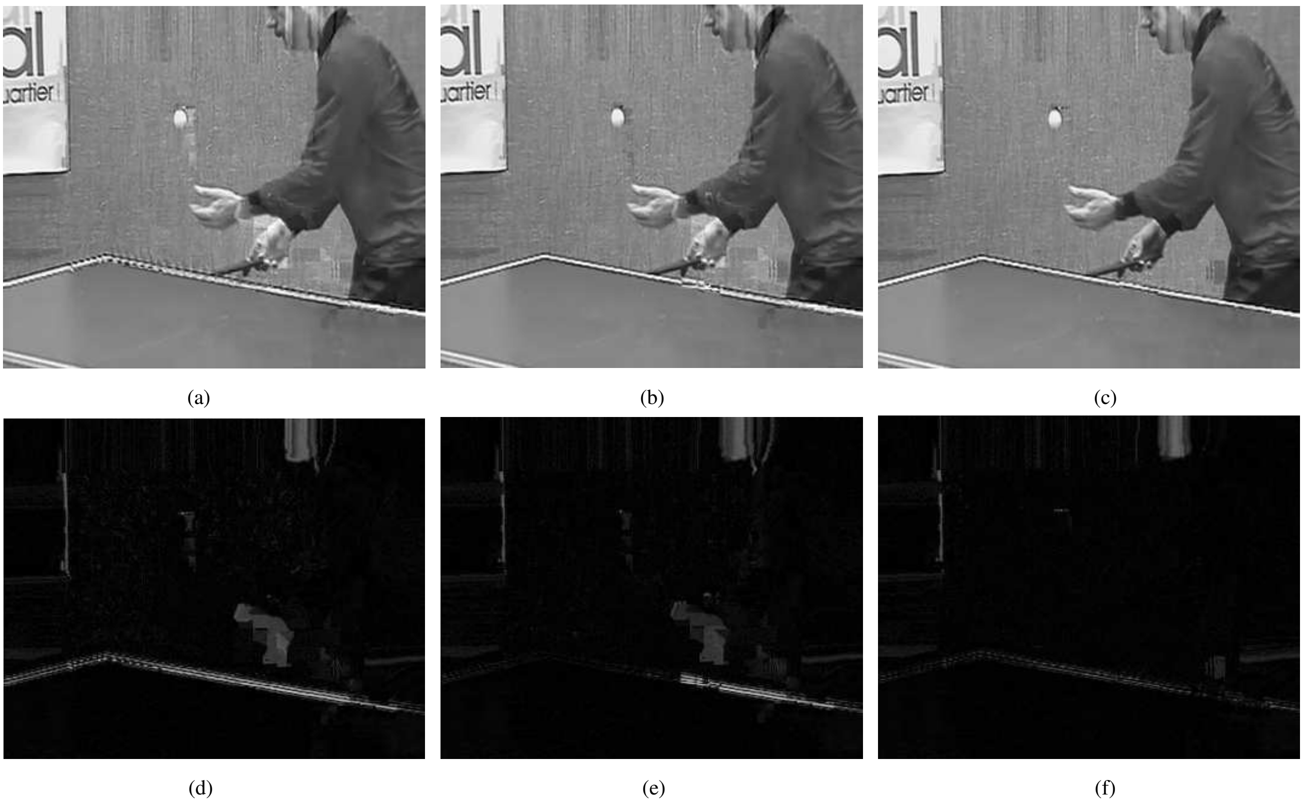


Figure 3. Comparison of the visual quality for the $8th$ concealed image of the *Table tennis* sequence. (a) FC method. (b) CA_B method. (c) Proposed method. The bottom row reports the error images between the correctly decoded frame and conceal frames. (d) FC method. (e) CA_B method. (f) Proposed method.

formats with CIF (352×288) resolution. The MV is only estimated for the luminance component and further used for the chrominance components. We also configure the test video sequences to use group of pictures (GOP) structure for effectively stopping error propagation; in particular, GOPs of 30 frames have been formed. In our experiments, quantization parameters (QPs) are set to 26, 32, and 38 for each test sequence. The weighting coefficient α is set to 0.5. The packets are randomly dropped according to the predefined packet loss rate (PLR). The values of PLR are simulated as 5%, 10% and 15% since they are typical values represented for wireless environment [2].

The performance of the proposed method is compared with that of two conventional EC methods, namely the frame copy (FC) in [5] and the concealment algorithm on blocks (CA_B) in [7]. In the FC, MVs of all MBs in the lost frame are set to zero. Although CA_B uses the optical flow concept of the concealment algorithm on pixels (CA_P) in [6], it works at block level instead of pixel level to meet the requirement of real-time applications.

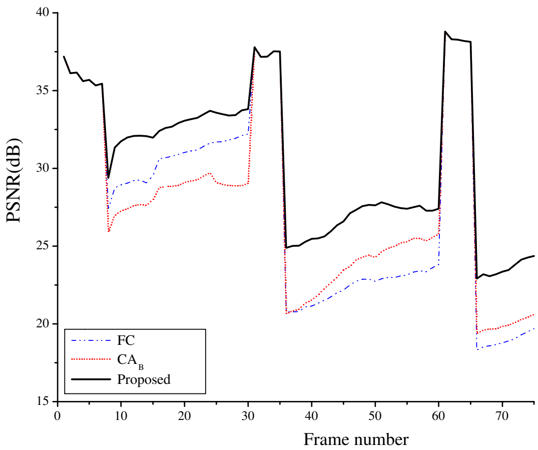
We compare the subjective image quality of the frames which are concealed by the FC, CA_B and proposed methods. Fig. 3(a), Fig. 3(b), and Fig. 3(c) report the concealed results of $8th$ frame in the *Table tennis* sequence of the FC, CA_B, and proposed methods, respectively. Error images in the bottom row in Fig. 3 represent the differences between correctly decoded frame and concealed frames by the FC in Fig. 3(d), CA_B in Fig. 3(e), and the proposed method in Fig. 3(f). Although the result of the FC in Fig. 3(a) seems to have no errors, it is not the current reconstructed frame but the previous frame. Therefore, the FC introduces much error areas on the concealed frame as shown in Fig. 3(d). In Fig. 3(b) and Fig. 3(e), the performance of the CA_B is better than that of the FC, but annoying blocking artifacts seen around the right hand areas of the tennis player in the reconstructed frame are considerable problem in this method. It can be observed from Fig. 3(c) and Fig. 3(f) that our proposed method yields the most faithful image quality and significant reduction of blocking artifacts.

The proposed method introduces not only better subjective image quality but also higher PSNR performance as com-

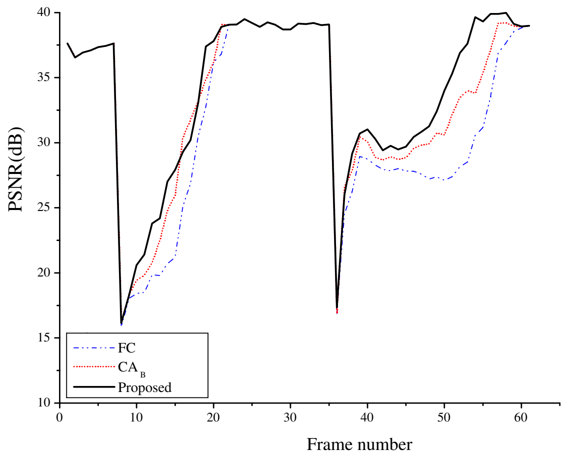
Table 1. Comparison of the average PSNR with different PLRs and QPs on *Foreman*.

(dB)

PLR	QP	FC	CA_B	Proposed
5%	26	31.45	32.13	**33.62**
	32	30.12	31.49	**32.78**
	38	28.24	28.60	**30.37**
10%	26	30.34	31.63	**32.51**
	32	29.44	30.12	**31.84**
	38	28.17	28.32	**29.68**
15%	26	29.84	30.73	**32.22**
	32	29.64	29.72	**31.24**
	38	27.97	27.82	**28.18**



(a)



(b)

Figure 4. Comparison of the PSNR performance with QP 26 and PLR 5%. (a) *Table tennis*. (b) *Football*.

pared with the conventional methods. Table I reports the average PSNR performance of the conventional and proposed algorithms on the *Foreman*. When calculating the average PSNR, we consider all frames on the test sequences, including the concealed frames and their next frames deteriorated by the error propagation. As reported in Table I, the proposed EC algorithm consistently provides better performance than the FC and CA_B. For example, the proposed method provides up to 2.17dB gain as compared with the FC and 1.49dB as compared with the CA_B for the *Foreman* sequence at the QP and PLR are 26 and 5%, respectively.

Fig. 4 shows the PSNR of each frame on the *Table tennis* and *Football* sequences at 5% PLR. The PSNR of the first frame after the loss one represents the error concealment ability of different EC methods, while that of the next following frames in the same GOP indicates the effects of the error propagation. As shown in Fig. 4, after concealing the lost frame, the effects of propagation errors on the next following frames are effectively suppressed in the proposed method as compared with the FC and CA_B to maintain high PSNR performance.

4. Conclusions

In this paper, we have proposed a new EC algorithm for the whole frame loss. Since all the received MVs of the neighboring frames are weighted and employed to estimate the MV of the missing block, the MV field obtained by the proposed algorithm is generally more realistic than that in the conventional methods to faithfully reconstruct the lost frame. Experimental results demonstrate that the proposed EC algorithm achieves the improved visual quality as well as the better PSNR performance than the conventional EC methods.

References

[1] M. Naghshineh and M. Willebeek-LeMair, "End-to-end QoS provisioning in multimedia wireless/mobile networks using an adaptive framework," *IEEE Commun. Mag.*, vol. 35, no. 11, pp. 72-81, Nov. 1997.

[2] J. Huang, R. Y. Yao, Y. Bai, and S. W. Wang, "Performance of a mixed-traffic CDMA2000 wireless network with scalable streaming video," *IEEE Trans. Circuits Syst. Video Techno.*, vol. 13, no. 10, pp. 973-981, Oct. 2003.

[3] J. Lu, "Signal processing for internet video streaming: A review," in *Proc. SPIE, Image Video Comm. Process.*, 2000.

[4] W.-Y. Kung, C.-S. Kim, and C.-C. J. Kuo, "Spatial and temporal error concealment techniques for video transmission over noisy channels," *IEEE Trans. Circuits Syst. Video Techno.*, vol. 16, no. 7, pp. 789-802, Jul. 2006.

[5] S. K. Bandyopadhyay, Z. Wu, P. Pandit, and J. M. Boyce, "Frame loss error concealment for H.264/AVC," 73rd MPEG meeting and 16th JVT meeting, Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-P072, Jul. 2005.

[6] 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, no. 2, pp. 316-329, Apr. 2005.

[7] P. Baccichet, D. Bagni, A. Chimienti, L. Pezzoni and F. S. Rovati, "Frame concealment for H.264/AVC decoders," *IEEE Trans. Consumer Electron.*, vol. 51, no. 1, pp. 227-233, Feb. 2005.

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