An Efficient Error Concealment for Whole Frame Loss with Violent Motion

Yu Chen[*†], Jun Wang[†‡], Zhiyi Yu[*†], Shaoju Wang[§] and Shilu Jiang[§]
Email: cyu2@andrew.cmu.edu, wangj387@mail.sysu.edu.cn, zhiyiyu@andrew.cmu.edu
[*] Department of Electrical and Computer Engineering, Carnegie Mellon University, USA
[†] School of Electronics and Information Technology, Sun Yat-sen University, China
[‡] SYSU-CMU Joint Research Institute, China
[§] School of Data and Computer Science,Sun yat-sen University, China

***Abstract*—In the video stream transmission, a packet loss would cause a whole frame loss, including pixel and motion information. Then seldom information could be utilized for lost frame (LF) reconstruction. Moreover, it is difficult to restore the LF in the scenario with chaotic and violent motion. In this paper, an efficient error concealment technique is proposed to resolve the whole frame loss problem by using priority-based block motion vector extrapolation (PBMVE). Most importantly, a Gauss filter (GF) with double zig-zag shifting path is designed for the smoothness of motion vector field (MVF) of LF. The experiment results demonstrate that the proposed technique PBMVE-GF outperforms other motion vector extrapolation methods in terms of objective and subjective quality, especially for the scenario with violent and chaotic motion.**

***Index Terms*—Video coding, error concealment, whole frame loss, violent motion, PBMVE-GF.**

I. Introduction

With the improvement of video compression efficiency, the video can be transferred in low bandwidth network. But at the same time, compressed video stream in the network is more sensitive to data loss. As a consequence, loss of a single bit may result in failing decoding the consecutive blocks in the frame. Besides, the congestion of network will cause the whole packet loss, which would lead to the whole frame loss. Moreover, the error of the LF will propagate to succeeding frames because video coding utilizes temporal prediction. In order to address these issues, many techniques have been reviewed in [1]. Among these techniques, error concealment technique is one of the most appropriate method to reconstruct LF and alleviate the error propagation [2].

In error concealment technique, there exit methods in temporal domain [3], spatial domain [4] or combinational domain [5]. If there is rightly received neighboring blocks of the missing block, block matching algorithm (BMA) [6] can be utilized to estimate the lost motion vector (MV). Moreover, in [7], [8], OBMA is proposed to obtain more accurate MV for MVF reconstruction of LF. Another strategy is to provide more MV candidates for missing block to select in [9]. Besides temporal methods above, spatial method of concealment is also developed to cope with frame reconstruction with changed scene or without motion vector candidates situation (e.g., I frame) [4], [10], [11], [12] .

With the development of compression efficiency, the network packet has the capacity to contain the whole frame. In another word, a packet loss can lead to the whole frame loss. Under this circumstance, the whole MVF of the LF needs to be reconstructed. The simplest and computation efficient method in [13] is to copy the reference frame (RF) to LF, which may produce a good result if the LF has little motion comparing to RF. In order to approximate the real motion of primitive LF, block-based motion vector extrapolation (BMVE) [3], [14] is proposed to combat the loss of a whole frame. This kind of method extrapolates MVs of blocks from the RF firstly, then the overlapped area is calculated between the damaged block and the motion extrapolation block. In BMVE, it selects the best estimated MV with the largest overlapped area to conceal the corrupted block by using general motion compensation. This technique is computational saving. But BMVE will cause block artifacts. To address the block artifacts, in [15], the method MV-OBEC with an overlapped window is proposed to smooth the edge effect. However, this window has limited effect to combat the whole frame loss in violent motion scenario. And in [16], pixel-based motion vector extrapolation (PMVE) is developed, which has higher perceptual quality and reserves more details. Later, in [17], [18], hybrid motion vector extrapolation (HMVE) is developed to combine the advantages of BMVE and PMVE to get more accurate MVF of LF. Besides, the multiframe extrapolation method like [19] can generate a better MV for LF. However, in some situations, there exist not enough MVs to extrapolate (e.g., the reference inter frame may contain lots of intra coded block, or there may be intra coded frame in reference frames). Meanwhile, the computational complexity of this method is much larger than one reference frame extrapolation.

Apart from using the MVF of previous frame, backward motion vector extrapolation is developed in [20], [21]. This method utilizes the MVF from succeeding frames and previous frames. This method will confront the situation where the succeeding frame of LF is also lost. In this situation, backward motion vector extrapolation does not work.

To the best of our knowledge, most of error concealment methods which utilize previous one frame information for whole frame loss could fail in MVF reconstruction to obtain acceptable visual quality in the scenario with violent motion.

An example in Fig.4(b,c,d,e) reflects this problem. This is because, in general, the MVF in the last received frame which is used for MVF extrapolation of current LF could not show its real motion.

Based on such motivation, PBMVE is proposed to restore MVF for LF with more accurate MV in this paper. Most importantly, a Gauss filter with double zig-zag shifting path is proposed to smooth the MVF of LF to get higher perceptual quality in violent and chaotic scenery. The experiments demonstrate the proposed method has better objective and subjective quality than conventional methods do.

The rest of the paper is organized as follows. In Section 2, the proposed error concealment technique PBMVE-GF and relating details are described. Section 3 shows subjective and objective experimental results with frame copy (FC), BMVE, MV-OBEC, PMVE, HMVE and the proposed technique. Section 4 is conclusion.

II. Proposed algorithm PBMVE-GF

The proposed algorithm performs in the following sequential steps: complete the whole MVF of RF by median filter, project MVF of RF to LF by PBMVE, complete MVF of reconstructed LF by colocate copy and use Gauss filter to smooth MVF of reconstructed LF. The whole process is illustrated in Fig. 1.

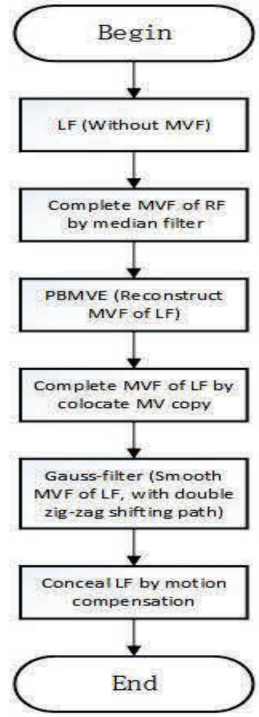

Fig. 1: Flowchart of Proposed Algorithm

Firstly, the MVF of RF F_{t-1} may contain intra blocks, so median filter is utilized to assign MV to intra blocks to complete the whole MVF of RF. Then, the MVF of RF F_{t-1} is fulfilled. The more information we get from frame F_{t-1}, the more accurate concealed frame F_t will be. Secondly, the block of RF F_{t-1} is projected into LF F_t with the help of MVF of RF. After PBMVE, the MVF of LF still has some blank blocks without any projected MVs. In order to fill these blank blocks, the collocated MVs in RF F_{t-1} are copied to LF F_t in the third step to utilize the temporal correlation. Then, as a very important step, a Gauss filter is used to smooth the MVF of LF, which can eliminate artifacts of concealed frame. During the filter process, a novel filter path named double zig-zag is proposed. At last, the motion compensation is utilized to conceal LF with the help of the reconstructed MVF. In the rest part of this section, two major steps PBMVE and GF will be described in detail.

A. PBMVE

In original BMVE [3], [14], the method first extrapolates MV of concealed frame from the last received frame, and then it calculates the overlapped areas between the lost block and the extrapolation block. Then the corrupted block would be assigned by a winning MV, which is the MV that has the largest overlapped area. Consequently, the MVF of LF is reconstructed.

In this paper, the priority is proposed to take the reliability of the MVF of RF into consideration. In the RF, there exist two kinds of MVs: primitive MV and estimated MV (e.g., MV in intra block without MV is regarded as an estimated one). The primitive MV has higher reliability than the estimated MV does. The estimated MVs of blocks are estimated by median filter among neighboring blocks. To distinguish estimated MVs and primitive MVs, a priority value $BP_{t-1}^{x,y}$ is assigned to block *(x,y)* in RF.

Then the blocks of RF F_{t-1} are projected into LF F_t, as shown in Fig. 2. The overlapped area $OA_{(x,y,i,j)}$ between the block *(x,y)* in RF and the block $B_t^{x,y}$ in LF can be calculated by:

$$OA_{x,y,i,j} = \sum_{p \in B_t^{i,j}} f_{t-1}^{x,y}(p) \tag{1}$$

$$f_{t-1}^{x,y}(p) = \begin{cases} 1, & p \in EB_{t-1}^{x,y} \\ 0, & p \notin EB_{t-1}^{x,y} \end{cases} \tag{2}$$

Where $EB_{t-1}^{x,y}$ denotes the extrapolated 4x4 block from RF in *(x,y)* to LF, and *p* stands for pixel.

After projection, the blocks in frame F_t may have multiple overlapped projected blocks. The MVs of blocks in LF would be assigned by one of their overlapped projected blocks. The selective results depend on the combination of priority $BP_{t-1}^{x,y}$ of the overlapped block and overlapped area $OA_{(x,y,i,j)}$.

The final priority of projected blocks can be calculated by

$$P_t^{i,j} = OA_{x,y,i,j} + BP_{t-1}^{x,y} \tag{3}$$

The chosen MV for block *(i,j)* in LF is selected from *M* candidates by

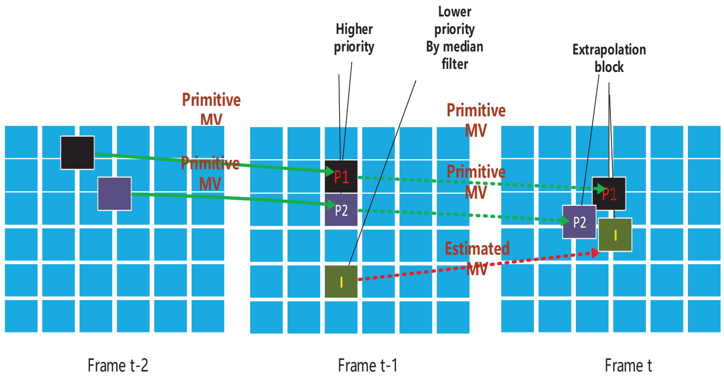

Fig. 2: Priority-based Block Motion Vector Extrapolation

$$MV_t^{i,j} = arg\ max\{P_{t,k}^{i,j}\}, k = 0, 1, 2, ..., M \quad (4)$$

In PBMVE, we take the reliability of estimated MVs and real MVs in RF into consideration to get more accurate MVs for LF.

B. Gauss Filter with Double Zig-Zag Shifting Path

After PBMVE, the MVF in the LF is estimated. However, the errors between the estimated MVF and original MVF may be still existed, which produce the edge artifacts in terms of video quality. A Gaussian filter is used for eliminating such artifacts.

In our algorithm, the Gauss filter is applied for a shifting window, whose size is 5x5 MVs. Every 4x4 block in pixel level has a MV. The Gauss kernel is as follows:

$$GF_{kernel} = \begin{bmatrix} 0.0285 & 0.0363 & 0.0393 & 0.0363 & 0.0285 \\ 0.0363 & 0.0461 & 0.0500 & 0.0461 & 0.0363 \\ 0.0393 & 0.0500 & 0.0541 & 0.0500 & 0.0393 \\ 0.0363 & 0.0461 & 0.0500 & 0.0461 & 0.0363 \\ 0.0285 & 0.0363 & 0.0393 & 0.0363 & 0.0285 \end{bmatrix} \quad (5)$$

The elements in the kernel follow Gauss distribution. As for σ of Gauss distribution, we tested different values of σ and found the optimal one is 2.5.

Therefore, the MV in block (i, j) is updated by following convolution:

$$MV_{t,GF}^{i,j} = \sum_{m=-2}^{2} \sum_{n=-2}^{2} GF_{kernel}(m, n) * MV_t^{i+m,j+n} \quad (6)$$

where (m, n) is the element index of Gauss kernel.

After all blocks' MVs were filtered, the MVF is updated. Note that the shifting window should be applied one by one for all blocks in the LF.

Moreover, the shifting window route is not normal as raster scanning. Here, a novel scan strategy is utilized. As showing in Fig. 3, the filer begins at the center of concealed frame, then the route diffuses in both sides with zig-zag shifting path. This kind of scan strategy can obtain better perceptual quality in the restored frame. This is because that compared with the MVs in the surrounding blocks, the MVs in the center blocks in a frame is more reliable in terms of real motion. Besides horizontal direction double zig-zag, we tested vertical direction double zig-zag shifting path, and we found the results between the horizontal one and vertical one are similar.

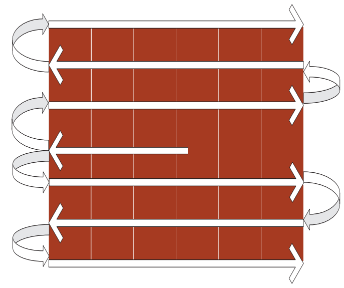

Fig. 3: Double Zig-Zag Scan Strategy

III. Experiments Results

In our experiments, 4 sequences including two violent motion FHD sequences ("basketball" and "tennis") and less violent motion HD sequences ("old town" and "raven") are chosen to demonstrate effectiveness of the proposed technique.

The platform we use is H.264/AVC JM reference software with baseline profile, and each sequence is encoded into one 60-GOP(group of picture) coded as IPPPPP...PPPPI.... QP is sustained in 28 for all test sequences. Furthermore, intra prediction is not constrained in inter frame and no other resilience feature is employed.

In the experiment, the proposed method is compared with FC[13], BMVE[14], MV-OBEC[15], PMVE[16] and HMVE[17] in objective and subjective quality. These methods utilize one reference frame, and they are more computational saving than multiframe extrapolation. As in experiments of [14], [16], [17], there is a constraint on intra encoded block in the original method. In comparing experiments, it is modified to be unconstrained on intra encoding in inter frame. As for blank MV block in reconstructed LF in [14], [16], [17], the missing MV is set with the left MV if left neighbor exists, otherwise the missing MV is set to zero.

TABLE I: average PSNR between different methods

Method \ Sequence	Basketball	Tennis	Old Town	Raven
FC	23.95	18.36	32.75	25.40
BMVE	26.60	23.40	32.09	27.23
MV-OBEC	26.57	22.19	32.10	27.50
PMVE	27.34	21.19	32.00	26.48
HMVE	25.90	21.63	32.05	26.85
PBMVE-GF	28.16	25.29	33.62	28.21

TABLE II: average SSIM between different methods

Method \ Sequence	Basketball	Tennis	Old Town	Raven
FC	0.9796	0.8180	0.8726	0.9645
BMVE	0.9787	0.8886	0.9563	0.9649
MV-OBEC	0.9672	0.8610	0.9566	0.9653
PMVE	0.9704	0.8541	0.9556	0.9584
HMVE	0.9689	0.8577	0.9562	0.9588
PBMVE-GF	0.9800	0.9145	0.9496	0.9648

In Fig.4, the proposed PBMVE and PBMVE-GF are compared with other methods in terms of subjective video quality. It can be observed that BMVE, PMVE, HMVE and MV-OBEC contain block artifacts, therefore it is hard to recognize the shape of the basketball player for audience. This shows that PBMVE and PBMVE-GF significantly outperform other methods in subjective quality. Furthermore, PBMVE-GF further enhances the quality of PBMVE.

In TABLE. I and TABLE. II, average PSNR and SSIM[22] for all frames are used for objective video quality comparison. The results demonstrate that the proposed method still achieves better performance than other methods. Moreover, as for chaotic motion sequences such as ”basketball” and ”tennis”, more significant enhancement of performance can be achieved than less chaotic motion sequences such as ”old town” and ”raven”. This indicates that our proposal works well in chaotic and violent motion scenarios, which is just the motivation of this paper.

IV. Conclusion

In this paper, PBMVE-GF is designed to conceal a whole frame loss especially in the chaotic and violent motion scenario. The obtained results demonstrate the potential of the proposed approach and show the benefits of restoring MVF of LF. Furthermore, the proposed method can be utilized for succeeding frames to get better perceptual quality and eliminate image blur for future work. All in all, the proposed algorithm can be effectively implemented to dispel the influence of the whole frame loss.

References

[1] Y. Wang and Q. F. Zhu, ”Error control and concealment for video communication: A review,” *Proc. IEEE,* vol.86, pp.974-997, May 1998.

[2] M. Usman, X. He, M. Xu and K. M. Lam, ”Survey of Error Concealment techniques: Research directions and open issues,” *Picture Coding Symposium (PCS), 2015.* IEEE, 2015.

[3] P. Qiang, T. W. Yang and C. Q. Zhu, ”Block-based temporal error concealment for video packet using motion vector extrapolation,” *Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on.* Vol. 1. IEEE, 2002.

[4] J. Koloda, J. Ostergaard and S. H.Jensen, ”Sequential error concealment for video/images by sparse linear prediction,” *Multimedia, IEEE Transactions on 15.4 (2013): 957-969.*

[5] P. Dheeraj and J. Patel, ”Hybrid spatio-temporal error concealment technique for image/video transmission,” *Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN), 2014 2nd International Conference on. IEEE,* 2014.

[6] W. M. Lam, A. R. Reibman and B. Liu, ”Recovery of lost or erroneously received motion vectors,” *Proc. ICASSP'93, Minneapolis,* pp. V417-V420, April 1993.

[7] Y. K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta and M. Gabbouj, ”The error concealment feature in the H. 26L test model,” *Image Processing. 2002. Proceedings. 2002 International Conference on.* Vol. 2. IEEE, 2002.

[8] Thaipanich, Tanaphol, P. H. Wu and CC. J. Kuo, ”Low-complexity video error concealment for mobile applications using OBMA,” *Consumer Electronics, IEEE Transactions on,* 54.2 (2008): 753-761.

[9] A. Kale and V. Hampiholi, ”Real time error concealment in H. 264 video decoder for embedded devices,” *Internet Multimedia Services Architecture and Applications, 2008. IMSAA 2008. 2nd International Conference on. IEEE,* 2008.

[10] J. Koloda, J. Ostergaard and S. H.Jensen, ”Sequential error concealment for video/images by weighted template matching,” *Data Compression Conference (DCC), 2012. IEEE,* 2012.

[11] J. Koloda, A. M. Peinado and V. Snchez, ”Kernel-Based MMSE Multimedia Signal Reconstruction and Its Application to Spatial Error Concealment,” *Multimedia, IEEE Transactions on 16.6 (2014): 1729-1738.*

[12] J. Koloda, V. Snchez and A. M. Peinado, ”Spatial error concealment based on edge visual clearness for image/video communication,” *Circuits, Systems, and Signal Processing 32.2 (2013): 815-824.*

[13] Z. Y. Wu and J. M. Boyce, ”An error concealment scheme for entire frame losses based on H. 264/AVC,” *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on. IEEE,* 2006.

[14] J. Zhou, Y. Bo and G. Hamid, ”Efficient motion vector interpolation for error concealment of H. 264/AVC,” *Broadcasting, IEEE Transactions on 57.1 (2011): 75-80.*

[15] K. Song, T. Chung, Y. Kim and Y. Oh, ”Error concealment of H. 264/AVC video frames for mobile video broadcasting,” *IEEE Transactions on Consumer Electronics 53.2 (2007): 704-711.*

[16] S. K. Bandyopadhyay, Z. Y, Wu, P. Pandit and J. M. Boyce, ”An error concealment scheme for entire frame losses for H. 264/AVC,” *Sarnoff Symposium, 2006 IEEE.* IEEE, 2006.

[17] B. Yan and H. Gharavi, ”A hybrid frame concealment algorithm for H. 264/AVC,” *Image Processing, IEEE Transactions* on 19.1 (2010): 98-107.

[18] P. C. Wang and C. S. Lin, ”Enhanced backward error concealment for H. 264/AVC videos on error-prone networks,” *Biometrics and Security Technologies (ISBAST), 2013 International Symposium on.*IEEE, 2013.

[19] C. Zhao, S. M, J. Zhang and W. Gao, ”A highly effective error concealment method for whole frame loss,” *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on.* IEEE, 2013.

[20] X. Fan, O. Au and J. Zhou, ”Backward error concealment of redundantly coded video, *in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP* 10., 2010, pp. 23062309.

[21] Y. Chen, K. Yu, J. Li and S. Li, ”An error concealment algorithm for entire frame loss in video transmission,” *Picture Coding Symposium.* 2004.

[22] A. Hore and D. Ziou, ”Image quality metrics: PSNR vs. SSIM,” *Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE,* 2010.

a. Original Error Free Frame

b. BMVE

c. MV-OBEC

d. PMVE

e. HMVE

f. PBMVE

g. PBMVE-GF

Fig. 4: Basketball Frame #2 Error Concealment with Different methods