Motion Compensated Error Concealment for HEVC Based on Block-Merging and Residual Energy

Yueh-Lun Chang*, Yuriy A. Reznik†, Zhifeng Chen† and Pamela C. Cosman*
*Dept. of Electrical & Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407
Email: {yuc018, pcosman}@ucsd.edu
†InterDigital Communications, Inc., San Diego, CA 92121
Email: {yuriy.reznik, zhifeng.chen}@interdigital.com

***Abstract*—We propose a motion-compensated error concealment method for HEVC and implement the method in reference software HM. The motion vector from the co-located block will be refined for motion compensation. Based on the reliability of these MVs, blocks will be merged and assigned with new MVs. The experimental results show both a substantial PSNR gain as well as an improvement in visual quality.**

I. Introduction

High Efficiency Video Coding (HEVC) is the latest video coding technology standard. As a joint project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG), the Joint Collaborative Team on Video Coding (JCT-VC) was formed for HEVC development, and the standard was approved and formally published in 2013. As a successor to H.264/AVC, HEVC promises to reduce the overall cost of delivering and storing video assets without decreasing the quality of experience delivered to the viewer, and two key issues have been a particular focus: increased video resolution and increased use of parallel processing architectures [1].

Packet loss happens due to various reasons when video is transmitted through the network: congestion, delay, limited bandwidth, etc. Though HEVC is designed to achieve multiple goals, including coding efficiency and ease of transport system integration and data loss resilience, it provides no guarantees of end-to-end reproduction quality and does not suggest any concealment when the bitstream is lossy. Various error concealment methods have been proposed to overcome packet loss in video transmission for prior standards [2]. In [3] and [4], the pixel values were recovered by spatially interpolating available pixels in neighboring macroblocks (MBs). The boundary matching algorithm (BMA) [5] and decoder motion vector estimation (DMVE) estimated lost motion vectors (MVs) based on taking the candidate MVs from its spatial and temporal neighbors, by minimizing a given distortion measure between the correctly received pixels. These spatial and temporal error concealment schemes were addressed to the coding characteristics of the MB-based codec by exploiting the correlation between a damaged MB and its adjacent ones in the same or previous frame. However, though HEVC is still under the block-based motion-compensation and transform coding structure, there is no MB in the codec. The macroblock concept has been extended by defining three types of variable size unit: Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU).

Starting from the largest CU (LCU), each CU allows recursive quadtree splitting into multiple sub-CUs, for sizes from 64×64 (CU depth=0) to 8×8 (CU depth=3). Each sub-CU can be further split into multiple PUs. PUs are the basic unit for motion prediction. After PU segmentation, a proper size of TU is determined for residual coding. Two encoding modes are supported in HEVC: intra- and inter-picture prediction. The mode is specified at the CU level, meaning all the PUs in a CU will be predicted under the same mode. In HEVC, a slice is the data structure that can be delivered and decoded independently, and it can either be an entire frame or a region of a frame. Slices are a sequence of LCUs, while they are composed of MBs in the prior standards. LCUs are usually set to be the size of 64×64, which is sixteen times larger than a 16×16 MB in the prior standards, thus slice losses from a HEVC bitstream will involve a larger area of a frame, although in some applications HEVC may be used on very high resolution data where the block is not a large portion of the frame. Under this condition, many of the prior error concealment methods would not be applicable, since they were usually designed for smaller lost blocks and often took the nearby correctly received pixels as reference. Because a loss in HEVC contains at least one LCU, the large lost block makes most of the lost pixels distant from the correctly received pixel border, thus the distortion measure between the block edge no longer serves as a good criterion for recovery.

Few studies have been done regarding error concealment in HEVC. In [6], a motion vector extrapolation based method was proposed for whole frame loss. A MV correlation from the co-located LCU was calculated for deciding whether to divide a large block into smaller ones or not. However, in block-based motion estimation at the encoder, the motion vector field is generated by minimizing the energy of prediction residuals and the rate-distortion cost, which may make those estimated MVs fail to represent the true motion [7]–[9], thus it could be improper to take every MV from the co-located LCU. In addition, a loss is not necessarily a whole frame loss and could be partial, which we call a slice loss in this paper. Spatial misalignment is more likely to happen and degrades the video quality.

Currently in HEVC reference software HM [10], only frame level concealment is implemented, where pixel copy from the previous frame is used. A slice loss is not yet detected and concealed. Among all types of error concealment, motion trajectory reuse from the co-located CTUs will be a relevant approach since the coding structure of HM preserves the information from the adjacent and co-located LCUs. In this paper, we propose a motion-compensated error concealment that can preserve the edge and object structure information without involving motion estimation or object detection at the decoder. First, residual energy of each block is analyzed to determine the reliability of each MV. Instead of refining the motion field by further partitioning each block into smaller blocks, we merge adjacent blocks that have unreliable MVs into a larger region. The merged block is assigned one single motion vector. Since the blocks with unreliable motion vectors are concealed using the same motion vector, the edges and the structure of the objects can be kept.

The paper is organized as follows: In Section II, we present the proposed algorithm in detail. The experimental results are presented in Section III. Section IV summarizes our conclusions.

II. PROPOSED METHOD

In this section, we propose a motion-compensated error concealment scheme based on the classification map of residual energy associated with each motion vector, and based on a block-merging algorithm.

Although network applications are one of the targets of HEVC, HEVC has not yet addressed transmission in networks other than to mandate byte stream compliance with Annex B of H.264/AVC. In [11], a streaming framework is designed and implemented. However, it used pre-generated encoder trace files and receiver trace files to detect loss, which is not a very realistic approach. The current HM so far is not able to detect and conceal a slice loss. However, a syntax element in HEVC, slice_segment_address, specifies the address of the first LCU in the slice segment, in a coding tree block raster scan of a frame. By modifying HM and tracking this syntax, we could detect a slice loss without auxiliary files. If the correctly decoded LCU number in a frame is discontinuous with the next slice_segment_address value in the same frame, a slice loss is detected.

Fig. 1 shows the block diagram of the proposed method. Each lost LCU will be concealed sequentially.

A. CU and PU segmentation

For a lost LCU, its CU/PU partition information is lost. In [12], [13], the authors showed that much of the time, the CU depth is highly correlated with its co-located one, and so is PU segmentation. In the first step of the block diagram, the algorithm assumes the lost LCU has the same CU partition and PU segmentation as the co-located one in the reference frame.

B. Motion vector classification based on residual energy

As HEVC still uses block-based motion estimation, it is possible that MV is selected for reasons of rate-distortion efficiency rather than because if represents the true motion. Therefore, when a PU has high residual energy, we can reasonably argue that the MV may not be reliable for representing the true motion. In our proposed method, we assume that the motion of the lost CU will follow the same trajectory as its co-located one, so the motion vector field from the co-located

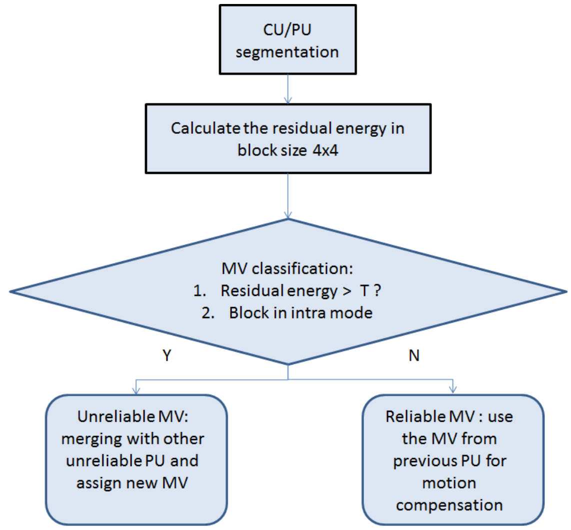

Fig. 1. Block diagram of the proposed method.

CU will be utilized for error concealment. In the second step in Fig. 1, to classify the candidate MVs from the co-located CU as reliable or not, the residual energy, E, of the co-located CU is calculated for each 4×4 block, $b_{m,n}$, by taking the sum of the absolute value of the luma reconstructed prediction error for each pixel.

$$E = \sum_{(i,j) \in b_{m,n}} | \, r_Y(i,j) \, |$$

Here, $r_Y(i,j)$ is the reconstructed residual signals of the Y component. If E is smaller than a threshold, the 4×4 block is classified as a reliable region, otherwise it is classified as unreliable. The threshold is selected from a heuristic search. In addition, an intra CU will also be categorized as unreliable. If any 4×4 block in a PU is unreliable, the whole PU is signaled as an unreliable PU, thus the corresponding MV of this PU is also unreliable. Fig. 2 demonstrates an example: an 8×8 CU is divided vertically into two 4×8 PUs, and there are two 4×4 residual blocks in each PU. The yellow denotes an unreliable block while blue stands for reliable ones. For the right PU, one of the residual blocks is unreliable, so the whole PU is considered unreliable.

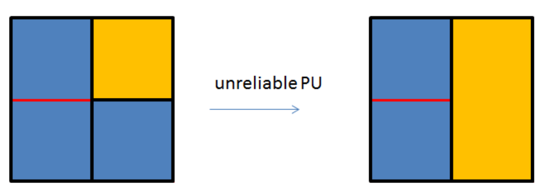

Fig. 2. Example of unreliable PU classification.

C. PU merging and MV reassignment

In the last step of Fig. 1, the unreliable PUs will be merged and reassigned with refined MV. The detail is described in this section.

When there is a group of adjacent PUs that have unreliable motion vectors, misalignment happens easily along the edges, and the shape of the objects usually could not be maintained. To keep the integrity of the object, it would be beneficial to group these units with one MV, so the structure would not be deformed. The merging process only happens between PUs in the same CU depth, so the merged PUs would not be too different in size. This reduces the blockiness effect. There is no merging at CU size 32×32 and 64×64 to refrain from losing too much detailed motion. Starting from the very top left point of a lost region, whenever we encounter an unreliable PU, we check its right, bottom, and bottom-right PUs. If any of these PUs are both unreliable and within the same CU level, they are merged into a larger piece. For each PU, it will only be merged once to keep the size of the group in a reasonable range. Fig. 3 depicts an example: Four 16×16 CUs contain five PUs. Three of the PUs are unreliable, so they are merged into one piece.

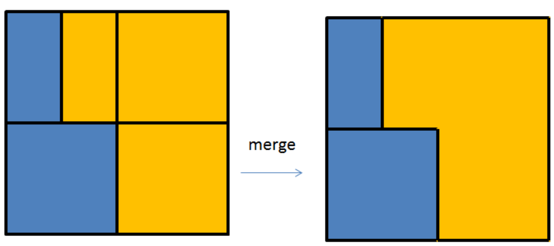

Fig. 3. Example of merging unreliable PUs.

After PU merging, the reliable MVs from its adjacent PUs are collected, and the average of these reliable MVs will be assigned. For a reliable PU, the MV from the co-located area will be used directly.

III. SIMULATION

In this section, we present experimental results to evaluate the proformance of the proposed method. The proposed method will be compared with two other schemes:

1) Pixel copy
2) Basic motion compensated error concealment (MCEC), where the MV from the co-located CU is applied directly with no refinement. For an intra CU, pixel copy is used.

Two video sequences, Soccer (720×480) and Drill (832×480) , are encoded by HM11.0 for 120 frames. The frame rate is 30 frame per second (fps) for Soccer and 50 fps for Drill; the intra period is every 20 frames with only P frames in between. The QP is 28 and each slice has a fixed 20 LCUs to simulate the loss of a region of a frame.

The proposed algorithm is implemented in HM10.0. We use the packet loss simulator developed by AHG14 to facilitate G.1050/TIA 921 packet loss simulations for HEVC [14]. The losses are randomly distributed in all the P frames. The packet loss rates (PLRs) of 1%, 3%, 5%, and 10% are tested, and each sequence is decoded for 100 random realizations.

Table I presents the PSNR performances averaged over all frames for different sequences with different PLRs. As shown in the table, the proposed method outperforms the copy and MCEC algorithm by up to 0.26 dB. Since the error propagation of HEVC is quite severe and the quality of succeeding frames degrades very fast for all types of the concealment methods, the gain of the proposed method is not much in terms of the PSNR over the whole sequence. However, if we only look at the erroneous frame itself, the PSNR gain is more prominent and the visual quality is also better. In order to evaluate the effectiveness of the concealment for the lossy frame solely, Table II presents the average PSNR of the first erroneous frame in each GOP, and hence the influence of the error propagation is excluded. As shown in the table, the proposed method yields higher PSNR than the copy and MCEC algorithm up to 1.3 dB. In both Tables I and II, MCEC performs the worst because the improper reuse of the co-located MVs makes the concealed area quite blocky. The copy method roughly maintain the shape of the object but fails at the boundary of the corrupted area.

Sequence	Method	Packet Loss Rate			
		1%	3%	5%	10%
Soccer	copy	28.63	24.58	22.07	19.84
	MCEC	28.56	24.51	21.99	19.77
	proposed	28.84	24.74	22.28	20.02
Drill	copy	30.93	27.05	25.32	23.13
	MCEC	30.86	26.97	25.25	23.06
	proposed	31.17	27.31	25.58	23.34

TABLE I. COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ALL FRAMES FOR DIFFERENT PLRS

Sequence	Total number of the first erroneous frame in a GOP from all realizations	Method	PSNR(dB)
Soccer	2114	copy	28.84
		MCEC	28.71
		proposed	29.82
Drill	2184	copy	30.98
		MCEC	30.81
		proposed	32.11

TABLE II. COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ONLY THE FIRST ERRONEOUS FRAME IN A GOP FOR ALL PLRS

The visual comparisons are presented in Figs. 4 and 5, demonstrating examples of frame 87 of the Soccer sequence and frame 24 of the Drill sequence, where (a) is the original compressed frame without loss, (b) is the corrupted frame and (c)-(e) are the frames concealed using copy, MCEC, and the proposed algorithm respectively. In both figures, the visual quality of our method is significantly better than the others. Our method successfully preserves the shape of the moving object with smooth edges while the copy and MCEC methods fail to maintain the structure of the moving object with blockiness and deformed boundary. Comparing with two other methods, the PSNR value of the proposed method is up to 2.9 dB higher for Soccer and 4.5 dB higher for Drill in these cases.

IV. CONCLUSION

We propose a motion-compensated error concealment method for HEVC and implement the method in reference software HM. Based on the received the residual information, the motion vector reliability is analyzed and classified. The CU with unreliable MVs will be merged and assigned with one new MV to maintain the structure of the moving object and edge information. Our method is effective yet simple without doing edge or object detection explicitly. It is noteworthy that simplistic use of MCEC actually performs worse than simple slice copy. However the proposed version of MCEC provides a significant PSNR improvement over both simplistic MCEC and over slice copy for the frame in which the loss occurs, with up to 1.3dB improvement.

REFERENCES

[1] G.J. Han, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," 2012.

[2] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 974–997, 1998.

[3] J.-W. Suh and Y.-S. Ho, "Error concealment based on directional interpolation," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 3, pp. 295–302, 1997.

[4] A. Raman and M. Babu, "A low complexity error concealment scheme for MPEG-4 coded video sequences," in *IEEE Symposium on Multimedia Communications and Signal Processing*. Citeseer, 2001.

[5] W.-M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993. ICASSP-93.,*. IEEE, 1993, vol. 5, pp. 417–420.

[6] C. Liu, R. Ma, and Z. Zhang, "Error concealment for whole frame loss in HEVC," in *Advances on Digital Television and Wireless Multimedia Communications*, pp. 271–277. Springer, 2012.

[7] H. Sasai, S. Kondo, and S. Kadono, "Frame-rate up-conversion using reliable analysis of transmitted motion information," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04).* IEEE, 2004, vol. 5, pp. 257–260.

[8] A.-M. Huang and T. Nguyen, "Motion vector processing based on residual energy information for motion compensated frame interpolation," in *IEEE International Conference on Image Processing, 2006*. IEEE, 2006, pp. 2721–2724.

[9] A.-M. Huang and T. Nguyen, "A novel multi-stage motion vector processing method for motion compensated frame interpolation," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*. IEEE, 2007, vol. 5, pp. V–389.

[10] "reference software of HEVC (HM)," https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/.

[11] J. Nightingale, Q. Wang, and C. Grecos, "Hevstream: a framework for streaming and evaluation of high efficiency video coding (HEVC) content in loss-prone networks," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 404–412, 2012.

[12] J. Leng, L. Sun, T. Ikenaga, and S. Sakaida, "Content based hierarchical fast coding unit decision algorithm for HEVC," in *International Conference on Multimedia and Signal Processing (CMSP)*. IEEE, 2011, vol. 1, pp. 56–59.

[13] H.-Y. Chen, "An efficient fast CU depth and PU mode decision algorithm for HEVC," National Central University, 2011.

[14] ITU-T SG16 WP3/ ISO/IEC JTC1/SC29/WG11 JCTVC-H0072, "Nal unit loss software," 2012.

(a) original

(b) corrupted

(c) copy

(d) MCEC

(d) proposed

Fig. 4. Reconstructed results of frame 87 of the Soccer sequence: (a) original frame, (b) corrupted frame, (c) concealed by copy, PSNR: 22.58dB (d) concealed by MCEC, PSNR: 22.57dB (e) concealed by the proposed method, PSNR: 27.16dB.

(a) original

(b) corrupted

(c) copy

(d) MCEC

(e) proposed

Fig. 5. Reconstructed results of frame 24 of the Drill sequence: (a) original frame, (b) corrupted frame, (c) concealed by copy, PSNR: 30.62dB (d) concealed by MCEC, PSNR: 30.54dB (e) concealed by the proposed method, PSNR: 33.58dB.