TEMPORAL ERROR CONCEALMENT ALGORITHM BASED ON ADAPTIVE SEACH RANGE AND MULTI-SIDE BOUNDARY INFORMATION FOR H.264/AVC

Myounghoon Kim, Soonhong Jung, Beumjoo Kang, Sanghoon Sull

School of Electrical Engineering
Korea University
Seoul, Korea
E-mail: mhkim@mpeg.korea.ac.kr

ABSTRACT

A compressed video stream is very sensitive to transmission errors that may severely degrade the reconstructed image. Therefore, error resilience is an essential problem in video communications. In this paper, we propose novel temporal error concealment techniques for recovering lost or erroneously received macroblock (MB). To reduce the computational complexity, the proposed method adaptively determines the search range for each lost MB to find best matched block in the previous frame. And the original corrupted MB split into for 8×8 sub-MBs, and estimates motion vector (MV) of each sub-MB using its boundary information. Then the estimated MVs are utilized to reconstruct the damaged MB. In simulation results, the proposed method shows better performance than conventional methods in both aspects of PSNR.

Keywords: block-matching, error concealment, block loss recovery and H.264/AVC

1. INTRODUCTION

For video transmission over bandwidth-limited networks, digital video compression standards such as H.264/AVC [1] use a combination of various advanced features to improve the compression. However, compressed video streams are vulnerable to transmission errors due to a packet loss and a delay that are almost inevitable in video transmission over wireless channels and the Internet. In these situations, error-free delivery of data packets can only be achieved by requesting retransmission [2] of lost or damaged packet through feedback channel and inserting redundant information into compressed video stream [3], [4]. They are efficient in stopping error propagation but may not be acceptable in some interactive application, due to the extra delay and more computation.

Recently, approaches for temporal error concealment (TEC) at the decoder side have been proposed, whose main objective is to estimate the lost information usually occurred due to transmission errors by utilizing the correctly received information without modifying source and channel coding schemes. They are hence suitable for a wide range of applications. The simple way to conceal the damaged MB is the temporal replacement [5] method, which uses zero motion vectors and copy image samples of the same MB within previous frame into the lost area, but performs worse when the motion is fast. In [6], the boundary matching algorithm (BMA) chooses one MV among surrounding error-free received MVs based on a spatial smoothes constraint. The selective motion vector matching (SMVM) algorithm in [7] improved the performance of BMA by using status flags of neighboring pixels and extending a set of the candidate MVs. These methods work well for video sequences with moderate motion. However, it often generates a false motion vector for the video sequence with scene change or complex motion. In [8], using the temporal matching principle, decoder motion vector estimation (DMVE) algorithm is looking for best match in a previous frame using neighboring pixels of the damaged MB in the current frame. However, the main drawback of DMVE is that blurred image areas and block artifacts are visible when the damaged MB contains non-translational motion such as rotation and multi-directional movements. A temporal recursive block-matching (RBM) algorithm [9] uses a two-step block-matching principle and additionally recursion steps for temporal EC. It results the high image quality, however, it requires a lot of computations to process its recursive step.

In this paper, we propose novel temporal error concealment algorithm based on block matching principles accurately estimates the MVs of any lost MBs. Firstly, it determines the search range to estimate the MV of damaged MB from reference frame. Secondly, best matching sub-MBs are searched in reference frame using boundary pixels of the corrupted MB in current frame. The rest of paper is organized as follow. In the next section, we describe the proposed method in detail and section 3 presents the simulation results. Finally section 4 concludes the paper.

2. PROPOSED METHOD

Several effective techniques [8], [9] for temporal error concealment have been provided. However, these methods have several drawbacks, which use fixed search range (SR) and do not consider whether lost MB is in the region of moving object boundary or not. Thus, these methods are not efficient when the lost MB is area in which little motion occurs. Also, blocking artifacts are presented in concealed results when damaged MB contains multi-directional movements or has more than two objects inside it, since they consider only one MV for damaged MB. In order to solve above problems, we propose a new algorithm to provide adaptive search range decision and temporal error concealment algorithm based on block matching with multi-side boundary information.

1.1 Adaptive search range decision

A general way of determining search range is to set a fixed value. This fixed value should be large enough to assure search window includes the best MV candidate. However, a fixed, large enough search range will put the error concealment process to continuous high computation consuming state. Therefore, it is necessary to adjust the search range at each block since motion in a single frame will vary from block to block and it will not be constant from scene to scene in a video sequence. A good way to solve this problem is to adaptively predict the search range when large search window is unnecessary.

The proposed method uses spatial correlation to decide search range, which aims to accurately determine the search range of lost MB using correctly received information at the decoder. For adaptive search range algorithm, search range is adaptively changed from one damaged MB to another. A search range ($\pm SR_x, \pm SR_y$) is directly calculated on the basis of a function of any adjacent motion vectors. In order to adjust the search range for damaged MB, firstly we construct the MV set as follows

$$S = \{d_T, d_B, d_L, d_R, d_C\}. \quad (1)$$

where d_T, d_B, d_L, d_R, are four neighboring MVs of correctly decoded MBs and d_C is the collocated MV in Fig. 1.

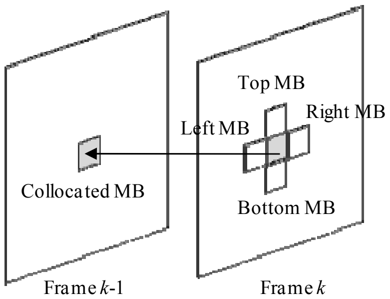


Fig. 1. Neighboring and Collocated MVs to Conceal the MB in Gray.

The d_L and d_R may not be available when the lost MBs are consecutive instead of isolated. Since the MV has two components, we denote the x and y parts of the MV set S by S_x and S_y respectively.

$$\begin{cases} S_x = \{d_x \mid d \in S\} \\ S_y = \{d_y \mid d \in S\} \end{cases}, \quad (2)$$

where d_x and d_y denote the x and y components of MV, respectively. Then we find the maximal components of sets S_x and S_y denoted by $\max(S_x)$ and $\max(S_y)$, respectively.

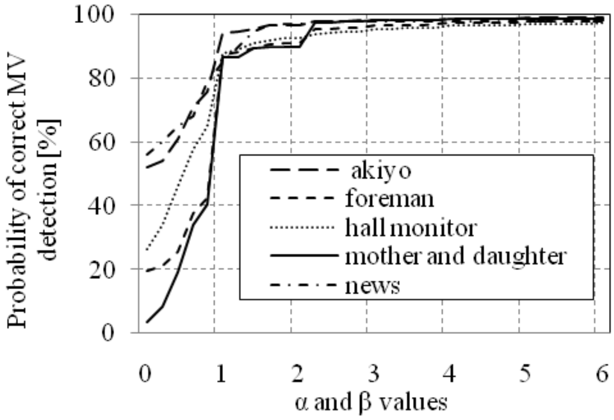


Fig. 2. Probability of correct MV detection.

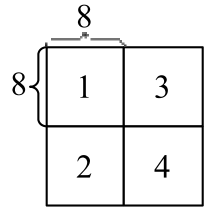


Fig. 3. Sub-MB partitions.

We use the function defined as

$$SR_x = \begin{cases} 16 & \text{if } \max(S_x) > 6, \\ 3 & \text{if } \max(S_x) = 0, \\ \alpha \times \max(S_x) & \text{otherwise,} \end{cases}$$
$$SR_y = \begin{cases} 16 & \text{if } \max(S_y) > 6, \\ 3 & \text{if } \max(S_y) = 0, \\ \beta \times \max(S_y) & \text{otherwise,} \end{cases} \quad (3)$$

where α and β are scale factors. According to α and β values, the performance of computational complexity to find MVs and obtain optimal MV for damaged MB are determined. If α and β values are small, search range will also small, so less computational complexity will be required. In this case, however, it will decrease the possibility of obtaining optimal MVs in scenes with large motions. Whereas, when α and β values are large, search range will also be large, it will be possible to obtain optimal MVs in scenes with large motion, but it requires much computational complexity. That is, it is necessary to decide optimal α and β values for a desired balance between search range and complexity. In order to determine optimal α and β values, we examined the probability that correct MVs might be included in search range determined by various α and β values. The correct MVs would be obtained by a DMVE [8] with fixed a $\pm$ 16-pixels search range. In this experiment, we utilize various test sequences. Fig. 2 shows the result of our experiments. As shown in this Fig., we use $\alpha=\beta=3.0$ since the probability of correct MV detection does hardly change when α and β are greater than 3.0.

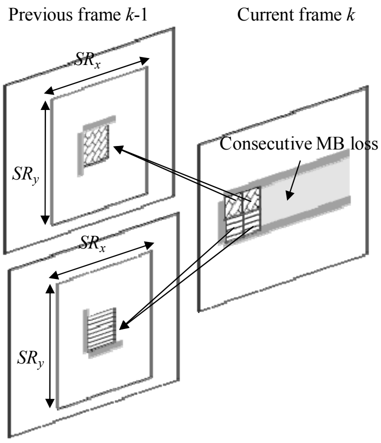
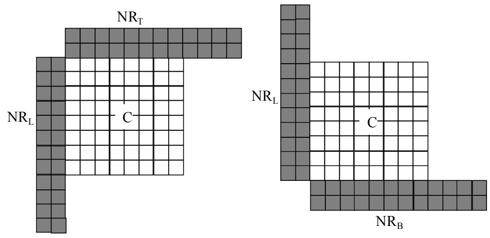


Fig. 4. Schematic illustration



(a) sub-MB 1 and 3 (b) sub-MB 2 and 4

Fig. 5. Damaged sub-MB and its neighboring regions.

1.2 Sub-MB matching EC Algorithm

In this section, we propose a method based on block matching principles, which is applied in 8×8 sub-MB level instead of MB level and each sub-MB can be recovering using temporal redundancy. Fig. 3 shows sub-MB partitions and recovering order of sub-MBs within a lost MB. As shown in Fig. 4, the proposed method utilizes the information available in the neighboring regions to conceal the error MB. The neighboring regions (NR_T, NR_B and NR_L) of corrupted sub-MB are depicted as dark-gray regions in Fig. 5. In this Fig, in order to perform temporal matching for all neighboring candidates in a search range in a previous frame F_{k-1}, the NR_T and NR_L neighboring regions on the top and left of a sub MB 1 and 3 and the NR_B and NR_L neighboring regions on the bottom and left of a sub MB 2 and 4 are used, respectively. The best matched block for concealing sub-MB is determined by minimizing mean absolute difference (*MAD*), which is used to conceal the corrupted sub-MB. *MAD* is calculated as

$$MAD(mv_x, mv_y) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} | NR(x, y, k) - NR(x - mv_x, y - mv_y, k-1) | \quad (4)$$

where *M* and *N* are the size of neighboring region, *NR*(*x*, *y*, *k*) is the boundary pixel at position (*x*, *y*) in current frame *k* and $NR(x+mv_x, y+mv_y, k\text{-}1)$ is the motion compensated pixel as position (*x*, *y*) in *k*-1 for MV(mv_x, mv_y). The size of top and bottom neighboring regions are 12×8 and left neighboring region is 8×12. To decide the MV of lost sub-MB, we calculate each *MAD* for neighboring regions. After that, the weighed *MADs* (*WMAD*) of the upper-half MB (sub-MB 1 and 3) and lower-half MB (sub-MB 2 and 4) are calculate as

$$\begin{aligned} WMAD_{Upper}(mv_x, mv_y) = &\ w(NR_T) \times MAD_{TOP}(mv_x, mv_y) \\ &+ w(NR_L) \times MAD_{LEFT}(mv_x, mv_y), \\ WMAD_{Lower}(mv_x, mv_y) = &\ w(NR_B) \times MAD_{BOTTOM}(mv_x, mv_y) \\ &+ w(NR_L) \times MAD_{LEFT}(mv_x, mv_y), \end{aligned} \quad (5)$$

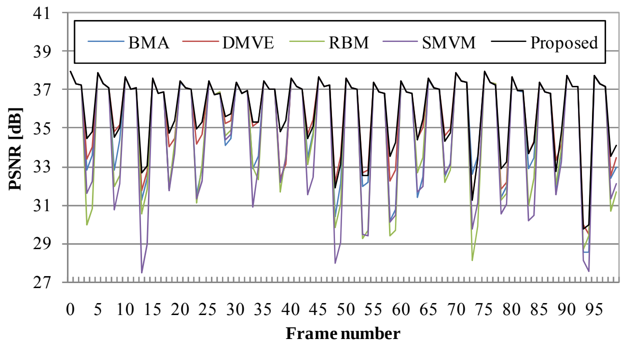
where MAD_T, MAD_B and MAD_L represent the *MAD* for top, bottom and left neighboring regions, respectively.

The weighting function *w*(*NR*) in (5) indicates whether the region *NR* is correctly received (*w*(*NR*)=1) or lost (*w*(*NR*)=0). The concealed regions are lower weighted that the correctly received pixels by setting *w*(*NR*) between 0.2 and 0.5. The best matched block has the minimum *WMAD*, the estimated MV is given by

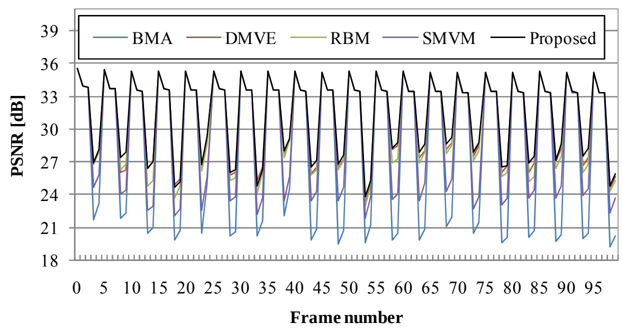
$$(\widetilde{mv}_x, \widetilde{mv}_y) = \arg \min_{mv_x, mv_y} WMAD(mv_x, mv_y) \quad (6)$$

Once the optimal MV of the corrupted sub-MB is obtained, the concealed pixel *f*(*x*, *y*) is obtained as follow:

$$f(x, y) = f(x + \widetilde{mv}_x, y + \widetilde{mv}_y, k-1) \quad (7)$$



(a)



(b)

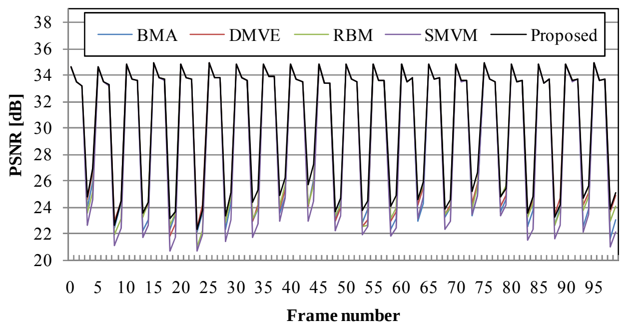


Fig. 6 The average PSNR performances: (a) "Foreman," (b) "Flower garden," (c) "Football."

Table II.
Test result of the computational complexity.

Average computation time (ms / frame)				
BMA	DMVE	RBM	SMVM	**Proposed**
1	199	1156	3	**29**

2. EXPERIMENTAL RESULTS

Our method was implemented into H.264/AVC reference software JM 9.8 [10], and whose performance was evaluated with three test sequences (*Foreman*, *Flower garden*, and *Football*). The frame structure in our experiment is IPPPPI.... Slice group map type is "Interleaved" such that the lost MBs can be concealed from the remaining slice groups. In our experiments, quantization parameter is set 28 for each sequence. The number of MBs is calculated according to the MB loss rate (MBLR) given by

$$MBLR = \frac{\#\ of\ corrupted\ MB}{\#\ of\ total\ MBs} \tag{8}$$

Four conventional methods, BMA [6], SMVM [7], DVME [8], and RBM [9], are compared with the proposed method. The BMA and DMVE are well-known and widely used temporal EC method utilizing the matching-based approach.

Figs. 7 to 9 compare the subjective quality of the reconstructed images concealed by the different algorithms. Fig. 7 shows the original and corrupted frames of *Foreman* and *Flower garden* sequences, respectively. The block lines in Fig 7 (b) and (d) present the corrupted MBs rows. Obviously, the proposed method is superior to other methods due to the more accurate block matching and fewer mismatches. As shown in difference images bottom row of Fig. 8, we see that the concealed images by conventional methods considerably distinguish from the original one especially around face, hat and wall of the building. Thus, reconstructed image contain the blocking artifacts. Other comparison results of the subjective visual quality are given in Fig. 9. The border of the tree and flowers are concealed with mismatching blocks by conventional method, the blurred areas can be seen at the border of the tree and the flowers behind the tree. Compared with these results, the concealed image obtained by the proposed method has the best visual quality, as shown in Figs. 8 (d) and 9 (d). Next Fig. 6 plots the average PSNR of each frame. It is assumed that the frames coded by intra mode are not corrupted and every fourth inter frames are corrupted. Note that, for most frames, the proposed method provides the highest PSNR performances. Table II shows a comparison of computational times for these methods. Our method needs less computational time compared to DMVE and RBM.

3. CONCLUSION

Error concealment is applied to reduce visible distortion using temporal redundancy left in the received information. In this work, we proposed new algorithms for error concealment in the video coding standard H.264. For temporal error concealment, we have point out the problems of conventional methods those require a lot of computation and produce the degradation of the visual quality of reconstructed image. To solve these problems, we adaptively determine search range and estimates motion vector of lost sub-MBs in search range. Experimental results demonstrate that the proposed technique achieves the better PSNR performance of the error concealment as compared with the other conventional methods.

4. REFERENCES

[1] ISO/IEC JTC1/SC29/WG11 (MPEG), "Coding of audio-visual objects – Part 10: Advanced Video Coding," *International Standard 14496-10*, ISO/IEC, 2004.

[2] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 87, no. 10, pp. 1707–1723, Oct.1999.

[3] J. Hagenauer, and T. Stockhammer, "Channel coding and transmission aspects for wireless multimedia," *Proc. IEEE*, vol. 87, no. 10, pp. 1764–1777, Oct. 1999.

[4] Y. Wang, M. Orchard, V. Vaishampayan, and A. R. Reibman, Multiple description coding using pairwise correlating transform. *IEEE Trans. on Image Processing*, 10, 3 (March 2001), 351-366.

[5] J. W. Suh and Y. S. Hu, "Error concealment based on directional interpolation," *IEEE Trans. Consumer Electronics*., vol. 43, no. 4, pp. 295–302, Aug. 1997.

[6] W.-M Lam, A. S. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing* (*ICASSP*), Minneapolis, April 1993, pp. V417–V420.

[7] B. Yan and K. W. Ng, "A novel selective motion vector matching algorithm for error concealment in MPEG-4 video transmission over error-prone channels," *IEEE Trans. on Consumer Electronics*, 49, 4 (Nov. 2003), 1416-1423.

[8] J. Zhang, J. F. Arnold, and M. R. Frater, "A cell-loss concealment technique for MPEG-2 coded video," *IEEE Trans. Circuits Syst. Video Technol*., vol. 10, no. 4, pp. 659–665, June 2000.

[9] M.-J. Chen, C.-S. Chen, and M.-C. Chi, "Temporal error concealment algorithm by recursive block-matching principle," *IEEE Trans. Circuits Syst. Video Technol*., vol. 15, no. 11, pp. 1385–1393, Nov. 2005.

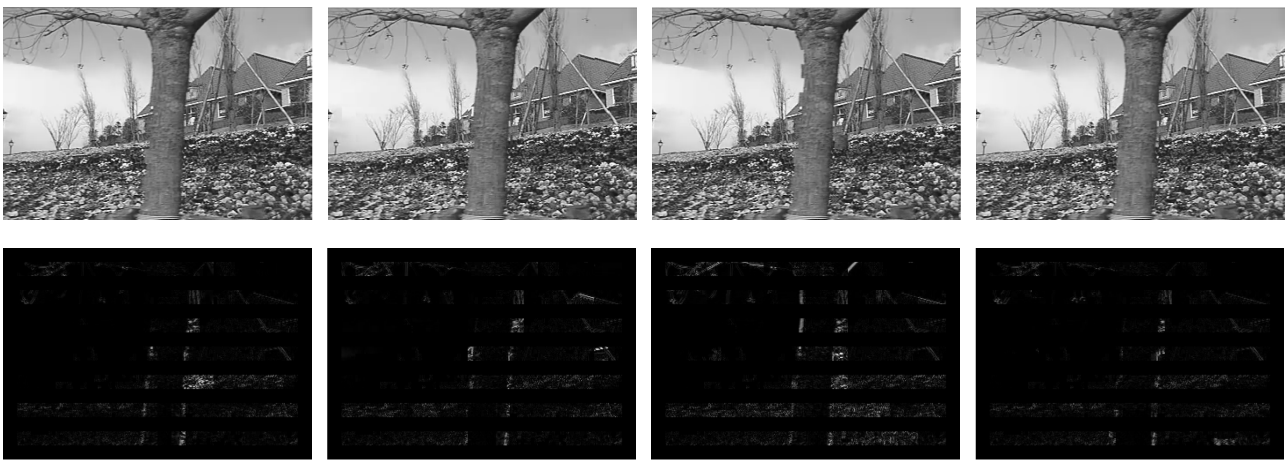
[10] JVT Model reference software from http://iphome.hhi.de/suehring/tml/download/oldjm/jm98.zip.

(a) (b) (c) (d)

Fig. 7. The frames of the “Foreman” and “Flower garden” sequences: (a) original image (36.93dB), (b) error image (7.45dB), (c) original image (33.74dB), (d) error image (8.29dB).

(a) (b) (c) (d)

Fig. 8. In the top row there are the reconstructed images with different concealment method, whereas the bottom row reports differences vs. correctly decoded images. (a) DMVE (33.36 dB) (d) RBM (29.93 dB) (e) SMVM (31.61 dB) (f) Proposed (34.47 dB)



(a) (b) (c) (d)

Fig. 9. In the top row there are the reconstructed images with different concealment method, whereas the bottom row reports differences vs. correctly decoded images. (a) DMVE (26.11 dB) (d) RBM (26.42 dB) (e) SMVM (24.10 dB) (f) Proposed (27.45 dB)