

Error Concealment Algorithm of Whole Frame Loss Using Inter-View Correlation for Multi-View Video Decoding

Yuan-Kai Kuan^a, Gwo-Long Li^b, Mei-Juan Chen^a, and Pin-Cheng Huang^a

Dept. of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan^a

E-mail: m9923037@ems.ndhu.edu.tw, cmj@mail.ndhu.edu.tw, m9923026@mail.ndhu.edu.tw

Industrial Technology Research Institute, Hsinchu, Taiwan^b

E-mail: gli@itri.org.tw

Abstract — In this paper, an error concealment algorithm for whole frame loss for multi-view video coding (MVC) is proposed. In our proposal, the disparity vectors similarity between continuous frames and the motion vectors similarity between adjacent views are considered and analyzed first. Based on analytical results, a joint sum of absolute difference (SAD) minimization approach is proposed to obtain the best motion vector to conceal the block in the error frame. By jointly considering the difference between neighboring views and frames, the error concealing results can be improved significantly. Simulation results demonstrate that our proposed algorithm can achieve higher PSNR performance as well as significant subjective quality improvement than previous work.

I. INTRODUCTION

As the multimedia technology advances in recent years, the applications of three-dimensional (3D) television and free viewpoint video (FVV) become attractive. To support multi-view video coding, the standard has been proposed in [1][2]. In multi-view video coding, the motion compensated prediction (MCP) technology in H.264/AVC[3][4] has been adopted to reduce the temporal redundancy with incorporating the disparity compensated prediction (DCP) technology to eliminate inter-view redundancy. The main concept of DCP technology is to use inter-view correlation to search the disparity vector (DV) so that the bit-rate of multi-view video sequences can be reduced.

In the error prone network environment, the packet error or the packet loss may occur very frequently due to the unpredicted interruption of noise sources and thus leads to the decline of the received video quality. Therefore, the error recovering mechanism becomes the important research issue. To deal with error problem for multi-view application, many works have been proposed [5]-[9]. Work [5] uses the intra-view difference, inter-view correlation and difference of the inter-view disparity vector projections on the neighbor views to conceal the error frames. Work [6] adopts the disparity estimation from the neighbor views to find the location of every pixel in previous frame. Afterwards, the corresponding motion vectors are considered to conceal the error frames. Work [7] compares the sum of absolute

difference (SAD) between the previous two frames and the SAD between adjacent views of the previous frame to achieve error concealment.

To conceal the error frame for multi-view coding (MVC), we propose an algorithm which applies predictive compensation approach with considering inter-view correlation to conceal error frame. By using the DV in the previous frame as the reference prediction DVs, the MVs inside the block referred by the reference DVs are collected for being the candidates for our error concealment process. Finally, once the candidate MVs have been successfully collected, the candidate MV with smallest joint SAD would be chosen as the best MV to conceal error block.

The organization of this paper is described below. In section II, the inter-view correlation in MVC is briefly described, and the proposed correlation-based motion vector prediction error concealment algorithm will be introduced in detail as well. Section III gives some simulation results to show the efficiency of our proposed algorithm. Finally, the conclusion is presented in Section IV.

II. PROPOSED ALGORITHM

A. Properties of Multi-View Video Coding

To create multi-view video sources, the cameras are usually placed linearly to capture the scene. In this case, the motion vectors of different views are very similar to each other. However, the differences are the inter-view disparity vectors caused by the distance between cameras as Fig.1 shown.

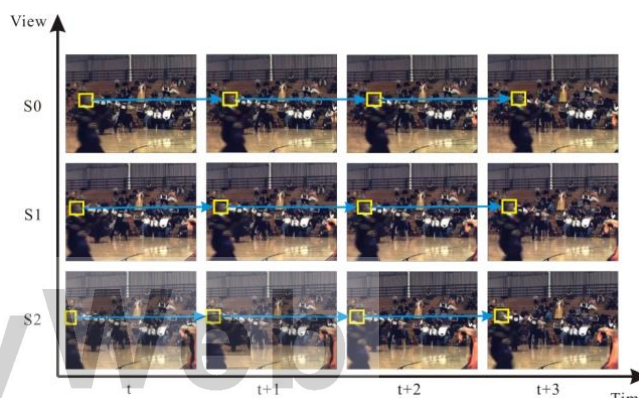


Fig.1 Frame similarity between views.

This work is supported by the National Science Council, Taiwan, under grant 100-2628-E-259-002-MY3.

Fig.2 gives an example to illustrate the relationship of object movement between views. From this figure, we can observe that the correlations of both motion vectors and disparity vectors are very high. Conceptually, the motion vectors and disparity vectors form a parallelogram-like motion relationship.

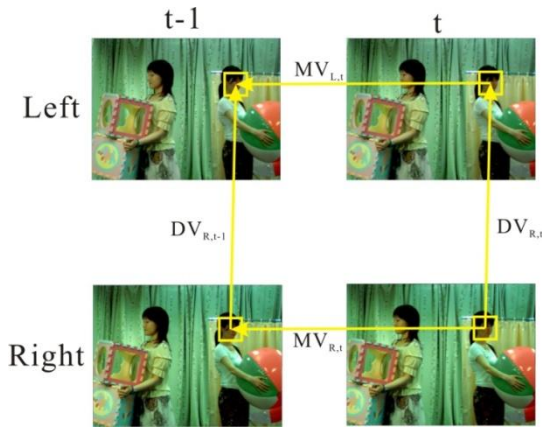


Fig.2 The in-view and cross-view parallelogram-like motion relationship. ($DV_{R,t-1} \approx DV_{R,t}$, $MV_{R,t} \approx MV_{L,t}$).

B. Proposed Error Concealment Algorithm

From the previous section, we can observe that the motion vectors between views have very high correlation. Based on such relationship, we propose an error concealment algorithm to deal with error problem by using correlation between views for MVC. It should be noted that the proposed algorithm works in the two-view coding pattern in MVC and the whole frame of right view is lost, as shown in Fig. 3. However, the proposed algorithm can be extended to several views application accordingly.

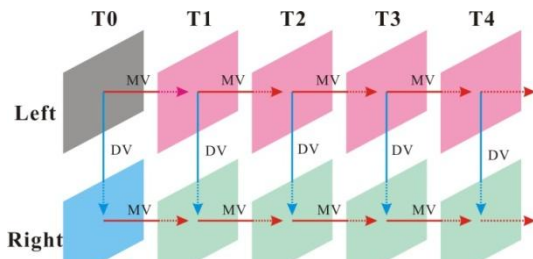


Fig.3 Encoding scheme of proposed algorithm

Fig. 4 shows the flowchart of proposed algorithm. In the proposed algorithm, the block size of 8×8 is adopted to conceal the erroneous frames since selecting block size 16×16 would bring significant blocking effect, and the block size of 4×4 would result in broken frame.

After deciding the block size, we extend 8 pixels all-around the corresponding block in the previous frame to form a 24×24 -size extended window called EW as shown in Fig. 5.

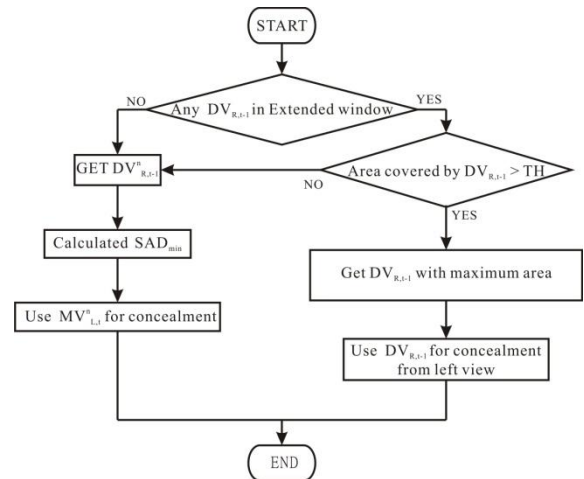


Fig.4 Flowchart of proposed algorithm.

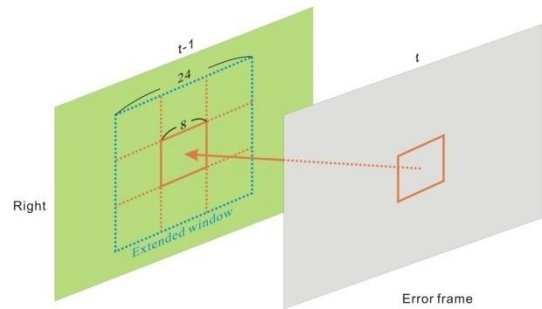


Fig.5 Schematic diagram of extended window.

Once the EW has been decided, we check if there is any disparity vector within EW. If there is, we will calculate the area covered by each disparity vector in EW and check whether any covered area has exceeded a predefined threshold. In our proposed algorithm, the default of threshold is set to half of the EW area. If the area is larger than the threshold, we use the DV with the biggest area size in the EW to conceal the error block. Otherwise, the proposed algorithm will construct a DV set which is composed by the DVs in the EW. If there is no any disparity vector in EW, the DV set will consist of the DVs in the previous frame.

$$DV_{R,t-1} = \{DV_{R,t-1}^0, DV_{R,t-1}^1, \dots, DV_{R,t-1}^{N-1}\} \quad (1)$$

Where R means the right view, and $t-1$ stands for the previous frame.

After obtaining the reference DV set, the DVs inside reference DV set will be used to be projected onto left view with a 8×8 window called covered window (CW) as shown in Fig.6 and the motion vectors with largest area coverage of all DVs will be considered to conceal the erroneous blocks. The selection of motion vector with maximum area coverage is defined in (2).

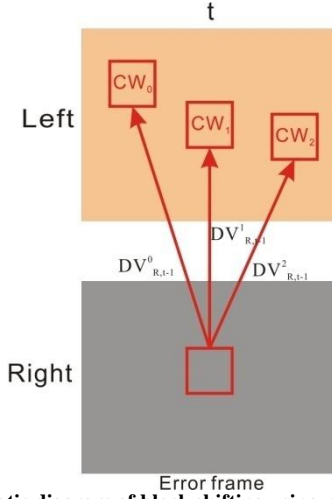


Fig.6 Schematic diagram of block shifting using reference DV

$$MV_{L,t}^n = MV_{L,t}^k | MV \text{ with maximum area in } CW_n \\ k = \text{index of different MV in } CW_n \quad (2)$$

Where L means the left view, t is the current time in timeline, k refers to the index of different MVs in the CW.

Based on the parallelogram-like motion relationship between inter-frame and inter-view correlation as shown in Fig.2, when t^{th} frame of the right view has error occurred, the MV obtained from the corresponding block in the left view shifted by the DV will be very similar to the original MV of the error frame if the corresponding DV in the previous frame is correct. Therefore, the SADs between B_1 and B_2 as shown in Fig. 7 are calculated for each MV corresponding to each CW to determine the block for concealing the current erroneous block. However, a situation might be faced when the block shifted from the wrong DV and the block from wrong MV is similar to each other. To solve this problem, we further consider SADs between the left and right views in the previous frame ($F_{L,t-1}$). As shown in Fig.7, the disparity vector of $DV_{R,t-1}^n$ has been selected and projected onto the left view to obtain the block B_1 pointed by $DV_{R,t-1}^n$. Afterwards, the motion vector $MV_{L,t}^n$ with the largest area coverage by B_1 will be selected and projected onto the previous frame of the right view to obtain B_2 . Then the corresponding block B_3 pointed by $DV_{R,t-1}^n$ from B_2 will be used to calculate the SAD between B_2 and B_3 . Finally, the motion vector with minimum joint SADs will be derived by (3) to conceal the error block B_c . By jointly considering the SADs between views and frames, the concealing results can be further improved.

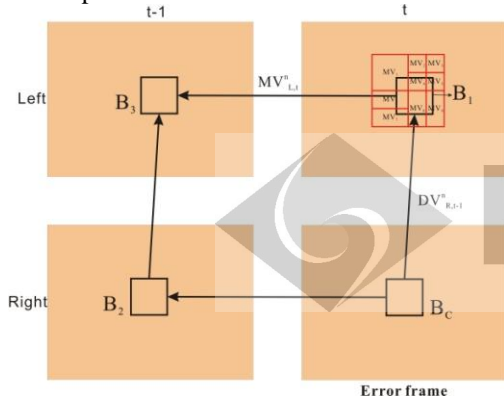


Fig.7 Illustration of the proposed algorithm

$$MV_{L,t}(i,j) = \underset{MV_{L,t}^n \in MV_{L,t}}{\operatorname{argmin}} \sum_{a=0}^7 \sum_{b=0}^7 (|F_{R,t-1}(8 \times i + MV_{L,t,x}^n + a, 8 \times j + MV_{L,t,y}^n + b) \\ - F_{L,t}(8 \times i + DV_{R,t-1,x}^n + a, 8 \times j + DV_{R,t-1,y}^n + b)| \\ + |F_{R,t-1}(8 \times i + MV_{L,t,x}^n + a, 8 \times j + MV_{L,t,y}^n + b) \\ - F_{L,t-1}(8 \times i + MV_{L,t,x}^n + DV_{R,t-1,x}^n + a, 8 \times j + MV_{L,t,y}^n + DV_{R,t-1,y}^n + b)|) \quad (3)$$

The notations of (3) are listed as follows,

- i and j : the horizontal and vertical indexes of the 8×8 block in a frame.
- a and b : the horizontal and vertical indexes of the pixel inside the block.
- $F_{R,t-1}$: the previous frame of right view.
- $F_{L,t}$: the current frame of left view.
- $F_{L,t-1}$: the previous frame of left view.
- $DV_{R,t-1,x}^n$: the horizontal component of the n^{th} DV in the block of the right view of the previous frame.
- $DV_{R,t-1,y}^n$: the vertical component of the n^{th} DV in the block of the right view of the previous frame.
- $MV_{L,t,x}^n$: the horizontal component of the n^{th} MV in the block of left view of the current frame.
- $MV_{L,t,y}^n$: the vertical component of the n^{th} MV in the block of the left view of the current frame.

III. SIMULATION RESULTS

In this section, several simulation results are given to demonstrate the efficiency of our proposed MVC error concealment algorithm.

The test sequences we used for simulation are Ballroom(640×480), Exit(640×480), Flamenco(640×480), Race1(640×480), AkkoKayo(640×480), and Vassar(640×480). In our simulation, we assume that only right view has whole frame error and the left view hasn't. The work [7] is adopted for comparison in this paper but we have made some modifications for [7] in order to let the algorithm of [7] could be able to support whole frame loss error concealment. The simulation settings are summarized in Table I.

Table I: Simulation parameters

Parameter	Value
View	2
GOP structure	IPPP...
Intra refresh	Only the first frame
Frame rate	25
Frame number	100
Reference frame number	2
Coding order	Left View \rightarrow Right view
QP	32
Packet lose rate (PLR)	5% · 10% · 15% · 20%

The simulation results are shown in Table II. Δ PSNR refers to the difference of the proposed and [7]. From this table, we can observe that the proposed algorithm has better average PSNR results of all frames than [7]. The proposed algorithm even has up to 4dB higher than [7] under 5% PLR in Race1 sequence.

Fig. 8 shows the results of subjective quality comparison. Here we compare the 55th frame of the Race1 sequence and the 10th frame of the Vassar sequence. Fig. 8(a) and (d) are the error free frame, (b) and (e) are concealed by [7], (c) and

(f) are concealed by the proposed algorithm. In the Race1 sequence, we can observe that our proposed algorithm can significantly improve the subjective quality when compared to [7] and thus reduce the error propagation. In the Vassar sequence, our result is more similar to the error frame. However, the result of [7] shows the unsatisfied quality in

which the person is far from the car while the people is near the car in the original sequence.

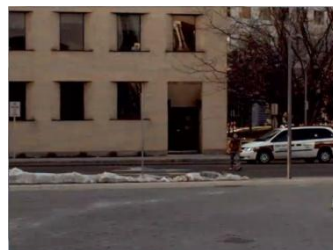
The average decoding time and overhead of our proposed algorithm of various error rates for the six sequences listed in Table II are shown in Table III.

Table II: Comparison of the average PSNR (dB) on all frames between [7] and the proposed algorithm

Sequences		Ballroom	Vassar	Race1	Exit	Akkokayo	Flamenco	Average
Error free		35.499	34.957	35.820	37.214	36.930	38.448	36.478
5%	[7]	29.292	34.388	22.780	33.614	28.184	29.743	29.667
	Proposed	30.241	34.541	26.926	33.608	31.232	30.239	31.131
	Δ PSNR	0.949	0.153	4.146	-0.006	3.048	0.496	1.464
10%	[7]	25.470	33.583	20.040	30.970	23.688	26.524	26.713
	Proposed	26.063	33.721	23.877	30.972	26.293	26.608	27.922
	Δ PSNR	0.593	0.138	3.837	0.002	2.605	0.084	1.210
15%	[7]	23.666	32.480	18.336	29.114	20.762	24.557	24.819
	Proposed	24.399	33.114	21.557	28.902	23.547	24.917	26.073
	Δ PSNR	0.733	0.634	3.221	-0.212	2.785	0.360	1.254
20%	[7]	22.334	31.964	17.693	27.213	20.726	22.858	23.798
	Proposed	23.439	32.726	20.698	27.811	22.921	23.446	25.174
	Δ PSNR	1.105	0.762	3.005	0.598	2.195	0.588	1.376



(a) error free(35.043dB)



(d) error free(35.043dB)



(b) [7] (14.985dB)



(e) [7] (27.332dB)



(c) proposed(19.021dB)



(f) proposed(30.437dB)

Fig. 8 Subjective quality comparison.(a)(b)(c): the right view of the 55th frame of the Race1 sequence, (d)(e)(f):the right view of the 10th frame of the Vassar sequence.

Table III: Average decoding time of the proposed algorithm(ms/frame)

Packet loss rate	Error free	Proposed	Overhead
5%	112.75	163.84	45.31%
10%	112.75	230.59	104.51%
15%	112.75	314.87	179.26%
20%	112.75	360.85	220.01%

IV. Conclusion

The proposed algorithm jointly considers the SAD between views and frames, and aims on error concealment for frame loss and reducing error propagation in efficient way for multi-view video decoding. Experimental results demonstrate that our proposed algorithm can achieve better performance than previous work.

Reference

- [1] Y. S. Ho and K. J. Oh, "Overview of Multi-view Video Coding," in *Proceedings of International Workshop on Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services*, pp. 5-12, June 2007.
- [2] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard," *IEEE Proceedings*, vol. 99, no. 4, pp. 626-642, April 2011.
- [3] "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14496-10 AVC)," in joint video team of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, 2003.
- [4] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, July 2003.
- [5] Y. Chen, C. Cai, and K. K. Ma, "Stereoscopic video error concealment for missing frame recovery using disparity-based frame difference projection," in *Proceedings of IEEE International Conference on Image Processing*, pp. 4289-4292, November 2009.
- [6] T. Y. Chung, S. Sull, and C. S. Kim, "Frame loss concealment for stereoscopic video plus depth sequences," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1336-1344, August 2011.
- [7] Y. Zhou, C. Hou, R. Pan, Z. Yuan, and L. Yang, "Distortion analysis and error concealment for multi-view video transmission," in *Proceedings of IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pp. 1-5, March 2010.
- [8] M. Ebian, M. El-Sharkawy, and S. El-Ramly, "Adaptive error concealment algorithm for multiview coding based on lost MBs sizes and using dynamic selection of lower candidates MBs," in *Proceedings of International Computer Engineering Conference*, pp. 26-29, December 2012.
- [9] S. Gao and K. K. Ma, "Error-Resilient H.264/AVC Video Transmission Using Two-Way Decodable Variable Length Data Block," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 340-350, March 2010.