Robust 3-D Video System Based on Modified Prediction Coding and Adaptive Selection Mode Error Concealment Algorithm

Walid El-Shafai
Dept. of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menouf 32951, Menoufia University, Egypt
Email: eng.waled.elshafai@gmail.com

***Abstract*—To improve the quality of 3D Multi-View Video (MVV) transmitted over noisy channels, we deploy Modified Prediction Multi-view Video Coding (MPMVC) algorithm at the encoder and Adaptive Selection Mode Error Concealment (ASMEC) algorithm at the decoder. MPMVC scheme exploits the best available correlations between video streams to construct the optimum MVC prediction structure. Deploying MPMVC at encoder can make the transmitted 3D MVV very robust to severe channel errors through aiding the proposed ASMEC algorithm to mitigate error propagation. To further improve the decoded MVV quality without increasing transmission rates, we deploy ASMEC algorithm at the decoder that can exploit the spatial, temporal and inter-view correlations in concealing transmission errors. Our results show that the proposed MPMVC-ASMEC schemes jointly improve the objective and subjective quality of reconstructed 3D multi-view video sequences and are more robust to transmission errors.**

***Index Terms*—3D video, multi-view video coding, prediction coding structure, error concealment**

I. INTRODUCTION

3D multi-view video has received wide attention lately and is expected to quickly replace traditional 2D video in many applications. In Multi-view Video Coding (MVC), video sequences are generated by capturing the same scene simultaneously with multiple cameras located at different view-point (angles). For efficient 3D video coding, MVC must exploit the spatial and temporal correlations within each video as well as the inter-view correlations between the video streams to increase the coding efficiency. Fig. 1 shows the general MVC Group Of Pictures (GOP) Prediction Structure (PS) [1] for the Joint Multi-view Video Model (JMVM) [2]. With this PS, frames are predicted from temporal neighbors, and spatial neighbors in different views. The choice of GOP Prediction Structure (GOP-PS) in JMVM is critical and must be carefully configured to achieve high coding performance and to increase the efficiency of the Error Concealment (EC) algorithms.

In multi-view video coding, the inter-view correlation always depends on the selection of the GOP-PS [3], [4]. So, a good choice of the GOP-PS and the stream coding order can reduce the coding bit rate and improve the inter-view correlations between the video streams [5]. In [5], the authors introduced a Self-Configurable Multi-view Video Coder (SCMVC) at the encoder, which efficiently exploits the inter-view correlations between video streams. The SCMVC adaptively estimates the GOP prediction structure without prior knowledge of the multi-camera arrangement. It is shown that deploying SCMVC reduces the transmitted bit rate and is very robust to camera failures and the severe channel errors. The inter-view correlations provided by SCMVC [5] can be used to enhance the performance of the proposed EC algorithm at the decoder side.

3D multi-view video transmitted over wireless networks is always subject to packet losses including both random and burst errors. It is not possible to retransmit all erroneous or lost packets due to delay constraints on real-time video transmission. Due to the predictive coding structure of MVC compressed video, which utilizes intra and inter coded frames, errors could propagate to the subsequent frames and to the adjacent views and result in poor video quality [6]. Therefore, there is a need for post-processing error-concealment (EC) methods at the decoder. EC algorithms can reduce the visual artifacts caused by channel errors or erasures without increasing the bit rate or transmission delay or requiring any difficult modifications at the encoder.

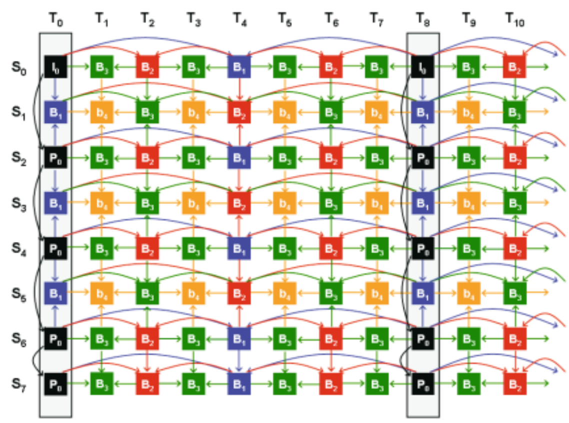

Figure 1. The general GOP prediction structure used in JMVM.

EC algorithms depend on the inter-view correlations between the multi-view video streams to conceal lost

Manuscript received January 24, 2012; revised June 25, 2014.

doi: 10.12720/ijsps.2.2.155-159

packets or frames in the received 3D video data [7], [8]. In this paper, we introduce an efficient Adaptive Selection Mode Error Concealment (ASMEC) algorithm for multi-view video encoded by Modified Prediction Multi-view Video Coding (MPMVC) algorithm. The MPMVC is used to blindly estimate the best GOP-PS together with the best stream coding order [5]; as a result the inter-view correlations between the multi-view video streams will be improved. These improvements will enhance the efficiency of the proposed ASMEC algorithm.

The rest of this paper is organized as follows: Section II presents the Adaptive Selection Mode Error Concealment (ASMEC) algorithm, Section III presents the proposed enhanced 3D multi-view video system with MPMVC algorithm at encoder and robust decoding with the proposed ASMEC algorithm, Section IV presents our experimental simulation results and Section VI concludes the paper.

II. Proposed ASMEC Algorithm

In this section, we present the Adaptive Selection Mode Error Concealment (ASMEC) algorithm that will be deployed in the proposed enhanced MVC system. ASMEC jointly exploits correlations in the space, time and inter-view domains. ASMEC can recover the lost Macro Blocks (MBs) of intra and inter coded frames. For intra-frames EC, EC exploits correlations in the space and time dimensions, and the hybrid Space-Time Domain Error Concealment (STDEC) mode is used. For inter-frame EC, EC is done in either the time or inter-view dimensions and three EC modes can be deployed: Time Domain Error Concealment (TDEC), Inter-view Domain Error Concealment (IVDEC) and joint Time and Inter-view Domain Error Concealment (TIVDEC).

Fig. 2 shows the flow chart of the ASMEC algorithm, which can detect errors in any received view (odd or even view) and in any received frame (I or P or B). The ASMEC algorithm can select one of the following EC modes depending on error location, as shown in Fig. 2.

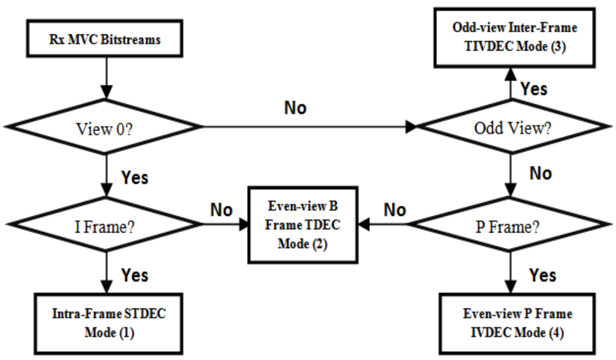

Figure 2. Flow chart of the proposed ASMEC algorithm.

A. *Mode (1): Space-Time Domain EC (STDEC)*

1. Find the 8x8 adjacent sub-blocks to the lost MB and their matching blocks in the reference frame.
2. Find the Motion Vectors (MVs) between the adjacent sub-blocks and their matching blocks.
3. Select the MBs that give the smallest Sum of Absolute Differences (SAD) [9].
4. Apply the Weight Pixel Averaging Algorithm (WPAA) [10] to find the matching pixels surrounding the lost MB's pixels.
5. Find the Disparity Vectors (DVs) between pixels inside the lost MB and pixels surrounding the lost MB.
6. Calculate the average value of the selected MVs and DVs found in the previous steps.
7. Replace the lost MBs with the averaged calculated value in step 6.

B. *Mode (2): Time Domain EC (TDEC)*

1. Apply the Outer Block Boundary Matching Algorithm (OBBMA) to find the matching pixels [7].
2. Find the most matched candidates MVs to the lost MB.
3. Average MVs values of the candidate MBs.
4. Replace the lost MBs with the candidates MBs by using the averaged calculated value.

C. *Mode (3): Time-Inter-View Domain EC (TIVDEC)*

1. Apply the WPAA [10] and OBBMA [7] algorithms.
2. Find the most matched candidate DVs and MVs to the lost MB.
3. Average DVs and MVs values of the candidate MBs.
4. Set appropriate coefficient values to the averaged values of MVs and DVs (avg (MVs) and avg (DVs), respectively) depending on Scene Change Detection Algorithm [10] by selecting between the following two cases:
   - Candidate MB = 1/3 avg (MVs) + 2/3 avg (DVs).
   - Candidate MB = 2/3 avg (MVs) + 1/3 avg (DVs).

This depending on "Is the Temporal information > Spatial information or vice versa?".

5. Replace the lost MBs with the candidates MBs by using the weighted average calculated value of MVs and DVs in the previous step.

D. *Mode (4): Inter-View Domain EC (IVDEC)*

1. Apply WPAA [10] to find the matching pixels.
2. Find the most correlated candidate DVs to lost MB.
3. Average DVs values of the candidate MBs.
4. Replace the lost MB with the candidate MBs by using the averaged calculated value.

III. Proposed MPMVC-ASMEC System

EC algorithms that were proposed in literature, e.g. [1], [7]-[10] for MVV transmission were used the traditional JMVM [2] which uses the fixed GOP-PS shown in Fig. 1. However, in this paper, we propose an enhanced multi-view video system that used the proposed ASMEC algorithm at the decoder based on using the MPMVC algorithm at the encoder, which is proposed in [5]. The MPMVC is a 3D video compression algorithm based on JMVM has the advantage of estimating the optimum GOP-PS with the best stream coding order in order to efficiently exploit the intra and inter-view correlations in the coding process [5]. Since MPMVC arranges adjacent views to have maximum inter-view correlation, the ASMEC algorithm is expected to significantly enhance the quality of its video on erroneous channels, as ASMEC exploits inter-view correlations for error concealment. ASMEC algorithm exploits this advantage in the decoder to enhance the quality of the reconstructed 3D multi-view video sequences.

In our proposed enhanced MPMVC-ASMEC system, the 3D video data is encoded based on configuration of GOP-PS using the final conclusion results of the proposed MPMVC algorithm that was presented in [5]. Then the encoded bit stream is transmitted over a noisy channel. The received noisy data is decoded and error concealed by the proposed ASMEC algorithm assuming the same configuration estimated with the MPMVC algorithm. In detail, the MPMVC algorithm was introduced in [5].

IV. Experimental Results

To evaluate the performance of the proposed MPMVC-ASMEC algorithm, we run some test experiments on well-known 3D video sequences, (Uli [11], Ballet [12] and Breakdancer [13]). For each sequence, the coded bit-streams are transmitted over a noisy channel with random Packet Loss Rates (PLRs) of 3%, 5%, 10% or 20%. The received bit-streams are then decoded and concealed by the MPMVC-ASMEC algorithm. The location of the lost MBs can be detected by using the received Motion Vectors (MVs) and Disparity Vectors (DVs) table.

To illustrate the effect of our robust MPMVC-ASMEC system, we compare its performance to that of the MVC system with conventional JMVM [2] when ASMEC is used to conceal erroneous frames (JMVM-ASMEC), as well as with that of conventional JMVM when no EC is deployed (JMVM-NoEC) and with that of MPMVC with no EC deployed (MPMVC-NoEC).

In our results, the MPMVC-ASMEC refers to our proposed enhanced system. The MPMVC algorithm adaptively changes the GOP-PS and the stream coding order to exploit the inter-view correlations [5]. Thus the estimated coding PS by the MPMVC is different from the PS used with JMVM. For example, Fig. 3 shows the conventional JMVM PS used with Uli, Breakdancer and Ballet sequences, while Fig. 4 shows the PS used with MPMVC for the Uli sequence. Also, Fig. 5 shows the PS used with the MPMVC for Ballet and Breakdancer sequences. Due to the difference in PS between the JMVM and the MPMVC, we add errors to the same frame types within the JMVM and the MPMVC. Thus for the selected Uli, Ballet and Breakdancer sequences, we choose S_o, S_1, S_2 and S_7 views for the test experiment and we take the average PSNR value over all the concealed frames as shown in Table I.

In Table I, we compare the objective PSNR values of the proposed MPMVC-ASMEC system with that of the JMVM-ASMEC system at different packet loss ratios. For Ballet and Breakdancer sequences, as it can be observed from Table I, Fig. 3, and Fig. 5, the views S_o and S_7 have opposite types (I and P) in JMVM and MPMVC respectively. By comparing the average PSNR, we observe that the proposed MPMVC-ASMEC system always achieves superior average PSNR. We can observe that the MPMVC-ASMEC algorithm has a significant average gain in objective PSNR about 0.92 dB over the JMVM-ASMEC algorithm. We can also observe that ASMEC is crucial as it provides about 10 dB average gain in both cases.

Fig. 6 shows the subjective simulation results for the Ballet sequence. We select the 16th intra-coded and inter-coded frames inside S_o, S_1 and S_2 views with PLR=20%. We compare the error concealment performance of the proposed MPMVC-ASMEC system with that of the JMVM-ASMEC system. We conclude that the robust MPMVC-ASMEC system gives better subjective results compared to the JMVM-ASMEC system.

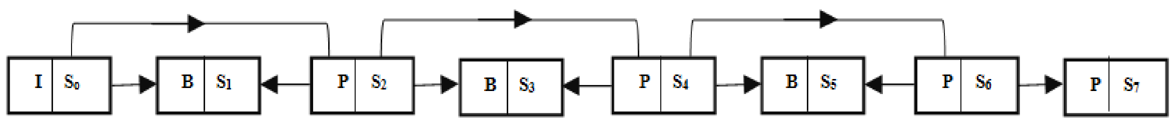

Figure 3. The JMVM PS for 8 views with parallel camera arrangement for Uli, ballet and breakdancer sequences.

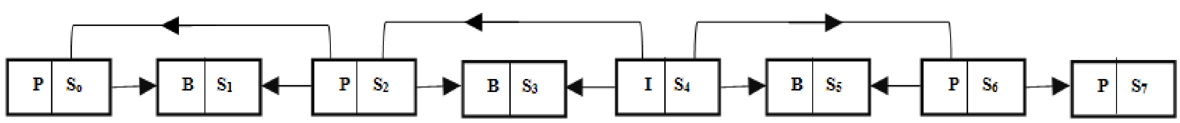

Figure 4. The MPMVC PS for the Uli sequence [5].

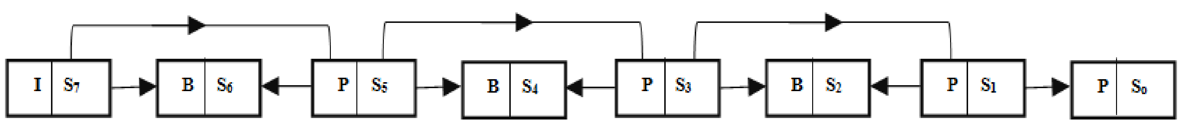

Figure 5. The MPMVC PS for the ballet and breakdancer sequences [5].

TABLE I. PSNR PERFORMANCE FOR ULI, BALLET AND BREAKDANCER VIDEO SEQUENCES WITH DIFFERENT PLR FOR THE SELECTED 16TH FRAME INSIDE S_O, S_1, S_2, AND S_7 VIEWS.

			JMVM-ASMEC						MPMVC-ASMEC				
			Packet Loss Rate (PLR) %						Packet Loss Rate (PLR) %				
		View (Frame)	**0%**	**3%**	**5%**	**10%**	**20%**	**View (Frame)**	**0%**	**3%**	**5%**	**10%**	**20%**
Uli	**PSNR (dB)**	**S_o (I)**	34.95	32.371	30.160	27.754	26.754	**S_o (P)**	35	34.982	34.481	32.252	31.126
		S_1 (B)	34.95	34.041	33.813	31.630	29.361	**S_1 (B)**	35	34.651	34.233	32.327	29.881
		S_2 (P)	34.95	34.472	33.773	32.276	30.558	**S_2 (P)**	35	34.972	34.567	32.316	31.173
		S_7 (P)	34.95	34.474	33.717	32.469	30.317	**S_7 (P)**	35	34.739	34.426	32.424	31.301
		Average (dB)	**34.95**	**33.840**	**32.866**	**31.032**	**29.248**	**Average (dB)**	**35**	**34.837**	**34.427**	**32.330**	**30.870**
		No EC	**34.95**	**27.776**	**25.803**	**21.152**	**18.984**	**No EC**	**35**	**28.443**	**26.446**	**21.736**	**19.526**
Ballet	**PSNR (dB)**	**S_o (I)**	38.7	36.121	33.912	31.711	30.504	**S_o (P)**	38.72	38.621	37.832	36.572	34.707
		S_1 (B)	38.7	37.792	37.553	35.380	33.058	**S_1 (P)**	38.72	38.621	37.833	36.573	34.717
		S_2 (P)	38.7	38.226	37.521	36.221	34.306	**S_2 (B)**	38.72	38.216	37.872	35.721	33.423
		S_7 (P)	38.7	38.169	37.603	36.402	34.193	**S_7 (I)**	38.72	36.401	34.213	32.039	30.901
		Average (dB)	**38.7**	**37.577**	**36.647**	**34.929**	**33.015**	**Average (dB)**	**38.72**	**37.965**	**36.938**	**35.226**	**33.437**
		No EC	**38.7**	**31.684**	**29.560**	**24.921**	**22.743**	**No EC**	**38.72**	**31.891**	**29.772**	**25.189**	**23.038**
Breakdancer	**PSNR (dB)**	**S_o (I)**	37.5	34.921	32.711	30.511	29.304	**S_o (P)**	37.52	37.351	36.531	35.262	33.401
		S_1 (B)	37.5	36.592	36.362	34.184	31.851	**S_1 (P)**	37.52	37.372	36.641	35.283	33.511
		S_2 (P)	37.5	37.023	36.321	35.026	33.162	**S_2 (B)**	37.52	36.879	36.656	34.331	32.102
		S_7 (P)	37.5	37.117	36.286	35.263	33.293	**S_7 (I)**	37.52	35.523	33.479	31.329	29.897
		Average (dB)	**37.5**	**36.413**	**35.420**	**33.746**	**31.903**	**Average (dB)**	**37.52**	**36.781**	**35.827**	**34.051**	**32.228**
		No EC	**37.5**	**30.443**	**28.353**	**24.37**	**21.277**	**No EC**	**37.52**	**30.685**	**28.586**	**24.622**	**21.895**

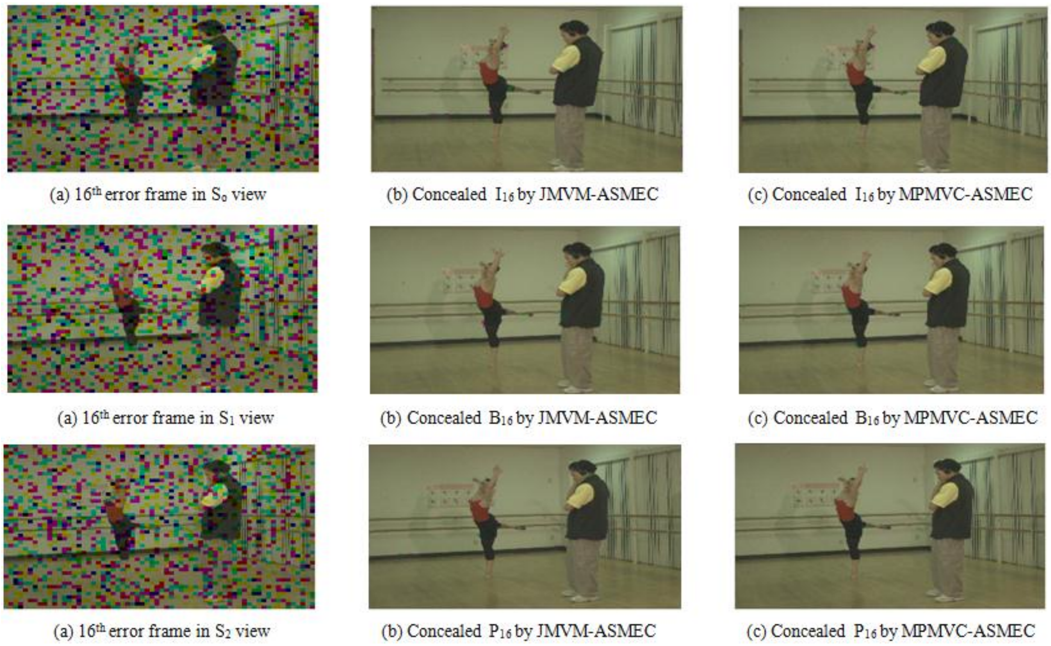

Figure 6. Subjective simulation results for the selected 16th frame within the So, S1, and S2 views within Ballet video sequence with PLR=20%.

V. CONCLUSION

In this paper, we introduced a 3D multi-view video system that is more robust to channel errors by enhancing the encoder with MPMVC algorithm that optimizes group of picture structure and enhancing the decoder with adaptive selection mode error concealment in time, space and inter-view dimensions. Our results show that the proposed system is more robust to channel errors while having a lower transmission bit rate. The proposed MPMVC-ASMEC algorithm is verified by experimental results on publicly available 3D multi-view video sequences. Our experimental results show that our proposed MPMVC-ASMEC system has a significant advantage, in the objective PSNR metric and the subjective video quality, over the conventional JMVM system with error concealment.

REFERENCES

[1] P. Nasiopoulos, L. Coria-Mendoza, H. Mansour, and A. Golikeri, “An improved error concealment algorithm for intra-frames in H.264/AVC,” in *Proc. IEEE International Symposium on Circuits and Systems*, 2005, pp. 320-323.

[2] A. Vetro, P. Pandit, H. Kimata, and A. Smolic, “Joint multiview video model (JMVM) 8.0,” JVT-AA207, Geneva, Switzerland, Apr. 2008.

[3] Y. Zhang, G. Jiang, M. Yu, and Y. Ho, "Adaptive multi-view video coding scheme based on spatiotemporal correlation analyses," *ETRI Journal*, vol. 31, no. 2, pp. 151-161, 2009.

[4] Z. Feng and A. Ping, "Multi-view video coding based on sequence correlation," in *Proc. International Conference on Audio Language and Image Processing (ICALIP)*, Shanghai, Nov. 2010.

[5] H. Hussein, M. El-Khamy, and M. El-Sharkawy, "Blind configuration of multi-view video coding," *in Proc. 30th IEEE International Conference on Consumer Electronics (ICCE) (ICCE2012)*, Las Vegas, Jan. 2012.

[6] R. Schäfer, "Review and future directions for 3D-video," in *Proc. 25th PCS*, 2006, pp. 1-11.

[7] K. Song, T. Chung, Y. Oh, and C.-S. Kim, "Error concealment of multi-view video sequences using inter-view and intra-view correlation," *Journal of Visual Communication and Image Representation*, vol. 20, no. 4, pp. 281-292, 2009.

[8] W. El-Shafai, B. Hrusovsky, M. El-Khamy, and M. El-Sharkawy, "Joint space-time-view error concealment algorithms for 3D multi-view video," in *Proc. 18th IEEE International Conference on Image Processing (IEEE ICIP2011)*, Brussels, Belgium, Sep. 2011, pp. 2249-2252.

[9] B. Hrušovský, J. Mochnáč, and S. Marchevský, "Temporal-Spatial error concealment algorithm for intra-frames in H.264/AVC coded video," in *Proc. Radioelektronika (RADIOELEKTRONIKA), 20th International Conference*, Apr. 2010, pp. 1-4.

[10] T. Chung, K. Song, and C. Kim, "Error concealment techniques for multi-view video sequences," in *Proc. 8th Pacific Rim Conference on Multimedia*, Hong Kong, China, Dec. 2007, pp. 619-627.

[11] (Mar. 2013). 3D Uli video sequence. [Online]. Available: http://www.3dtv-research.org/3dav_CfP_FhG_HHI/

[12] (Apr. 2013). 3D ballet video sequence. [Online]. Available: ftp://ftp.ne.jp/KDDI/multiview/

[13] (Mar. 2013). 3D breakdancer video sequence. [Online]. Available: http://www.research.microsoft.com/vision/ImageBasedRealities/3DVideoDownload/

Walid El-Shafai was born in Alexandria, Egypt, on April 19, 1986. He received the B.Sc degree in Electronics and Electrical Communication Engineering from Faculty of Electronic Engineering (FEE), Menoufia University, Menouf, Egypt in 2008 and M.Sc degree from Egypt-Japan University of Science and Technology (E-JUST) in 2012. He is currently working as a Teaching Assistant and Ph.D. Researcher in ECE Dept. FEE, Menoufia University. His research interests are in the areas of Wireless Mobile and Multimedia Communications, Image and Video Signal Processing, 3D Multi-view Video Coding, Error Resilience and Concealment Algorithms for H.264/AVC and H.264/MVC Standards.