3DR EXPRESS

Joint Adaptive Pre-processing Resilience and Postprocessing Concealment Schemes for 3D Video Transmission

Walid El-Shafai

Received: 5 December 2014 / Revised: 7 February 2015 / Accepted: 22 February 2015 / Published online: 27 February 2015 © 3D Research Center, Kwangwoon University and Springer-Verlag Berlin Heidelberg 2015

Abstract 3D video transmission over erroneous networks is still a considerable issue due to restricted resources and the presence of severe channel errors. Efficiently compressing 3D video with low transmission rate, while maintaining a high quality of received 3D video, is very challenging. Since it is not plausible to re-transmit all the corrupted macro-blocks (MBs) due to real time applications and limited resources. Thus it is mandatory to retrieve the lost MBs at the decoder side using sufficient post-processing schemes, such as error concealment (EC). In this paper, we propose an adaptive multi-mode EC (AMMEC) algorithm at the decoder based on utilizing preprocessing flexible macro-block ordering error resilience (FMO-ER) technique at the encoder; to efficiently conceal the erroneous MBs of intra and inter coded frames of 3D video. Experimental simulation results show that the proposed FMO-ER/ AMMEC schemes can significantly improve the objective and subjective 3D video quality.

W. El-Shafai (\boxtimes) Department of Electronics and Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt e-mail: eng.waled.elshafai@gmail.com

Graphical Abstract

Keywords 3D video - Error resilience and concealment - Flexible macro-block ordering - Macro-blocks (MBs) - Wireless networks

1 Introduction

3D video (3DV) has received a huge attention recently. It is expected to replace traditional video in the near future in many applications including 3DTV, entertainment, gaming, medicine, education, and so on. 3DV created by capturing the same scene simultaneously with multiple cameras located at different perspectives, offering interactivity as well as 3D perception [\[1](#page-11-0)]. 3D multi-view video sequence exhibits high inter-view correlations, in addition to spatiotemporal correlations within each view. Therefore how to efficiently compress and transmit 3D videos with high quality has been a hot and urgent issue to meet future bandwidth constraints.

Efficiently compressing 3DV with low transmission rate, while maintaining a high quality of received 3D video, is very challenging. 3DV transmitted over wireless networks is always subject to packet losses including both random and burst errors. There are varieties techniques can be used to limit the impact of packet losses prior to transmission. In literature [[2,](#page-12-0) [3](#page-12-0)],

forward error correction (FEC) and automatic repeated request (ARQ) methods are introduced, however they increase the transmission rate and present latency. Subsequently, it is not possible to retransmit all erroneous or lost packets due to bandwidth and delay constraints on real-time video transmission. Error concealment (EC) algorithms have the advantage of improving the received video quality without modifications in the transmission rate or in the encoder hardware or software, which aim to retrieve the lost MBs from the available information at the encoder.

Fig. 1 Different types of flexible macro-block ordering (FMO)

To further improve the objective and subjective qualities of the decoded 3D video without increasing transmission rates, we employ flexible macro-block ordering error resilience (FMO-ER) at the encoder to aid the proposed EC algorithms at the decoder to reconstruct the lost macro-blocks (MBs) and frames efficiently. FMO is one of the new error resilience tools [\[4](#page-12-0)], which can be used to mitigate the effects of losses in error prone environments. It also provides a way to spread the erroneous MBs within the frame and take advantage of the spatial locations of the successfully decoded MBs for facilitating EC mechanism. Therefore, FMO-ER is a powerful interleaving technique at the encoder when conjunction with EC algorithms at the decoder.

In this paper, we focus on concealing the precompressed 3DV sequences generated by joint multiview video coding (JMVC) reference software [\[5](#page-12-0)], based on H.264/AVC codec [[6\]](#page-12-0). We propose an adaptive multi-mode error concealment (AMMEC) algorithm at decoder based on using FMO-ER technique at the encoder. The proposed FMO-ER/ AMMEC scheme is used to reconstruct the 3DV sequences transmitted over noisy channel. It exploits the intra-view and inter-view correlations between frames and views to conceal lost MBs of intra and inter-frames. The rest of this paper is organized as follows: Sect. [2](#page-4-0) presents background error resilience (ER) and error concealment (EC) in 2D H.264/AVC standard. In Sect. [3](#page-6-0), we introduce the proposed joint

Fig. 3 Flow chart of the proposed AMMEC algorithm

 \blacktriangleleft **Fig. 4** Subjective simulation results for the selected I₁₀₅, B₁₀₅, and P_{105} intra and inter frames within the S_0 , S_1 and S_2 views, respectively, of the ''Ballroom'' sequence at channel PLR 20 %: a original error free frame, b corrupted frame (No EC–No FMO), c AMMEC–No FMO, d AMMEC–DFMO, e AMMEC– IFMO, f No EC–DFMO, g No EC–IFMO

FMO-ER/AMMEC scheme. Section [4](#page-9-0) shows our extensive experimental simulation results and finally Sect. [5](#page-11-0) concludes the paper.

2 Bases of Error Resilience and Concealment

2.1 Error Resilience (ER)

Error resilience (ER) mechanisms are introduced at the encoder in order to make the transmitted video stream more resilient to potential errors and to improve EC at the decoder. The ER schemes adopted by H.264/AVC codec [\[4](#page-12-0)] to mitigate the effect of packet loss are: (a) slice coding that confines the spatial error propagation (b) insertion of regular intra coded frames that limits the temporal error propagation, and (c) flexible macro-block ordering (FMO) that restricts both spatial and temporal error propagation without increasing the transmission rate and allows more flexibly deciding what slice MBs belong to, in order to spread out errors.

FMO allows flexibility in changing the encoding and transmission order of MBs on top of the normal raster scan order. Its main advantage is the ability to contain the spatial and temporal propagation of error within the slice boundary. Since each slice is designed to be decodable independently of other slices. Thus FMO allows the encoder and decoder to resynchronize their states at the slice boundary in the event that there is an error in the bit-streams. The H.264/AVC standard supports six common different FMO map types as shown in Fig. [1;](#page-1-0) they were introduced in detail at $[4]$ $[4]$. The most efficient map types include, amongst others, Dispersed (DFMO) and interleaved (IFMO) macroblock allocation. Both can lead to improved concealment performance (especially at higher packet error rates or in the presence of bursty errors) by increasing the likelihood of having correctly received MBs adjacent to the ones lost [\[7](#page-12-0)].

Dispersed FMO (DFMO) type 1, where consecutive MBs are transmitted in different slice-groups to protect the neighborhood MBs. It uniformly scatters possible errors to the whole frame to avoid error accumulation in a limited region. For all the concealment results presented in this work, DFMO is used and also we compare its increased improvement performance over the IFMO type. Therefore, we propose to use DFMO-ER as an interleaving technique at the encoder to enhance the performance of the proposed AMMEC algorithm at the decoder.

2.2 Error Concealment (EC)

3DV transmission over wireless networks may suffer from random and burst packet losses due to channel errors that seriously degrade the received 3DV quality. Therefore, it is challenging to provide error resilient and concealment for reliable 3DV communications over such wireless lossy networks. EC is an effective way to fix the errors by replacing the missing parts of video content by previously corrected decoded parts of the video sequence in order to eliminate or reduce the visual effects of bit stream error.

Due to the predictive coding structure of 3DVcodec which is shown in Fig. [2](#page-2-0), that is used to compress the transmitted 3DV, which utilizes intra and inter coded frames, therefore errors could propagate to the subsequent frames and to the adjacent views and result in poor video quality [[1\]](#page-11-0). Since it is not possible to retransmit all erroneous or lost packets due to delay constraints on real-time video transmission. Thus there is a need for post-processing EC methods at decoder. EC algorithms are attractive since they have the advantage of reducing the visual artifacts caused by channel errors or erasures without increasing the bit rate or transmission delay. Therefore, we propose using of EC algorithms to enhance the 3DV quality at the decoder through exploiting the inter-view correlations between the 3DVstreams in addition to spatiotemporal correlations between frames within each view, as shown in Fig. [2](#page-2-0).

EC algorithms were proposed for mono view H.264/AVC against transmission errors [\[8–10\]](#page-12-0); can be adopted with specific adaptive modifications to conceal erroneous frames in 3DV sequences. The main difference between EC in 2D and 3D video sequences is that EC in 2D videos is employed only in spatial direction (with the frame itself) or in temporal direction (forward or backward) in the same view. But EC for 3D video sequences can be exploited in all

 (5.3) B₇₆ Inter-Frame

Fig. 5 Subjective simulation results for the selected B_{76} , B_{76} , and B_{76} inter frames within the S_0 , S_1 and S_2 views, respectively, of the "Exit" sequence at channel PLR 20 %: a original error free frame, b corrupted frame (No EC–No FMO), c AMMEC– No FMO, d AMMEC–DFMO, e AMMEC–IFMO, f No EC– DFMO, g No EC–IFMO

spatial, temporal and interview (between different views) directions. Therefore the using of 2D video EC algorithms are expected to be more reliable for concealing errors in 3D H.264/MVC; as they take the advantage of the inter-view correlations between views.

Scene Change Detection Algorithm [\[8](#page-12-0)] is an operation used in our work to determine the degree of motion for the 3D video content. It divides the 3D video content into continuous frames then measures the matching between each two consecutive frames by calculating the luminance color and edge difference. It calculates the motion vector between the two consecutive frames to clarify the degree of change between them and thus the 3D video content type; slow or fast video.

Frame temporal replacement algorithm (FTRA) is a simple temporal EC that is reasonable for static or very slow 3D video sequences, where the lost MBs are replaced by the MBs spatially located at the reference frame. Outer block boundary matching algorithm (OBBMA) [[9\]](#page-12-0) is more sophisticated temporal EC technique that is suitable for fast moving videos. OBBMA determines the MVs between the two pixels wide outer boundary of the replacing MB and the same external boundary of the lost MB. It only uses the outer borders of the reference MBs to check the highly correlated neighborhood MVs, and it is useful in identifying the replacing MBs that minimize the boundary distortion error.

The weight pixel averaging algorithm (WPAA) and directional interpolation EC algorithm (DIECA) algorithms [\[8](#page-12-0)] are used for spatial and inter-view EC. WPAA conceals the damaged pixels using the horizontal and vertical pixels in the neighboring blocks. DIECA conceals the damaged MBs by calculating the object edge direction from the neighboring blocks, where the object edge direction with the largest magnitude is chosen as the direction to be used to conceal the damaged MBs.

3 Proposed Joint Adaptive DFMO-ER/AMM-EC Schemes

In this section, we present our joint proposed dispersed flexible macro-block ordering error resilience (DFMO-ER) and AMMEC for intra-frames (I frames) and inter-frames (P and B frames) of 3D video. Intraframes EC are not only essential for improving the 3DV quality of reconstructed intra-frames but also for improving the 3DV quality of reconstructed interframes in the subsequent frames and views. We propose to use DFMO-ER type 1 as an interleaving method at the encoder, which known as scattered slices technique. It uses a function which is known to both the encoder and decoder to spread the MBs; to increase the probability that a corrupted MB has distortion-free neighbors which can be employed to aid the proposed AMMEC at decoder.

AMMEC algorithms adapt to the motion characteristics of the received 3DVand to the error patterns. AMMEC jointly exploits correlations in the space, time and inter-view domains to recover the lost macro blocks (MBs) of intra and inter coded frames. For intra-frames EC, EC is exploits correlations in the space and time dimensions, and the hybrid adaptive space–time mode error concealment (ASTMEC) is used. For inter-frame EC, EC is done in either the time or inter-view dimensions and three EC modes can be deployed: adaptive time mode error concealment (ATMEC), adaptive inter-view mode error concealment (AIVMEC) and joint Adaptive time and interview mode error concealment (ATIVMEC).

Figure [3](#page-2-0) shows the flow chart of the AMMEC algorithm, which can detect errors in any received view (odd or even view) and in any received frame (I or P or B). The AMMEC algorithm can select one of the following EC modes depending on error location, as shown in Figs. [2](#page-2-0) and [3.](#page-2-0) We can employ the appropriate EC mode depending on the erroneous frame type and its location within the erroneous view.

- 3.1 EC Mode (1): Adaptive Space–Time Mode EC (ASTMEC)
- 1. Find the lost MBs locations inside erroneous Iframe.

 (6.3) P_{217} Inter-Frame

- **Tig. 6** Subjective simulation results for the selected I_{217} , B_{217} , and P_{217} intra and inter frames within the S_0 , S_1 and S_2 views, respectively, of the "Objects2" sequence at channel PLR 20 %: a original error free frame, b corrupted frame (No EC–No FMO), c AMMEC–No FMO, d AMMEC–DFMO, e AMMEC– IFMO, f No EC–DFMO, g No EC–IFMO
	- 2. Check ''Is the received 3DV fast or slow moving video?'', by using scene change detection algorithm.
	- 3. Select between the FTRA or EBBMA algorithms [\[9](#page-12-0)], depending on step 2.

If the received 3DV is a **slow moving** video, Then

- Apply the FTRA algorithm.
- Replace the lost MBs by the MBs located at the same spatial positions in the reference frame.

If the received 3DV is a fast moving video, apply EBBMA algorithm, Then

- Find the 8×8 adjacent sub-blocks to the lost MB and their matching blocks in the reference frame.
- Find the motion vectors (MVs) between the adjacent sub-blocks and their matching blocks.
- Select the candidates MBs that give the smallest sum of absolute differences (SAD) [[10\]](#page-12-0).
- 4. Select between the WPAA or DIECA algorithms [\[8](#page-12-0)], to find the matching pixels surrounding the lost MB's pixels depending on the location of the lost MBs.
	- a. If ''the lost MBs are at the edge or at the corner of the frame''

• Apply the WPAA algorithm. Else

- Apply the DIECA algorithm.
- b. Find the disparity vectors (DVs) between pixels inside the lost MB and pixels surrounding the lost MB for the WPAA algorithm or DIECA algorithm.
- 5. Calculate the average value of the selected candidates MVs and DVs found in the previous steps.
- 6. Check if temporal information \ge spatial information or vice versa.
- 7. Set appropriate coefficient values [[1\]](#page-11-0), to averaged values of MVs and DVs (avg (MVs) and avg (DVs), respectively):

If Temporal information \leq Spatial information

Candidate $MB = 1/3$ avg $(MVs) + 2/3$ avg (DVs).

Else

- Candidate $MB = 2/3$ avg (MVs) + 1/3 avg (DVs).
- 8. Replace the lost MB with the candidates MB calculated using the weighted average of MVs and DVs as in step 7.
- 3.2 EC Mode (2): Adaptive Time Mode EC (ATMEC)
- 1. Find the lost MBs locations inside erroneous evenview B-frame.
- 2. Apply the steps 2 and 3 of the EC Mode (1), to find the matching pixels inside the previous and subsequent reference frames depending on the received 3DV characteristics (slow or fast).
- 3. Find the most correlated candidate MVs to lost MB.
- 4. Average MVs values of the most matched candidates MBs.
- 5. Replace the lost MBs with the candidates MBs by using the averaged calculated value.
- 3.3 EC Mode (3): Adaptive Time-Inter-View Mode EC (ATIVMEC)
- 1. Find the lost MBs locations inside erroneous oddview B-frame.
- 2. Apply the steps 2 and 3 of the EC Mode (1), to find the matching pixels inside the previous and subsequent reference frames depending on the received 3DV characteristics (slow or fast).
- 3. Apply the step 4 of the EC Mode (1),to find the matching pixels inside the left and right reference frames depending on the locations and the pattern of the lost MBs (at the edge or at the corner).

- 4. Find the most matched candidates MVs and DVs
- 5. Average DVs and MVs values of the candidate MBs.

to the lost MB.

- 6. Set appropriate coefficient values [[1\]](#page-11-0), to the averaged values of MVs and DVs (avg (MVs) and avg (DVs), respectively) 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 informa $tion$ $>$ Spatial information or vice versa?".

- 7. Replace the lost MBs with the candidates MBs by using the weighted average calculated value of MVs and DVs in the previous step.
- 3.4 EC Mode (4): Adaptive Inter-View Mode EC (AIVMEC)
- 1. Find the lost MBs locations inside erroneous evenview P-frame.
- 2. Apply the step 4 of the EC Mode (1), to find the matching pixels inside the left reference frame depending on the locations and the pattern of the lost MBs (at the edge or at the corner).
- 3. Find the most correlated candidate DVs to lost MB.
- 4. Average DVs values of the candidate MBs.
- 5. Replace the lost MB with the candidate MBs by using the averaged calculated value.

4 Simulation Results

In order to evaluate the performance of the proposed joint DFMO-ER/AMMEC schemes, we run some test experiments on well-known standard 3DV sequences, (ballroom, exit and objects $2 \begin{bmatrix} 11 \end{bmatrix}$). The moving objects in the ''exit'' sequence are in simple slow motion, while the ones in the ''objects2'' and ''ballroom'' sequences involve fast and complex motions. JMVC [\[5](#page-12-0)] reference software is employed as a platform for our proposed simulation work, based on H.264/AVC codec [[6\]](#page-12-0). All encoding parameters are set sche

according to the JVT common test condition [[9\]](#page-12-0). For each sequence, the encoded bit-streams are produced with applying DFMO-ER technique at the encoder, and then transmitted over a communication noisy channel with various random Packet Loss Rates (PLRs) (3, 5, 10 and 20 %) and then concealed at the decoder by the proposed AMMEC algorithms. We used the peak signal to noise ratio (PSNR) value as the objective measure of the recovered and concealed MBs to evaluate the performance of the proposed DFMO-ER/AMMEC scheme.

To illustrate the effect of our robust proposed DFMO/AMMEC scheme, we compare its performance to that of using of AMMEC only (AMMEC–No FMO), as well as with that of using DFMO or IFMO ER only (No EC–DFMO & No EC–IFMO), and with that when no AMMEC and no FMO are deployed (No EC–No FMO). In our results, the AMMEC–DFMO refers to our proposed improved scheme; we compare its performance also to that of using AMMEC with Interleaved FMO (AMMEC–IFMO). For each sequence, we select random erroneous frames within S_0 , S_1 and S_2 views and conceal each erroneous frame by employing the appropriate proposed EC mode depending on its reference frames as shown in Fig. [2.](#page-2-0) In our simulation results, we assume that the erroneous

frames inside the first three S_0 , S_1 and S_2 views, respectively, are I_{105} , B_{105} , and P_{105} for ballroom sequence and B_{76} , B_{76} , and B_{76} for exit sequence. But we select the lost frames I_{217} , B_{217} , and P_{217} for objects2 sequence. Thus we will select the appropriate adaptive EC mode to conceal each erroneous frame. As shown in Fig. [2,](#page-2-0) for example the erroneous I_{105} or B_{76} or I_{217} frames can be concealed using the ASTMEC algorithm. Also, the corrupted B_{105} or B_{217} inter-frames can be reconstructed by applying ATIVMEC algorithm.

But, the lost MBs within the P_{105} or P_{217} interframes can be recovered by exploiting AIVMEC algorithm. By the way, the corrupted B_{76} and B_{76} inside S_0 and S_2 views for exit sequence can be concealed by applying ATMEC algorithm.

Figures [4](#page-4-0), [5](#page-6-0), and [6](#page-8-0), show the subjective experimental results for the "ballroom", "exit", and "objects2" test video sequences, respectively, which are different in motion characteristics. For each sequence, we select the corrupted intra-coded and inter-coded frames at channel PLR $= 20\%$. We recovered the selected erroneous MBs inside the lost frames with the appropriate proposed EC algorithms with using DFMO in the encoder. We compare the performance of the proposed AMMEC–DFMO scheme to that of

10 Page 12 of 13 3D Res (2015) 6:10

using AMMEC–No FMO, No EC–DFMO, No EC– IFMO, No EC–No FMO and AMMEC–IFMO. The corresponding objective PSNR results for the same selected frames of the same 3D video sequences are shown in Fig. [7,](#page-9-0) Table [1](#page-10-0), and Fig. 8 at different channel PLRs, respectively.

From all the results, we observe that using of joint AMMEC–DFMO scheme has the best subjective and objective results compared to using of AMMEC only (AMMEC–No FMO) or DFMO–ER only (No EC– DFMO). So, the pre-processing DFMO-ER method can be used in the encoder to aid the proposed postprocessing AMMEC algorithms in the decoder to mitigate the channel errors efficiently. Our proposed AMMEC–DFMO algorithm achieved improved PSNR results comparing to the AMMEC–IFMO algorithm, this is because of DFMO is more reliable than IFMO-ER type. Also, we detect that our proposed AMMEC– DFMO scheme gives sufficient results for different characteristics 3DV sequences at various PLRs.

5 Conclusions

In this paper, we have proposed adaptive postprocessing multi-mode error concealment (AMMEC) algorithm at the decoder based on utilizing preprocessing DFMO-ER technique at the encoder; to conceal the erroneous MBs of intra and inter coded frames of 3D video corrupted by random channel errors. The main advantage of our proposed AMMEC–DFMO scheme is to jointly utilize the spatial, temporal and inter-view correlations in 3D video sequences for EC of both intra-frames and interframes. Our simulation results show that the proposed scheme is more robust to channel errors while having a lower transmission bit rate. Our results demonstrate the significance of using DFMO-ER at the encoder in addition to AMMEC at the decoder to enhance the subjective video quality, as well significant gain in objective PSNR. We demonstrated that the proposed scheme can recover different characteristics of 3D multi-view videos efficiently with high video quality.

References

1. El-Shafai, W., Hrusovsky, B., El-Khamy, M., & El-Sharkawy, M. (2011). Joint space-time-view error concealment algorithms for 3D multi-view video. In 18th IEEE international conference on image processing (IEEE ICIP2011) (pp. 2249–2252). Belgium: IEEE.

- 2. Kuo, C., Shih, C., & Hwang, W. (2012). Modeling and analysis of frame-level forward error correction for MPEG video over burst-loss channels. Applied Mathematics & Information Sciences, 8(4), 1845–1853.
- 3. Sachs, D., Kozintsev, I., & Yeung, M (2001). Hybrid ARQ for robust video streaming over wireless LANs. In IEEE international conference on information technology: Coding and computing, 2–4 April (pp. 317–321).
- 4. Kumar, S., Xu, L., Mandal, M., & Panchanathan, S. (2006). Error resilienc schemes in H.264/AVC standard. Journal of Visual Communication and Image Representation, 17(2), 425–450.
- 5. WD 4 Reference Software for Multiview Video Coding (MVC). (2014). [http://wftp3.itu.int/av-arch/jvt-site/2009_](http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD207.zip) [01_Geneva/JVT-AD207.zip](http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD207.zip). Accessed Aug 2014.
- 6. <http://iphome.hhi.de/suehring/tml/>. (2014). Accessed Nov 2014
- 7. [ftp://iva07.ntua.gr/pub/4dkonto/WIAMIS-2005/defevent/](ftp://iva07.ntua.gr/pub/4dkonto/WIAMIS-2005/defevent/papers/cr1099.pdf) [papers/cr1099.pdf](ftp://iva07.ntua.gr/pub/4dkonto/WIAMIS-2005/defevent/papers/cr1099.pdf). (2005). Accessed Aug 2014
- 8. Seiler, J., Kaup, A. (2008). Adaptive joint spatio-temporal error concealment for video communication. In Multimedia Signal Processing, IEEE 10th Workshop on, 8–10 October (pp. 229–234).
- 9. Hrusovsky, B., Mochnac, J., & Marchevsky. S. (2010) Temporal-spatial error concealment algorithm for intraframes in H.264/AVC coded video. In Radioelektronika, 20th international conference, 19–21 April (pp. 1–4).
- 10. Chung, Y., Chen, L., Chen, X., Yeh, W., & Bae, C. (2008). A novel intra-fame error concealment algorithm for H. 264/AVC. In IEEE third international conference on convergence and hybrid information technology, ICCIT'08, 11–13 November (Vol. 1, pp. 881–886).
- 11. ISO/IEC JTC1/SC29/WG11 (2006) Common test conditions for multiview video coding. JVT-U207, Hangzhou, China.