Error Concealment Techniques for H.264/MVC Encoded Sequences

Brian W. Micallef [1], Carl J. Debono [2], and Reuben A. Farrugia [3]

Department of Communications and Computer Engineering,
University of Malta, Msida, MSD 2080, Malta
[1] brian.micallef@ieee.org, [2] c.debono@ieee.org, [3] reuben.farrugia@um.edu.mt

Abstract

The H.264/MVC standard offers good compression ratios for multi-view sequences by exploiting spatial, temporal and interview image dependencies. This works well in error-free channels, however in the event of transmission errors, it leads to the propagation of the distorted macro-blocks, degrading the quality of experience of the user. This paper reviews the state-of-the-art error concealment solutions and proposes a low complexity concealment method that can be used with multi-view video coding. The error resilience techniques used to aid error concealment are also identified. Results obtained demonstrate that good multi-view video reconstruction can be obtained with this approach.

1 Introduction

The rapid development of information technology and its supporting hardware is pushing multimedia services to include more realistic 3D depth impressions of observed scenes. Such systems can be exploited in a number of new applications including entertainment, education, tele-presence, surveillance, medicine, remote manipulation, gaming and art [1]. Therefore, 3D Television (3DTV) and Free View-point Television (FVV) are expected to be the next revolution in TV broadcasting [2] and personal entertainment. Practical implementation of these technologies requires the deployment of multiple cameras to capture the same scene from different view-points at the same time [3]. The 3D effect is obtained by projecting two of the captured perspectives. This multiple camera setup creates a Multi-View Video (MVV) that brings new challenges when it comes to its delivery due to the huge amount of data that needs to be transmitted.

The significant increase in bandwidth demand requires higher compression by video codecs. However, the increase in compression efficiency generally comes at the cost of an increase in sensitivity to transmission errors [4]. Even a single bit in error may cause degradation in picture quality due to loss in bit-stream synchronization and loss of all encoded data which is dependent on the effected region. The wireless channel is highly susceptible to errors due to impairments. These degrade the quality of the transmitted multimedia sequences and therefore the design of reliable video coding systems is needed. This involves the development of error resilient tools and adequate channel coding techniques that limit the errors, their propagation and their influence. Furthermore, the use of error concealment techniques can be exploited to hide the effect of transmission errors in the video.

This paper first gives a review on the MVV concealment techniques found in literature. Different techniques that operate on part or the whole frame and use the depth frame or the view-point frame to conceal the lost data are presented. This work also evaluates the most appropriate error resilience techniques that can be applied with MVV Coding (MVC) and proposes a low complexity error concealment technique which uses only the neighborhood data to estimate the missing information for the corrupted slices in the frame.

The rest of the paper is organized as follows: Section II gives an overview of the MVC together with the type of errors found in error-prone channels, why these degrade the video quality and why error resilience and error concealment techniques become so important. Section III discusses the approach taken by some authors to conceal multi-view sequences while section IV proposes a low complexity error concealment technique that can be used with multi-view video coding. Section V presents some simulation results while Section VI provides a conclusion for this work.

2 Backgroud

2.1 Overview of Multi-view Video Coding

MVV sequences contain a huge amount of spatial, temporal and inter-view redundancy which can be exploited in order to improve the compression efficiency of the video codec. The spatial correlation between pixel values is reduced by using Intra coding algorithms, such as predictive coding. Temporal redundancy stems from the high video frame rates which make the scene look near-stationary between frames. Therefore correlation between successive frames is very high and this redundancy can be reduced by using Inter frame coding algorithms, such as Motion Compensation (MC) techniques. These Intra and Inter coding algorithms are used as in normal single-view videos and are specified in the H.264/AVC standard [5]. However, MVVs also have a large amount of inter-view dependencies which can be exploited for further compression. To remove this redundancy, MC techniques can be extended between view-points to obtain disparity vectors (DVs), thus obtaining a prediction structure similar to that shown in Fig. 1 [6]. This multi-view extension is defined in annex H of the H.264/AVC standard [5]. Coding MVVs in this way gives significantly better results compared to applying the single-view H.264/AVC solution on each view [7].

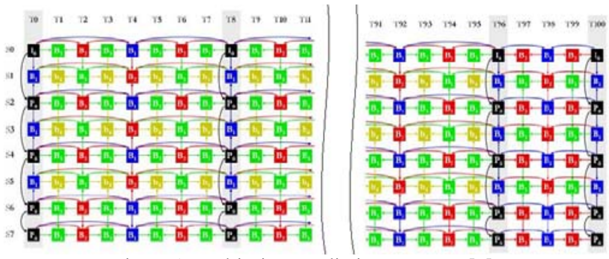

Figure 1. Multi-view prediction structure [7].

2.2 Transmission errors

Transmission of compressed video content over wireless channels is severely affected by packet loss. These result in bit-stream errors that can be classified into two types; random bit or burst errors. Since H.264/AVC has a spatial-temporal block-based structure, the produced bit-streams are very sensitive to transmission errors leading to spatial and temporal error propagation. Spatial error propagation occurs when there is a loss of synchronization in predictive or entropy decoding. Temporal and inter-view error propagation result due to the use of MC techniques between temporal and inter-view frames respectively [8]. The standard only considers the proper definition of the syntax and semantics of the bit-stream and it does not give any solution for erroneous bit-streams. It assumes that the lower network layer has the capability to detect and drop corrupted packets so that the decoder is only presented with intact packets [4]. Therefore, error resilient coding is required to limit the propagation of visual artifacts while error concealment can be adopted to minimize the visual artifacts caused by the lost slices.

2.3 Error resilience coding

Error resilient mechanisms are introduced at the encoder to make the transmitted video bit-streams more robust to potential errors and to facilitate error concealment at the decoder. The error resilient schemes adopted by H.264/AVC to mitigate the effect of packet loss are: slice coding that limits the spatial error propagation, insertion of regular Intra coded frames to limit the temporal propagation of the error, and Flexible Macro-block Ordering (FMO) that aids in error concealment. An overview of the performance of slice coding together with the effect of the cyclic-Intra coded period on H.264/MVC bit-streams can be found in [8]. Slice coding is also a requirement for most concealment methods since it prevents an entire picture from being lost [4]. FMO allows the Macro-blocks (MBs) to be grouped in slice-groups and each of them can be partitioned into one or more slices. Various FMO types are available [9], but the most efficient is the dispersed type FMO, where consecutive MBs are transmitted in different slice-groups to protect the neighborhood. FMO uniformly scatters possible errors to the whole frame to avoid error accumulation in a limited region [10]. Both slice coding and FMO increase the probability that a corrupted MB has distortion-free neighbors which can be used to aid concealment.

2.4 Error concealment techniques

Error concealment methods are used at the decoder to estimate the lost information by taking advantage of the inherent correlation that exists between the missing MB and the neighbourhood MBs. For multi-view concealment, each view can be concealed separately using only the single-view spatial and temporal error concealment techniques. However, for better error concealment, the neighbourhood view-point frame together with the depth video can be exploited since this has also a high correlation with the corrupted frame.

3 Multi-view Error Concealment

In literature, algorithms that consider the depth information for error concealment have been proposed. In [11], the authors propose a frame concealment method using shared MVs between the color and depth data. If the color frame is received corrupted, the MVs from the corresponding uncorrupted depth frame are used to form the concealed frame, and vice versa. If both corresponding frames are lost, then conventional single-view concealment algorithms are used to recover the lost frame. The authors in [12] consider the correlation between the color video and the depth video to propose a temporal error concealment technique for the lost MB. Another way to conceal a MVV is to use view interpolation [13] to obtain the lost frames. However, since MVV can be encoded efficiently without using the depth information, these techniques are not feasible as they require more processing resources and bandwidth.

In [14], the authors calculate a global DV, for the inter-view referenced frame, relative to the base view and this is transmitted with the anchor frames. When a frame is lost, the corresponding MBs in the dependent frame are located using the global DV. For each MB, the mode and the MVs are copied and MC is used to generate the concealed frame. If the matching MB does not contain MVs, then depending on the frames' type, it is either spatially concealed or a MV is estimated for it.

In [15], the authors propose another method for full frame concealment where, following a number of steps, different algorithms are used to estimate the MVs and DVs for each MB from the DV or the MV or both fields of the source frame. Then, a median filter is applied to filter both the MV and DV fields, to fill the empty spaces and to filter irregularities. These vectors are used with motion and disparity compensation to form the concealed frame. Finally, the resulting picture is filtered again using a median filter to fill the empty regions.

The latter methods consider the loss of a whole frame but the effect of transmission errors can be reduced by transmitting the frames in smaller slices. This is obtained through slice coding that result in slices with a smaller probability of error [4]. Since only the corrupted parts of the image will be concealed, the reconstructed quality of the frame improves. This increases the bandwidth required for MVVs [4] but provides better concealment since higher fidelity neighborhood data is available to predict lost regions. In literature, algorithms that consider concealment on multi-view slices can be found in [16-19]. The authors in [16] consider a four-dimensional frequency selective extrapolation process which exploits the surrounding information within the same image and information from temporal and view-point neighboring frames to estimate the missing samples in the corrupted area. In [17] and [18] the authors consider a method that identifies the corresponding region in the reference frame through feature points. This region is used together with the boundary pixels in a weighted sum to obtain the replacing MB. However, the latter two methods are highly dependent on the quality of the reference frames. They only consider the pixel values within the reference frames and they do not make use of MVs or DVs which are useful for better concealment especially in highly dynamic scenes.

In [19], the authors estimate the lost MBs by estimating the MVs and DVs from the neighborhood temporal and inter-view frames, respectively. The outer boundary of the lost MB is considered and a full search for the replacing MB that minimizes the boundary distortion error is searched in the temporal and the view-point frames using the Decoder Motion Vector Estimation (DVME) technique. This method is adopted from single-view concealment [20]. All the MVs and DVs which give a small distortion error are used for motion compensation and the formed MBs are weighted to form the replacing MB. This technique provides good concealment however the complexity of the decoder is drastically increased since the decoder must search for the optimal vectors for the replacing MBs.

Other work, such as [21-22], uses Forward Error Correction (FEC) schemes to introduce redundancies in the codewords to ease their correction. These are used to form unequal error protection on different MVV elements such that better protection is given to more important data. In [21], MBs are classified into slice-groups by examining their relative significance to the video and more important MBs are transmitted with better protection, by using the explicit type FMO [9]. In [22] the unequal error protection is formed by protecting different frame types with different levels of protection. Intra coded frames are the most protected, followed by temporal predicted frames, followed by inter-view predicted frame. Although this provides good error resilience, it increases the transmission bandwidth. If this is compensated for by decreasing the video bandwidth to accommodate it, the quality of the uncorrupted frames is reduced. With these schemes simple error concealment techniques can be used.

4 Proposed Multi-view Error Concealment

Since the H.264/MVC is an extension of the H.264/AVC, error resilience and error concealment techniques that are available for single-view video coding, can be adopted. To reduce the effect of transmission errors and to be able to transmit the video on a wireless channel, the frame has to be transmitted in smaller slices. To increase the probability of an intact neighbourhood, the dispersed type FMO should be used. These increase the bit-rate but are essential in obtaining good video quality on error-prone channels [4]. Further redundancies such as FEC are not essential since error resilience together with good error concealment is enough to provide acceptable video quality for low to high Packet Error Rates (PERs) [23] and these techniques increase the complexity.

For error concealment techniques, methods that operate on corrupted slices should be used and low complexity algorithms are preferred since the video must be decoded in real-time. Thus, to search for a replacing MB in the reference frames using the DVME technique is not practicable. However, a less complex technique [24] can be used. This gives good results that are comparable to the DVME, for low to medium PERs [25]. This method uses the outer boundary distortion measurement to find a replacing MB, like DVME, but instead of a full search in a defined area, it only examines the neighbourhood vectors. Statistically, these vectors are highly correlated. In MVC, the neighbourhood MBs can also be predicted using DVs. In this paper we use these DVs together with their respective reference view-point as potential candidates for the replacing MB. If one of the neighbourhood MBs is predicted by a DV, then the view-point reference frame is also used for concealment and this avoids a full search in all the reference frames. This forms a temporal/inter-view error concealment method where the best MC MB from the temporal or inter-view frame is selected.

The Intra-coded frames that do not contain neighbourhood vectors can be concealed using pixel-wise techniques. However, to reduce the complexity needed in finding the corresponding MBs in the reference frames, only the neighbourhood pixel values in the boundary are used. Therefore, the simple concealment technique [26] used for H.264/AVC concealment [4] can be adopted. This technique uses the boundary pixels of a corrupted MB and gives a weighted sum to the missing pixels depending on the distance of the corrupted pixel from the boundary.

5 Simulations and Results

The Joint Multi-view Video Model [27] (JMVM) was used to obtain the H.264/MVC coded bit-streams. The proposed error resilience and error concealment techniques were implemented in this reference software.

To demonstrate the performance of the proposed concealment algorithm, the *Ballroom*, a multi-view test sequences (8 views, VGA, YUV 4:2:0, 25Hz) [28] was used. The simulation parameters were chosen to have an encoder with low complexity and to be realistic for low delay real-time applications [29]. Slice sizes of 120 and 150 bytes (low delay and low probability of being in error [4]) where corrupted with five different error patterns and their results were averaged. Figures 2 and 3 show the results obtained with this concealment technique. These results show that applying the low complexity single-view error concealment techniques with the extension of inter-view prediction, can provide good MVV concealment even at high PERs.

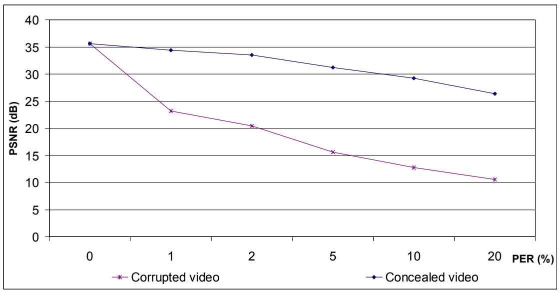

Figure 2: Objective results for the ballroom sequence


(a) (b) (c)

Figure 3: Subjective image quality comparison of frame 77 from view 1 with 10% PER and (a) no errors, (b) no concealment, (c) concealed.

6 Conclusion

This paper presented a review on MVV error concealment techniques and proposed a low complexity concealment technique. This method is adopted from the single-view error concealment techniques but enhanced with the view-point prediction extension. Results show that with this modification, it becomes adequate for MVV concealment too. Furthermore, low complexity error resilience techniques, such as slice coding and FMO, are deemed fundamental in supporting the error concealment solution.

Acknowledgement This work is partially funded by the Strategic Educational Pathways Scholarship Scheme (STEPS-Malta). This scholarship is partly financed by the European Union–European Social Fund (ESF 1.25).

References

[1] H. M. Ozaktas and L. Onural, Three-Dimensional Television, Capture, Transmission, Display, Springer, 2008, USA.

[2] L. Onural, T. Sikora, J. Ostermann, A. Smolic, M. R. Civanlar and J. Watson, *An assessment of 3DTV Technology*. [Online]. Available: http://www.3dtv-research.org

[3] W. Matusik and H. Pfister, "3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes," in *ACM Transactions on Graphics*, vol. 24, no. 3, pp 814-824, Aug. 2004.

[4] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in Wireless Environments," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 13, no. 7, pp. 657-673, July 2003.

[5] *Advanced Video Coding for Generic Audiovisual Services*, ITU-T Rec. H.264 – ISO/IEC IS 14496-10, March 2009.

[6] Y. S. Ho and K. J. Oh, "Overview of Multi-view Video Coding," in *Proc. ICSSIP'07*, pp 5-12, June 2007.

[7] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient Prediction Structures for Multiview Video Coding", *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 17, no. 11, pp. 1461-1473, Nov. 2007.

[8] B. W. Micallef and C. J. Debono, "An Analysis on the Effect of Transmission Errors in Real-time H.264-MVC Bit-streams," in *Proc. MELECON 2010,* Malta, April 2010.

[9] Y. Dhondt and P. Lambert, "Flexible Macroblock Ordering, an Error Resilience Tool in H.264/AVC," 5th FTW PhD Symp., Faculty of Engineering, Ghent University: paper no. 106.

[10] S. Kumar, L. Xu, M. K. Mandal and S. Panchanathan, "Error Resiliency Schemes in H.264/AVC Standard," *Journal of Visual Comm. and Image Rep.*, Vol. 17, no. 2, pp. 425-450, April 2006.

[11] C. Hewage, S. Warrall. S. Dogan, and A. M. Kondoz, "Frame Concealment Algorithm for Stereoscopic Video Using Motion Vector Sharing," in *Proc. ICME 2008*, Germany, June 2008.

[12] Y. Liu, J Wang, and H. Zhang, "Depth Image Based Temporal Error Concealment for 3D Video Transmission," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 20, no. 4, pp. 600-604, April 2010.

[13] C. L. Zitnick, S.B. Kang, M. Uyttendaele, S. Winderm and R. Szeliski, "High-Quality Video View Interpolation using a Layered Representation," *ACM SIGGRAPH and ACM trans. on Graphics*, pp. 600-608, Los Angeles, CA, USA, August 2004.

[14] S. Liu, Y. Chen, Y.-K. Wang, M. Gabbouj, M. Hannuksela and H. Li, "Frame Loss Error Concealment for Multiview Video Coding," in *Proc. ISCAS 2008*, pp. 3470-3473, USA, May 2008.

[15] C. Bilen, A. Aksay, G. Bozdagi Akar, "Two Novel Methods for Full Frame Loss Concealment in Stereo Video," in *Proc. PCS'07*, Lisboa, Portugal, Nov. 2007.

[16] U. Fecker, J. Seiler and A. Kaup, "4-D Frequency Selective Extrapolation for Error Concealment in Multi-View Video," in *Proc. MMSP 2008*, Australia, Oct. 2008.

[17] S. Knorr, C. Clemens, M. Kunter and T.Sikora, "Robust Concealment for Erroneous Block Bursts in Stereoscopic Images," in *Proc. 3DPVT 2004*, Sep. 2004.

[18] C. Clemens, M.Kunter, S. Knorr, and T. Sikora, "A Hybrid Approach for Error Concealment in Stereoscopic Images," in *Proc. of WIAMIS*, 2004.

[19] T. Chung, K. Song and C.-S. Kim, "Error Concealment Techniques for Multi-view Video Sequences," *Adv. in Multimedia Info. Proc. Pacific-Rim Conf. on Multimedia 2007, Lecture Notes in Comp. Science*, pp-619-627, Dec. 2007.

[20] S. Tsekeridou, I. Pitas, "MPEG-2 Error Concealment Technique Based on Block-matching Principle," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 10, no. 4, pp. 646-658, June 2000.

[21] S. Argyropoulos, A. S. Tan, N. Thomos, E. Arikan, and M.G. Strintzis, "Robust Transmission of Multi-view Video Streams Using Flexible Macroblock Ordering and Systematic LT Codes," in *3DTV CON'07*, Kos, Greece, May 2007.

[22] A. S. Tan, A. Aksay, G. B. Akar, and E. Arikan, "Rate-Distortion Optimization for Stereoscopic Video Streaming with Unequal Error Protection," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 14 pages, 2009.

[23] Y. Wang and Q. Zhu, "Error Control and Concealment for Video Communications: A Review," in *Proc. IEEE*, vol. 86, pp. 974-997, May 1998.

[24] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of Lost or Erroneously Received Motion Vectors," in *Proc. ICASSP'93*, vol. 5, Mar. 1993, pp. 417–420.

[25] H. R. Wu and K.R. Rao, Digital Video Image Quality and Perceptual Coding, Taylor & Francis Group, 2006, USA.

[26] P. Salama, N. B. Shroff, and E. J. Delp, "Error Concealment in Encoded Video," in *Image Recovery Techniques for Image Compression Applications*, Norwell, MA: Kluwer, 1998.

[27] MPEG & VCEG, "Reference software - Joint Multi-view Video Coding Model (JMVC 5.0.5)," JVT-AE207, June 2009.

[28] MERL sequences, [Online]. Available: ftp://ftp.merl.com/pub/avetro/mvc-testseq/orig-yuv/

[29] S. Wenger, "H.264/AVC over IP," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 13, no. 7, pp. 645-656, July 2003.