Robust Multi-View Video Streaming through Adaptive Intra Refresh Video Transcoding

Sagir Lawan[*], Abdul Hamid Sadka

Department of Electronic and Computer Engineering, Brunel University, London, UK.

Abstract

A multi-view video (MVV) transcoder has been designed. The objective is to deliver maximum quality 3D video data from the source to the 2D video destination, through a wireless communication channel using all of its available bandwidth. This design makes use of the spatial and view downscaling algorithm. The method involves the reuse of motion information obtained from both the reference frames and views. Consequently, highly compressed MVV is converted into low bit rate single view video that is compliant with H.264/AVC format. Adaptive intra refresh (AIR) error resilience tool is configured to mitigate the error propagation resulting from channel conditions. Experimental results indicate that error resilience plus transcoding performed better than the cascaded technique. Simulation results demonstrated an efficient 3D video streaming service applied to low power mobile devices.

1. Introduction

In everyday life, we communicate both through vision and language. Visual communication is the most natural way of human communication. More recently, multimedia 3D video system has been employed for visual communication. Most 3D video communication spans across the whole spectrum of government, corporate and private operational activities. Public and private organization are girding for 3D video information communication for them to get greater corporate realistic information. For example, Policemen across the world engaged in the ongoing ill-defined asymmetric global war on terror are equipped with body-worn cameras to acquire and transmit 3D video data over wireless links [1]. The fundamental purpose of wireless 3D video communication is the exchange of information between two or more parties. However, the best-effort wireless service cannot guarantee high quality delivery of 3D video packets due to 3D video sensitive to delay, jitter and packet loss that often occur in wireless environments. Unlike the conventional 2D video, 3D video is realized through MVV mechanism and contains redundancies in time, space and across views.

A typical MVV distribution scenario is simply explained in Figure 1. Referring to this figure, multiple cameras simultaneously capture the same scene from different angles. The processing part analyzes the acquired data, extracts features from data and compresses data for transmission. Despite having many advantages, wireless services pose a number of challenges that have prevented MVV from reaching their full potential. In the wireless channel, a large number of busty random errors severely degrade the quality of service (QoS) of MVV communication. Therefore, error resilience technique is required to control error propagation within frame and among views. The AIR error resilience tool [2-10] is an effective method to mitigate temporal error propagation. This is achieved by periodically encoding macroblocks (MBs) in INTRA

---

[*] Corresponding author. E-mail address: sagir.lawan@Brunel.ac.uk

Tel.: +44-7778-737614

mode, at the encoder. The encoder builds up a considerable data based on high motion MBs in the video stream known as refresh map. These MBs are encoded in intra (I) mode without any prediction.
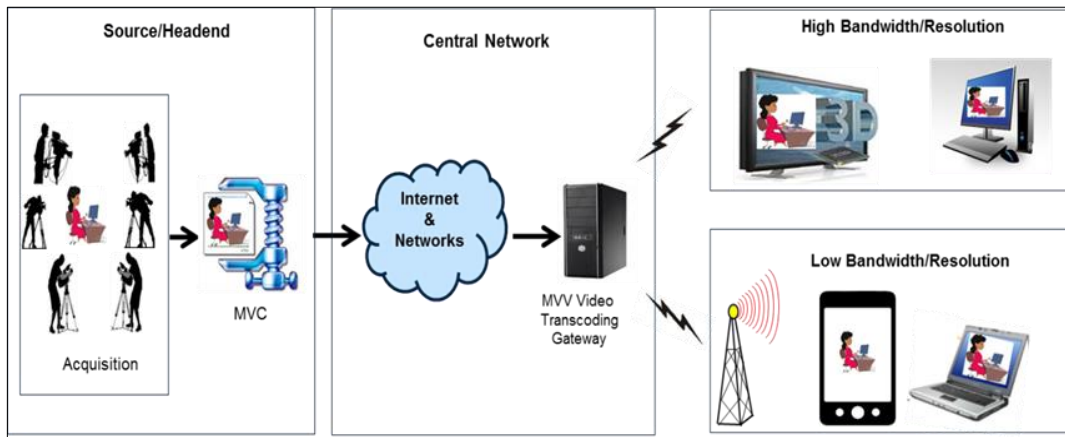

Fig. 1 Multi-view Video Communication System with Error Resilience Transcoding

Similarly, from the user's perspective, there are indeed some users who would want to watch the live streaming or record 3D video on their mobiles and tablets especially over internet. But, many of the mobile devices such as laptops, smartphones and the personal digital assistant (PDAs) are not capable of receiving bitstreams that are encoded by multi-view video coding (MVC) system. In such cases, it is desirable to design a MVV video transcoder that can dynamically adapt the MVV content to single-view video that is suitable for H.264/AVC decoders [11]. The MVV transcoder actually decodes the MVC bitstream, scales down the bitstream to single view and then encodes it for bandwidth-efficient transmission. In our propose, MVV transcoder, the design of transcoder is not done in isolation, but error resilience AIR tool is introduced to the already encoded video stream at a later stage of the transcoder to enable robust MVV transmission.

The rest of the paper is organized as follows. Section 2 provides a relevant literature review. Section 3 gives a brief introduction to adaptive intra refresh. Section 4 discusses the proposed MVV error resilience transcoding approach. Experimental results and discussions are shown in Section 5. Section 6 concludes the paper with a summary.

2. Review of Existing Related Work

Video transcoding and intra refresh error resilience is a study that has generated lots of interest. Hence, literatures on the subject matter are quite numerous. Attempt can only be made to review a few of these. In this section, an overview of video transcoding is outlined. A number of significant efforts have been placed in improving video transcoder. Historically, video transcoding has developed within a number of distinct content forms, such as bitrate, spatial, and temporal resolutions, format variation etc [12-21]. Most of the early transcoding techniques were based on open loop, closed loop or cascaded schemes. Figure 2 shows an example of cascaded transcoder which is a widely used method. It involves in decoding and then re-encoding the input video stream [22].
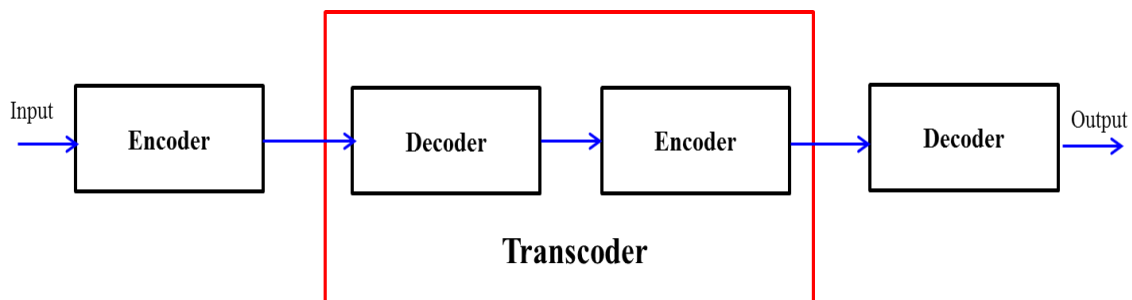

Fig. 2 Multi-view Video Communication System with Error Resilience Transcoding

In [23-30] suggested an error-resilient transcoder that can improve video quality in the presence of errors. The transcoder operates by maintaining the input bit rate over wireless channels and the performance was satisfactory. In the approaches in [31] various method to maintain quality of video transcoding for wireless channels stream based on analytical models were described. An appraisal on a cloud-based transcoding system for mobile devices for high definition video was also treated. The goal is to decrease both network traffic and the computational overhead by performing optimal VM allocation based on the workload prediction model.

Gaps, therefore, still exist to be filled. For instance, it is not yet explained how transcoding MVV to single view video with AIR which could enhance the application of 3D video to mobile device such as smartphone and PDAs. It is also not proved whether certain stages of MVV transcoder that converts MVC bitstreams into one single bitstream in H.264/AVC can operate with low rate mobile devices to display 3D videos. The impact of MVV on wireless environments has not been assessed and there is also the need to determine the categories of error resiliency in MVV transcoding. Finally, the specific types of MVV transcoding that characterised the conversion of 3D video on 2D video device are yet to be investigated. In seeking to achieve the objective of this study, these gaps will be adequately filled. In the meantime, the next section focuses on AIR process.

3. Adaptive Intra Refresh

Adaptive intra refresh has long been recognised in many studies as an effective error resilience tool that mitigates error propagation. Most of AIR error resilient features tend to increase the computational complexity of video encoder. However, the design of real-time wireless MVV transcoder requires both error robustness and low complexity. AIR error resilience is also useful during transmission of 3D video over unreliable channel condition. In time varying channel, it is imperative to adjust the error control parameters based on the estimated channel condition. Transmission errors emanating from burst and random bit errors can severely degrade the QoS of video communication over wireless channels [32]. These transmission errors often propagate among frames and other views as shown in Figure 3. One way to address the problem of error propagation in video transcoding is to protect the high motion area from transmission errors. It is evident that the high motion areas in a video frame are more susceptible to channel errors and more difficult to be concealed [33-35].
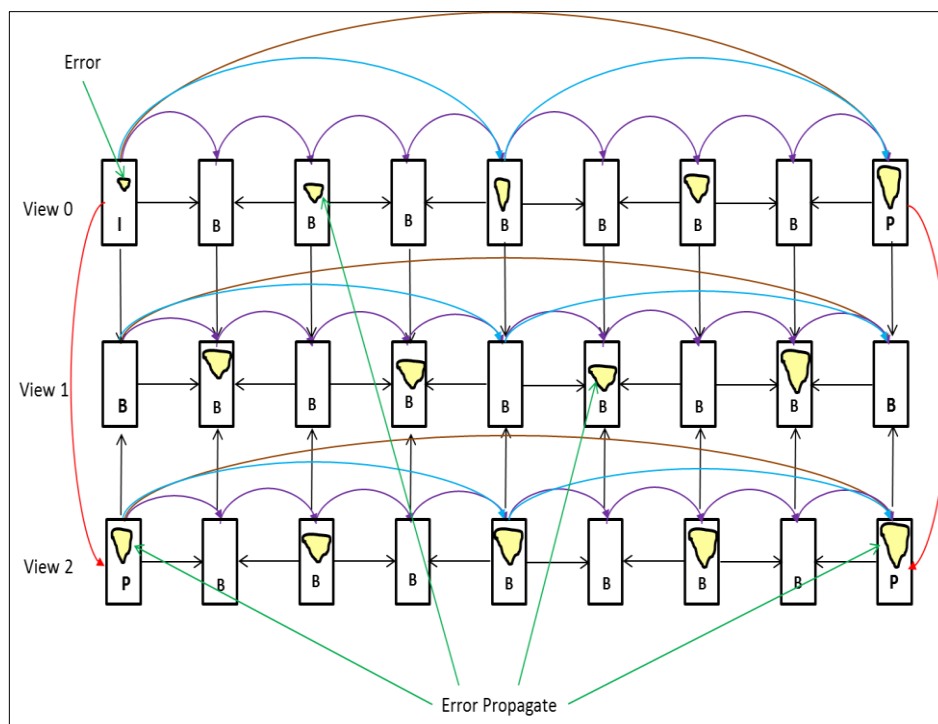

Fig. 3 Transmission errors propagate among frames and other views [32]

In this paper we illustrate that a video streams can be made more robust to channel errors by the periodic insertion of intra MBs in error-infected frame. For instance, a video frame (QCIF: 640 x 480 pixel) shown in Figure 4 consists of horizontally continues inter (P)/(B) mode MBs. The yellow area depicts high motion region that is prone to transmission error. For quick recovery of corrupted frame, a number of encoded intra mode MBs is periodically inserted to refresh the frames. In AIR error resilience process, the challenging issue is to determine the high motion region.
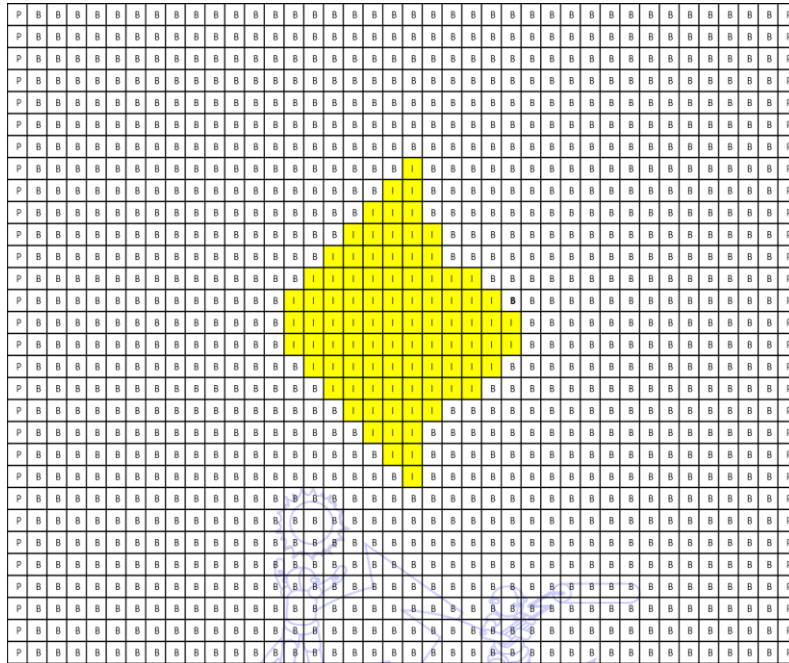
Fig. 4 AIR algorithm video frame (QCIF: 640 x 480 pixels)

This paper uses the sum of the absolute difference (SAD) to obtain the region with high motion. In this method, the SAD values are compared with the pre-determined threshold of the particular MBs. In the event that SAD value exceeds the pre-determined threshold, the encoder marks that particular MB as high motion MB. For each inter frame, a refresh map is generated to represent which MBs refer to or are referred to by full high motion MB. Table 1 shows an example of refresh map generation for two frames. If the given MB belongs to motion region, a bit in the refresh map is set to one.

Table 1 Refresh Map Generation.

A cyclic intra MBs refresh method among the various intra refresh techniques is employed and implemented in this work. An example of intra cyclic technique is illustrated in Figure 5 with a frame sequence (640 x 480). The varying level of refresh details has been segmented from top-to-bottom for efficient Intra coded MBs. This method comes into action starting from row #1, #2 up to the last row #30, and then it is reset to first row #0. In this process, one block consists of 40 MBs and by refreshing 30 times, the entire frame of 1200 MBs is intra refreshed. In [36], the increasing the number of MBs that are refreshed in each frame speeds up the recovery from errors. However, this is subject to a decrease in error-free quality at a given target bit rate.
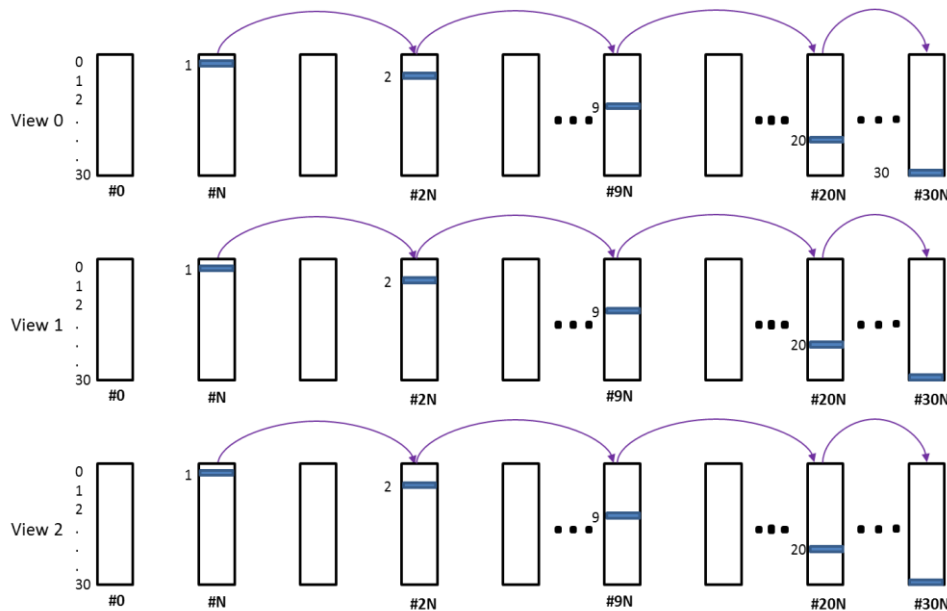
Fig. 5 A cyclic Intra MBs refresh method [32]

4. Error Resilient Multi-view Video Transcoding

The use of error resilience has greatly improved compressed video streams over communication networks [37]. The (MVV-AIR) error resilient transcoder proposed in this study is illustrated in Figure 6. The cutting edge approach consists of a central flow control block that extracts the rate distortion and channel conditions characteristics from the networks. It further, analyses the data and makes decisions about the threshold values mentioned in Section III. The threshold values are fed into the AIR block to generate the refresh map.
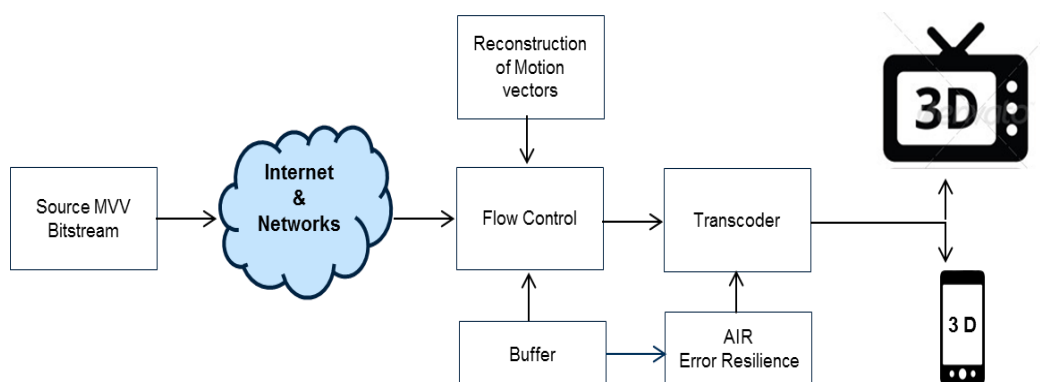
Fig. 6 AIR Error Resilience Multi-view Video Transcoder

The transcoding mechanism has been modified as illustrated in Figure 7. From the figure, it is shown that the MVV-AIR transcoder actually decodes the input MVC bitstream and scales down the bistream from MVV to single view. In this architecture, only VLD and inverse quantization are performed to get DCT value of each block in the decoder end.

VLD: Variable Length Decoding
Q_S^{-1} Inverse Quantization of the Source
Q_R :   Quantization of the Target Video
Q_R^{-1}:   Inverse Quantization of Target Video
DCT: Discrete Cosine Transform
IDCT: Inverse DCT
RPM   Reference Picture Memory
MRE: Motion Refine Estimation
MC: Motion Compensation
VLC   Variable Length Coding
AIR   Adaptive Intra Refresh
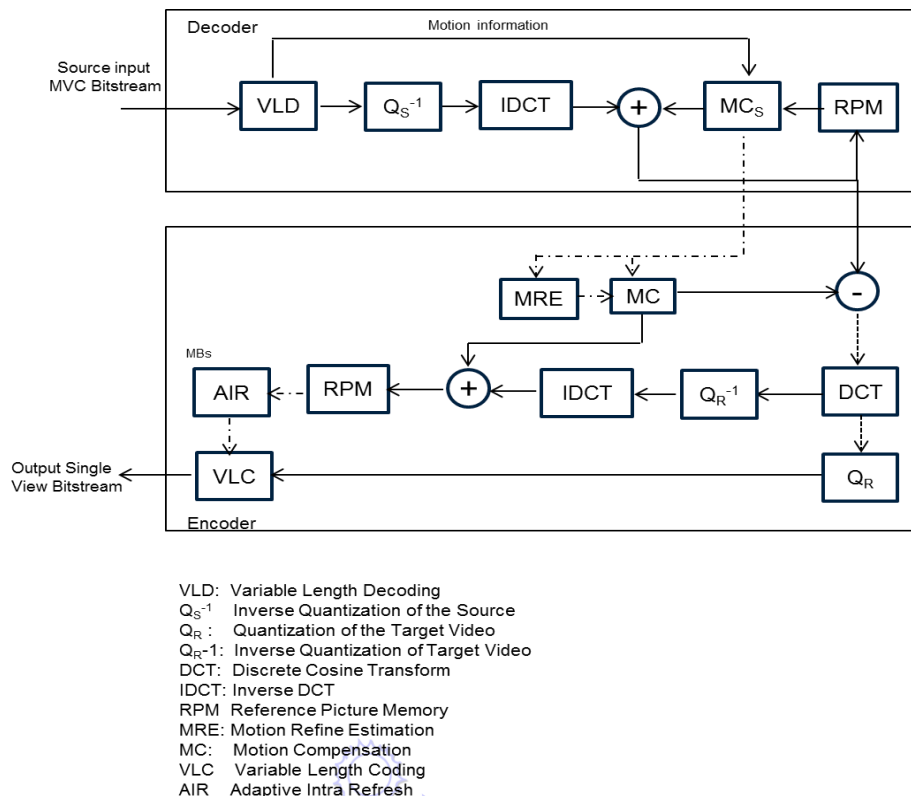
Fig. 7 Basic MVV –AIR Transcoding with motion information re-use scheme

At the encoder end, the motion compensated residual errors are encoded through re-quantization, and VLC. The reference frame memory in the encoder end stores the DCT values after inverse quantization, which is then fed into the frequency-domain MC module to reduce the drift error. Note that the effect of drift error leads to image gradual dissolution. To this end, drift error need to compensate through the processes of the second quantization Q2. Motion compensation is performed in the frequency domain using a MV reusing algorithm. Figure 8 shows a flow chart for the information re-use inter-coded MBs.
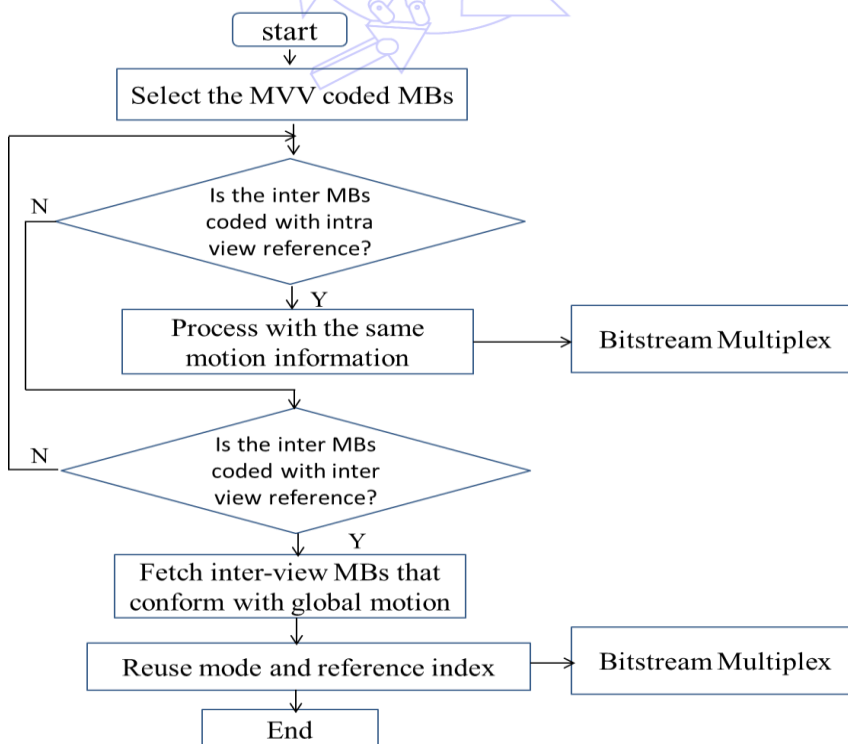
Fig. 8 A Flow Chart for the Information Re-Use

The transcoder is developed to improve the quality of compressed MVV delivery over error prone networks. However, in case of loss of information reports from the network, the transcoder searches its data base of refresh map to find out the affected MBs. The feedback channel informed the transcoder what MBs have been corrupted. Because the transcoder has the refresh map it knowns what MBs are affected by errors by virtues of the inter-MB dependencies. These MBs are re-encoded in intra mode making the system perform as an effective error resilience transcoding.

5. Experiments and Discussion

The experiments were conducted based on the JMVC version 8 Reference Software and MVC common test conditions defined in [38]. To assess the performance of proposed MVV-AIR transcoding, frames from view 0, 1 and 2 of Ballroom, Exit and Vassar were considered. The H.264/AVC code was then employed to generate error pattern for video sequence (640 x 480) with QP 22, 27, 32 and 37. The main goal of the experiment was to convert MVC bitstream to a single view bitstream with a view of producing a robust targeted bitstream for mobile devices. The proposed MVV-AIR algorithm is compared with H.264/AVC and Cascaded schemes whose inter prediction is a trade-off between the complexity and the rate distortion performance. The perceptual quality of each reconstructed view is measured in terms of peak signal to noise ratio (PSNR). The average experimental values for Ballroom, Exit and Vassar test sequences are recorded in Figure 9. According to these graphs, our propose method compared to cascaded technique show better performance approximately 0.4dB in Ballroom, 0.2 dB and about 1.0 dB for Exit and Vassar.
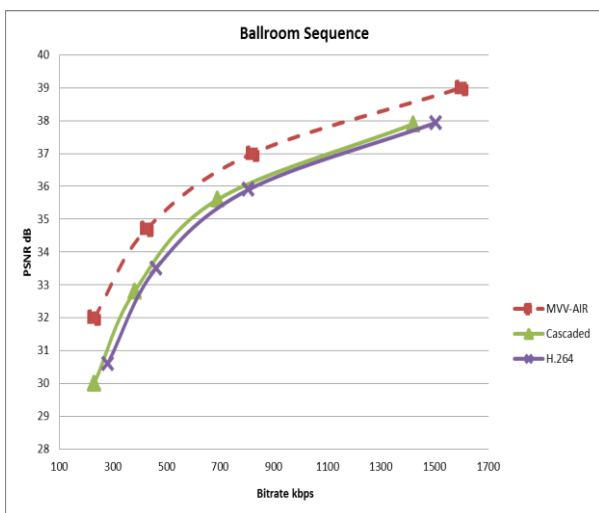
(a) Ballroom Sequence
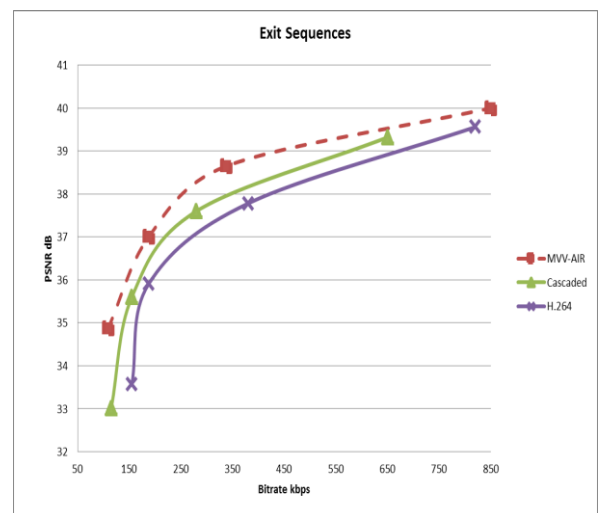
(b) Exit Sequence
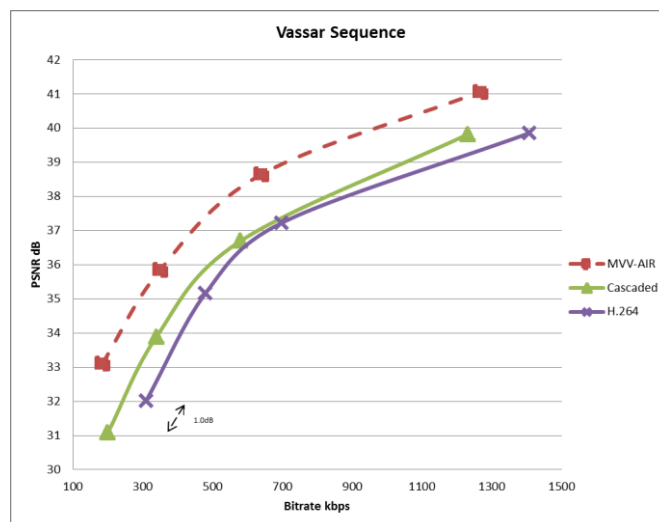
(c) Vassar Sequence

Fig. 9 Rate Distortion Performance

To overcome the effect of errors such as packet loss on a packet network or bit or burst errors on a wireless link, simulation was conducted with SIRANNON. Sirannon is simulator software that supports variety of network protocols such as RTP, RTSP, RTMP and HTTP [39]. Setting up Sirannon usually involves linking several modules in blocks, as shown in Figure 10. The scenario for the network conditions employed in our test replicates the characteristic of typical wireless channel suitable for mobile communication devices. Simulation for each bitrate and PSNR condition was conducted for 50 times.
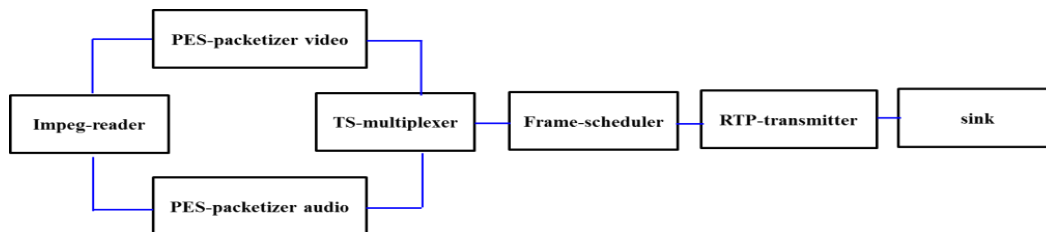
Fig. 10 SIRANNON network

Table 2 shows average PSNR of MVV-AIR, Cascaded and H.264 schemes under packet loss of 5%, 10%, 15% and 20%. It shows that our proposed MVV-AIR scheme provides better PSNR gain than the Cascaded scheme.

Table 2 PSNR Comparison

Sequence	Scheme	% Packet Loss Rate PSNR (dB)			
		5%	10%	15%	20%
Ballroom	H.264	30.27	30.17	29.79	29.06
	Cascaded	31.39	31.23	30.78	29.78
	MVV-AIR	**32.89**	32.03	31.33	30.04
Exit	H.264	29.67	29.37	29.11	29.01
	Cascaded	30.59	30.23	30.18	29.78
	MVV-AIR	**31.89**	31.63	30.33	30.04
Vassar	H.264	29.27	29.17	29.09	29.01
	Cascaded	30.56	30.22	30.08	29.78
	MVV-AIR	**32.05**	31.78	31.33	30.74

In Figure 11, the average PSNR performances comparison of the three methods for Ballroom, Exit and Vassar sequences are shown when the data is transported over wireless channel with bit errors. It can be seen that the performance of the AIR method are better, the reason is that MVV-AIR transcoding method has taken account of the changes of the channel conditions. It uses intra to refresh all the regions with high motion based on the channel state. Thus, in the MVV-AIR method, the MBs are periodically intra coded in a fixed order which mitigate the error propagation effectively.
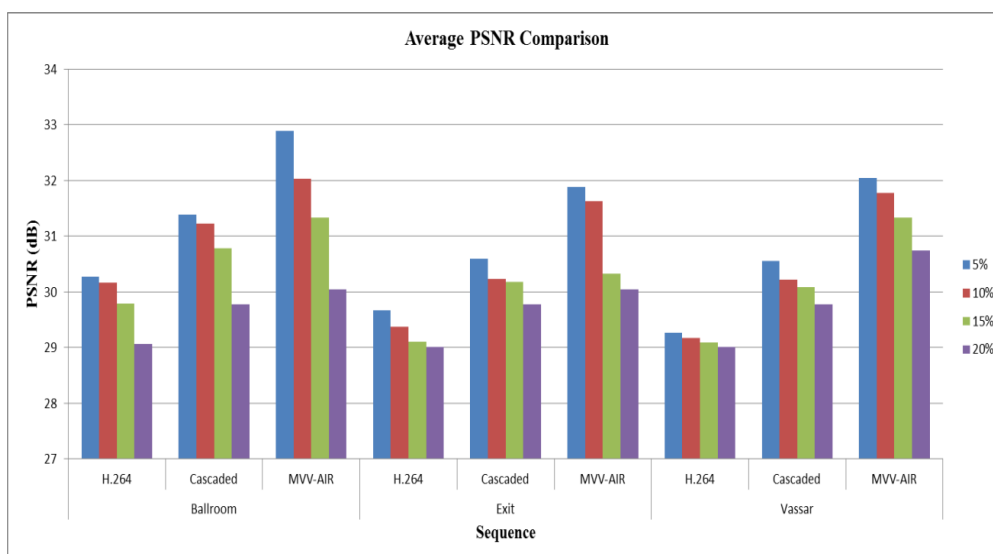
Fig. 11 Transmission over Internet

The ability to produce high-quality video output is a key attribute of MVV-AIR transcoding. The actual goal of injecting error resilience into the transcoder is quality preservation. It is obvious from Figure 12 that MVV-AIR results generate perfect perceptual visual object position. The loss of residual information in H.264 and Cascaded techniques as a result of channel errors has affected the full recovery of their frame information. While the subjective results of 5%, 25% and 20% are more relative to the objective result, however, the 10% result was similar to 15%. Hence it is not presented in this paper.

Fig. 12 Comparison of Subjective Quality for Ballroom, Exit and Vassar

In terms of time spent in the encoder, the MVV-AIR scheme takes triple less time to compare with the cascaded as shown in Figure 13. This also presents the fact that lower threshold value enables a faster and better performance. The relative computational complexity of the proposed MVV-AIR with cascaded method is much less. Even though MVV-AIR scheme has the lowest computation amount of time, the MVV-AIR scheme itself shows poor rate-distortion performance and visual quality.

Fig. 13 Rate Distortion Performance Ballroom

6. Conclusion

In this paper, we highlight MVV-AIR transcoding. The goal is to send compressed MVV bitstream over error prone wireless networks. Simulation results demonstrated an efficient 3D video streaming service applied to l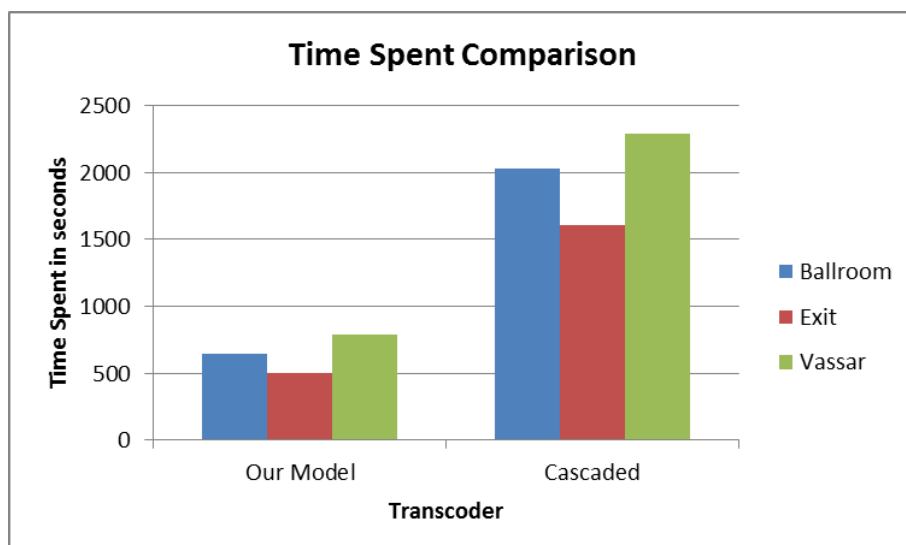ow power mobile devices. This is appropriately achieved through motion information reuse of both current view and reference view. The AIR error resilience is embedded with model transcoder to efficiently cope with error propagation in the target bitstream. Therefore, complying with this, our model has presented an improvement in the transcoding of multi-view video to single view.

References

[1] D. A. Harris, "Picture This: Body Worn Video Devices ('Head Cams') as Tools for Ensuring Fourth Amendment Compliance by Police," Texas Tech Law Review, vol. 43, no. 1, pp. 357-372, 2010.

[2] A. Vetro, J. Xin and H. Sun, "Error resilience video transcoding for wireless communications," IEEE Wireless Communications, vol. 12, pp. 14-21, 2005.

[3] I. A. Ali, Cross-Layer Enhancements for Error Resilient Video Delivery over Wireless Networks, 2012.

[4] Y. Zhang, W. Gao, H. Sun, Q. Huang and Y. Lu, "Error resilience video coding in H. 264 encoder with potential distortion tracking," International Conference on Image Processing (ICIP'04), IEEE press, Oct. 2014, pp. 163-166.

[5] M. Fleury, I. A Ali and M. Ghanbari, "Video Intra Coding for Compression and Error Resilience: A Review," Recent Patents on Signal Processing, vol. 4, pp. 32-43, 2014.

[6] 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," the 8th International Computer Engineering Conference (ICENCO'12), IEEE press, Dec. 2012, pp. 26-29.

[7] O. H. Salim and W. Xiang, "A novel unequal error protection scheme for 3-D video transmission over cooperative MIMO-OFDM systems," EURASIP Journal on Wireless Communications and Networking, vol. 2012, p. 269, 2012.

[8] Y. Zhou and Y. Chen, "Error-resilient video coding of H. 264/AVC based on network-adaptive intra refresh and reference selection refresh," Optical Engineering, vol. 49, pp. 077401-077401-11, 2010.

[9] Y. Sun, X. Zhang, F. Tang, S. Fowler, H. Cui and X. Dong, "Layer-aware unequal error protection for scalable H. 264 video robust transmission over packet lossy networks," the 14th International Conference on Network-Based Information Systems (NBiS'11), IEEE press, Sept. 2011, pp. 628-633.

[10] R. Talluri, "Error-resilient video coding in the ISO MPEG-4 standard," Communications Magazine, IEEE, vol. 36, pp. 112-119, 1998.

[11] S. Liu and C. W. Chen, "Multiview video transcoding: From multiple views to single view," Picture Coding Symposium, (PCS '09), IEEE press, May 2009, pp. 1-4.

[12] B. A. Adedayo, Q. Wang, J. M. A. Calero and C. Grecos, "Dynamic resource allocation engine for cloud-based real-time video transcoding in mobile cloud computing environments," IS&T/SPIE Electronic Imaging, SPIE proc. press, Feb. 2015, pp. 94000O-94000O-8.

[13] A. Cedillo-Hernandez, M. Cedillo-Hernandez, M. Garcia-Vazquez, M. Nakano-Miyatake, H. Perez-Meana and A. Ramirez-Acosta, "Transcoding resilient video watermarking scheme based on spatio-temporal HVS and DCT," Signal Process, vol. 97, pp. 40-54, 2014.

[14] C. F. Good, On The Fly Transcoding of Video on Demand Content for Adaptive Streaming, 2015.

[15] S. Liu and C. W. Chen, "3D video transcoding for virtual views," Proc. of the 18th ACM International Conference on Multimedia, 2010, pp. 795-798.

[16] A. Vetro, C. Christopoulos and H. Sun, "Video transcoding architectures and techniques: an overview," Signal Processing Magazine, IEEE, vol. 20, pp. 18-29, 2003.

[17] J. Xin, C. Lin and M. Sun, "Digital video transcoding," Proc. IEEE, vol. 93, pp. 84-97, 2005.

[18] I. Ahmad, X. Wei, Y. Sun and Y. Zhang, "Video transcoding: an overview of various techniques and research issues," IEEE Transactions on Multimedia, vol. 7, pp. 793-804, 2005.

[19] S. Moiron, M. Ghanbari, P. Assunção and S. Faria, "Video transcoding techniques," Recent Advances in Multimedia Signal Processing and CommunicationsAnonymous Springer, 2009, pp. 245-270.

[20] A. H. Sadka, "Video Transcoding for Inter-network Communications," Compressed Video Communications, pp. 215-256, 2002.

[21] S. Dogan, A. Sadka and A. Kondoz, "Tandeming/transcoding issues between MPEG-4 and H. 263, mobile and personal satellite communications 3," Proceedings of the Third European Workshop on Mobile/Personal Satcoms (EMPS'98), pp. 339-346.

[22] X. Zhang, Y. Li, J. Li, K. Zhao and T. Zhang, "Proximate control stream assisted video transcoding for heterogeneous content delivery network," IEEE International Conference on Image Processing (ICIP'14), IEEE press, Oct. 2014, pp. 2552-2555.

[23] I. Ahmad, X. Wei, Y. Sun and Y. Zhang, "Video transcoding: an overview of various techniques and research issues," IEEE Transactions on Multimedia, vol. 7, pp. 793-804, 2005.

[24] M. T. Beck, S. Feld, A. Fichtner, C. Linnhoff-Popien and T. Schimper, "ME-VoLTE: Network functions for energy-efficient video transcoding at the mobile edge," the 18th International Conference on Intelligence in Next Generation Networks (ICIN'15), IEEE press, Feb. 2015, pp. 38-44.

[25] C. Chen, C. Lin, H. Wei and Y. Chen, "Robust video streaming over wireless LANs using multiple description transcoding and prioritized retransmission," Journal of Visual Communication and Image Representation, vol. 18, pp. 191-206, 2007.

[26] S. Dogan, A. Cellatoglu, M. Uyguroglu, A. H. Sadka and A. M. Kondoz, "Error-resilient video transcoding for robust internetwork communications using GPRS," IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, pp. 453-464, 2002.

[27] T. Grajek, J. Stankowski, K. Wegner and M. Domanski, "Video quality in AVC homogenous transcoding," International Conference on Systems Signals and Image Processing (IWSSIP'14), IEEE press, May 2014, pp. 211-214.

[28] C. Kao, T. Huang, H. H. Chen and J. Wu, "Perceptully lossless video re-encoding for cloud transcoding," IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP'14), IEEE press, July 2014, pp. 301-305.

[29] D. K. Krishnappa, M. Zink and R. K. Sitaraman, "Optimizing the video transcoding workflow in content delivery networks," Proceedings of the 6th ACM Multimedia Systems Conference, 2015, pp. 37-48.

[30] A. H. Sadka, "Video Transcoding for Inter-network Communications," Compressed Video Communications, pp. 215-256, 2002.

[31] M. Song, Y. Lee and J. Park, "Scheduling a video transcoding server to save energy," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 11, no. 2s, article 45, 2015.

[32] L. Sagir and A. H. Sadka, "Robust Adaptive Intra Refresh for Multiview Video," International Journal of Computer Science, Engineering and Applications, vol. 4, no. 6, pp. 1-12, 2014

[33] A. H. Sadka, Compressed Video Communications. Halsted Press, 2002.

[34] S. Worrall, A. Sadka, A. Kondoz and P. Sweeney, "Motion adaptive intra refresh for MPEG-4," Electron. Lett., vol. 36, pp. 1924-1925, 2000.

[35] A. H. Sadka, "Error resilience in compressed video communications," Compressed Video Communications, pp. 121-176, 2002.

[36] S. Dogan, A. Sadka and A. Kondoz, "Tandeming/transcoding issues between MPEG-4 and H. 263," Mobile and Personal Satellite Communications 3, Springer Lodon, 1999, pp. 339-346.

[37] A. Vetro, J. Xin and H. Sun, "Error resilience video transcoding for wireless communications," IEEE Wireless Communications, vol. 12, pp. 14-21, 2005.

[38] Y. Su, A. Vetro and A. Smolic, "Common test conditions for multiview video coding," JVT-T207, Klagenfurt, Austria, 2006.

[39] A. Rombaut, N. Vercammen, N. Staelens, B. Vermeulen and P. Demeester, "Sirannon: Demonstration Guide," ACM Multimedia, vol. 9, pp. 1-4, 2009.