

ERROR RESILIENT FOR MULTIVIEW VIDEO TRANSMISSIONS WITH GOP ANALYSIS

A.B Ibrahim and A.H Sadka

Department of Electronic & Computer Engineering,
Brunel University, London, United Kingdom

ABSTRACT

The work in this paper examines the effects of group of pictures on H.264 multiview video coding bitstream over an erroneous network with different error rates. The study considers analyzing the bitrate performance for different GOP and error rates to see the effects on the quality of the reconstructed multiview video. However, by analyzing the multiview video content it is possible to identify an optimum GOP size depending on the type of application used. In a comparison test, the H.264 data partitioning and the multi-layer data partitioning technique with different error rates and GOP are evaluated in terms of quality perception. The results of the simulation confirm that Multi-layer data partitioning technique shows a better performance at higher error rates with different GOP. Further experiments in this work have shown the effects of GOP in terms of visual quality and bitrate for different multiview video sequences.

KEYWORDS

Multiview Video Coding, Group of Pictures, Error rates, Bitrate, and Video quality.

1. INTRODUCTION

Three-dimensional (3D) technology has transformed many fields of discipline, such as entertainment, communications, medicine, and many more. 3D technology can be perceived in a number of different ways. In this paper, we shall restrict our understanding to multiview video coding in this paper. Generally, the main concept of video coding is to exploit the statistical correlation between consecutive frames. The MVC extension of the H.264/AVC exploits the similarities between frames, simplifies the decoding process, and advances new features specific to multiview video coding [1]. Multiview video coding has emerged as advancement in video coding technology. The multiview video coding system enables efficient encoding of sequences captured from different cameras at different locations at the same time. The H.264 MVC codec takes as an input several synchronized bitstream captured from several different cameras and generate a single bitstream as an output for storage or transmission [2]. The work in [3] gives a detailed overview of the MVC standard. The structure of MVC is defined by a concept known as matrix of pictures (MOP). In this technique, each row consists of a group of pictures (GOP) normally captured by the base view and each column represents the time domain of the video.

2. BACKGROUND

The H.264/AVC international standard [4] has specified a coding standard of video data. H264 defines three picture types namely: I-frame, P-frame, and B-frame. In a standard reference multiview video encoder, all the pictures are encoded with a fixed GOP length depending on the settings and applications. The arrangement of these three picture types in a sequence is distributed statistically within the group-of pictures. The special type of I-frame at the beginning of a sequence, also known as an IDR frame serves as an entry point to facilitate random seeking or switching between channels. This can further be used in providing coding robustness to transmission errors [5], which are only coded with moderate compression to reduce the spatial redundancies in the multiview video sequence. I frames are generally larger than P and B frames, which means the less you have the longer the GOP size and the more compression you can get. But in multiview video content transmission, especially in error prone channels, very long GOP can have an adverse effect of propagating error spatially, temporally, and in interview direction. P frames are coded in an efficient way through the concept of motion compensation from either a past I or P frame which are mostly used as a reference to predict further. B frames have a very high compression ratio that requires the presence of both a past and future reference pictures for motion compensation.

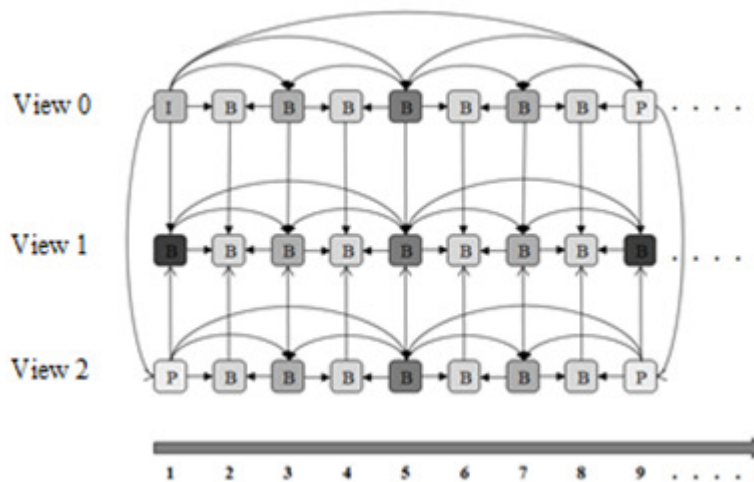


Figure 1. MVC prediction structure with GOP size of 8

Fig. 1 depicts a multiview video coding prediction structure with GOP size of 8, where I, P, and B represents the encoding of pictures in intra mode, predicted mode and bi-predicted mode respectively. The compressed multiview video data are highly sensitive to noise and information is loss due to the removal of statistical and subjective redundancy in the video by the compression scheme [6]. H.264/AVC employs variable length coding (VLC) to achieve higher compression gain. This type of predictive coding technique makes the video data highly sensitive to bit errors, and the effects of errors on the perceptual video quality can be quite severe. Thus, it is necessary to provide an effect technique and configuration settings that can make the MVV bitstream more robust to transmission errors and to improve the visual quality of the reconstructed multiview video [7]. The effectiveness of H.264/AVC coding depends on many coding parameters one of which is GOP size and its internal organization [8]. Most standard reference H.264 codecs use a fixed size for the GOP to encode video sequences. The GOP size can have different values as specified by the standard, however, once a given size is chosen, it

becomes applicable to the entire coding process and the corresponding standard decoder can be able to sort out the positioning of these frames during decoding process.

2.1. Concept of Data Partitioning in H.264/AVC

The H.264/MPEG-4 AVC standard is established to represent complete video information in a much lower level called the slice. A H.264 video slice consists of an arbitrary integer number of successive macroblocks that represent different types of video data [9]. The slice header conveys information common to all the MBs in the slice, such as the slice types which, determine which MBs types are allowed, and frame number that the slice corresponds to, reference picture settings, and default quantization parameter. The slice data section consists of a series of MBs that make a slice.

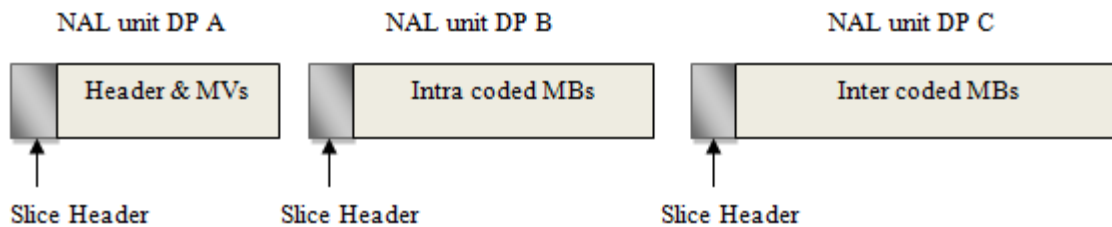


Figure 2. H.264/AVC Slice layout with data partitioning

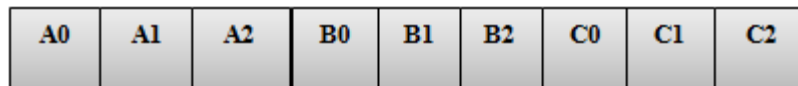


Figure 3. Multi-Layer data partitioning technique

In DP technique, MV and the residual information are separated by a boundary marker which is a uniquely decodable codeword. The codeword indicates the end of header information in a slice and the beginning of residual information [10]. Recent study on the concept of DP can be found in [11]. Data partition, nonetheless, creates more than one bit string (partitions) in every slice, and rearrange all symbols of a slice into a separate partition that have a close semantic relationship with each other Fig. 3. In H.264/AVC, when data partition is enabled, each slice of the coded bitstream is divided into three separate partitions with each of the partitions being from either type A, type B or type C partitions. Type A partition consists of header information, Quantization parameter (QP), Macroblock type, reference indices and motion vectors. The intra partition also called type B consists of the Discrete Cosine Transform (DCT) intra coded coefficients and the inter partitions also known as type C partitions contain DCT coefficients of motion compensated Inter-frame coded MBs. Type C partition in many cases is the biggest partition of a coded slice and yet the least sensitive to error because its information does not synchronise the encoder and the decoder [12]. Each partition is placed in a separate Network Abstraction Layer (NAL) unit and may be transmitted separately over a network. The use of both Types B and C will require a type A partition and not vice-versa.

2.2. Previous Work

The implementation of data partitioning technique for MVC is presented in [13]. A video slice without any ER mechanism may be affected by transmission errors that can lead to the loss of the entire information within the slice. Implementation of error resilience techniques such as data partitioning in the JMVC reference software is necessary because there is no provision for any ER technique in the MVC in the reference software.

Therefore, in order to analyse the performance of MVC in error-prone networks, implementation of a valid error resilience technique such as data partitioning as shown in Fig. 2 is employed and implemented in the JMVC 8.5 reference software. From the H.264 data partitioning technique, a video slice can be recovered when either partitions B or C, or both, are affected by transmission errors as long as the partition A is not affected or lost as a result of losing the header and motion information contained therein. It has been observed that the performance of H.264/AVC data partitioning technique in MVC is not too encouraging and further error performance improvements can be made through the introduction of the proposed multi-layer data partitioning technique depicted in Fig.3.

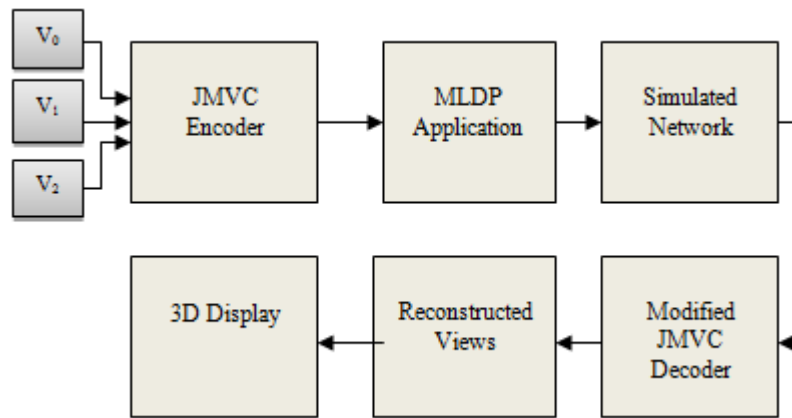


Figure 4. Flow diagram of the Multi-Layer DP technique

2.3. Multi-Layer Data Partitioning Technique

In an effort to provide error resilience to the MVV bitstream against losses in erroneous wireless network, a method is proposed that create a second layer of partitioning for each slice in the multiview video bitstream. Fig. 4 depicts the general architecture of the technique. The multiview video bitstream is data partition into a multi-layer partitioning structure for improved robustness against the transmission losses in an error-prone wireless network.

The partitioned bitstream is received by the modified JMVC reference decoder in order to decode and reconstruct the multiview video bitstream for viewing at the display. Multi-Layer DP adopts a mechanism that restructures a video slice as shown in Fig. 3. A_0 partition consists of the header information of frame 0 from view 0, and A_1 partition consists of the header and motion information of frame 1 from view 1 and A_2 partition consists of the header and motion information of frame 2 from view 2. B_0 consists of the residual information of intra coded MBs of frame 0, B_1 consists of the residual of intra coded MBs in frame 1 and B_2 consists of the residual

of intra coded MBs of frame 2 and C_0 is an empty partition, C_1 consists of residuals of inter coded MBs and C_2 consists of the residual of inter coded MBs of frame 2 and in that sequence it continues till nth view and nth last slice of the multiview bitstream. It is worth mentioning that partition C_0 is empty because there is no residual information of inter-coded MB's in frame 0 which is an intra-coded frame. I-frames are self-referential and do not require any information from other frames for prediction, so it consists of only intra coded MBs. The H.264 compliant encoder does not need to send empty partitions to the decoder because a standard H.264 decoder will assume missing partitions are empty partitions and are designed to handle the multiview bitstream accordingly [14].

The reference decoder is modified to cope with the losses in the bitstream due to errors in the wireless channel during decoding. The effects of transmission losses in a reconstructed frame can severely degrade visual perception by introducing artefacts. In order to support the MLDP technique more effectively and to minimise the effects of channel errors in the multiview video bitstream, a simple technique is employed. A simple error concealment technique is employed, which can replace the luma and chroma components of the corrupted MBs in a slice with that of the previous slice that is correctly received. Lost data in the bitstream can be concealed by copying the information from previously received error free slices. Frames that are generated by copying related video data in order to replace lost information are not always perceptually noticeable by a viewer that is an advantage of this technique especially in low-activity scenes [15]. In the approach, the multi-layer data partitioning technique can be supported with improved quality by employing frame copy error concealment, which works fairly well with MVC and is simple to implement. However, there are more complex techniques that use an elaborate approach to exploit the redundancy within the video frame in order to come up with a more efficient estimate of the lost data [3].

In MVC, the time first coding depicted in Fig. 5 is important and allows all views to be encoded and organized in a time domain for suitable transmission. The decoder can receive and reorder the bitstream in the right decoding order, which can allow it to decode all the pictures of different views in the same time domain. The time first coding ensures that the display of videos in the correct order. The implementation of frame copy concealment scheme in the reference decoder exploits the time first coding structure of MVC. It can be achievable because of the display nature of all the video frames across the views in the same time domain, which makes it easier to conceal missing pictures from previously received images in the reference list.

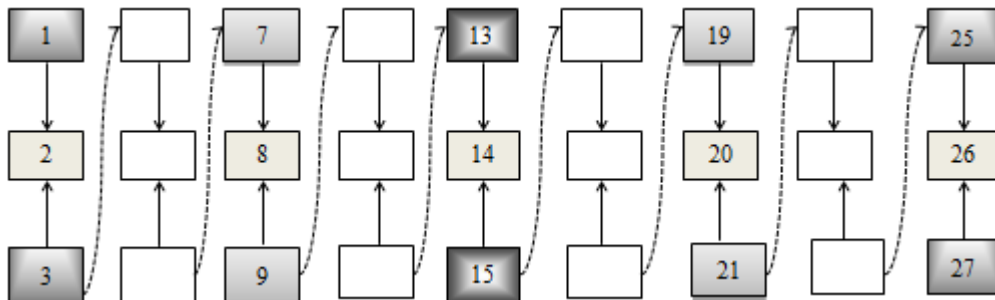


Figure 5. Time first coding [1].

Currently, the MVC reference decoder only accepts H.264 compliant bitstream and does not support the decoding of erroneous coded video sequence. In order to be able to decode the corrupted multiview video bitstream, the H.264/AVC frame copy error concealment technique is implemented in the JMVC reference decoder to adapt and cope with the losses within the bitstream. Frame copy error concealment technique is simple and usually quite effective in a video content where the motion is not large [16]. In Addition, the JMVC 8.5 reference codec has two types of reference frame lists that is also part of the standard and can be used to support frame copy error concealment in MVC. The first list is a reference list 0 which can be used for both P and B frames while reference list 1 is only applicable for B-frames. The main difference between the two reference lists is that list 0 utilizes the temporally earlier key frames (I or P) within the GOP in a sequence while in the case of the reference picture list 1; it utilizes temporally closer reference frames which can be a B frame [17]. Conceptually, reference list 1 can ensure smoother pictures because the frame to be copied is nearer to the picture to be reconstructed.

2.4. Proposed decoding scheme

The H.264/AVC frame copy error concealment technique is implemented in the JMVC reference decoder and further modified to decode the multi-layer DP bitstream with losses, as discussed in the previous section. The technique is optimized to reconstruct all the views successfully from the multiview coded bitstream with a higher level of quality in conformance with the standard [18]. Part of the reason and motivation to adopt frame copy error concealment technique in our work is its convenience to replace missing pictures, especially in the case of packet loss network.

The flowchart in Fig. 6 illustrates the implementation of frame copy error concealment technique. The technique can conceal lost information in the MVV bitstream with an improved perceptual quality based on experimental results presented in section 3.3. When the ML data partitioned bitstream is transmitted over the network and is received, it is first buffered and rescheduled back to the standard H.264 DP format for processing. Note that, the multi-layer data partitioning technique employed during source coding is only to make the multiview video bitstream more resilient to channel errors during transmission or streaming over the simulated wireless network. After successfully delivering the bitstream across the network, the received bitstream is rescheduled back to the standard H.264 data partitioned format for decoding. The decoder checks if the buffer is full then all the frames are sent directly for decoding. Also, note that all the slices are partitioned into three different partitions encapsulated into VCL NAL units of DP A, DP B and DP C respectively. The decoding of these types of slices is such that the loss of one partition might make another partition useless. To decode partitions B and C correctly, it is important for the H.264 standard compliant decoder to know how each macroblock is predicted within a slice. This information is stored in partition A as part of header information. Therefore, loss of partition A can render partitions B and C useless even when correctly received and decoded. Partition A does not necessarily require the information from partition B and C to be decoded correctly.

Equation (1) below shows how to compute the pixel value during motion compensation

$$E_{x,y} = I_{x,y} - P_{x,y} \quad (1)$$

Therefore, the pixel value or reconstructed value can be expressed as

$$I_{x,y} = E_{x,y} + P_{x,y} \quad (2)$$

Where $I_{x,y}$ is the pixel value and $P_{x,y}$ is the predicted value, and for each pixel, residual error $E_{x,y}$ is calculated. The values of x, y gives the coordinates of the variables pixel, predicted, and residual error respectively. The predicted value can be obtained from the motion vectors (in the case of inter coded MB) or intra prediction (in the case of intra predicted MB). It is known that, motion vectors and intra predicted modes are placed in partition A. The residual information is placed in the form of transform coefficients for intra-coded and inter-coded MBs in partition B and C respectively.

When the residual information is lost, then, $E_{x,y} = 0$ and the pixel value becomes

$$I_{x,y} = P_{x,y} \quad (3)$$

Because we lose some part of the video data in the form of residual information, the effect on the reconstructed video is usually grey scales around the pictures.

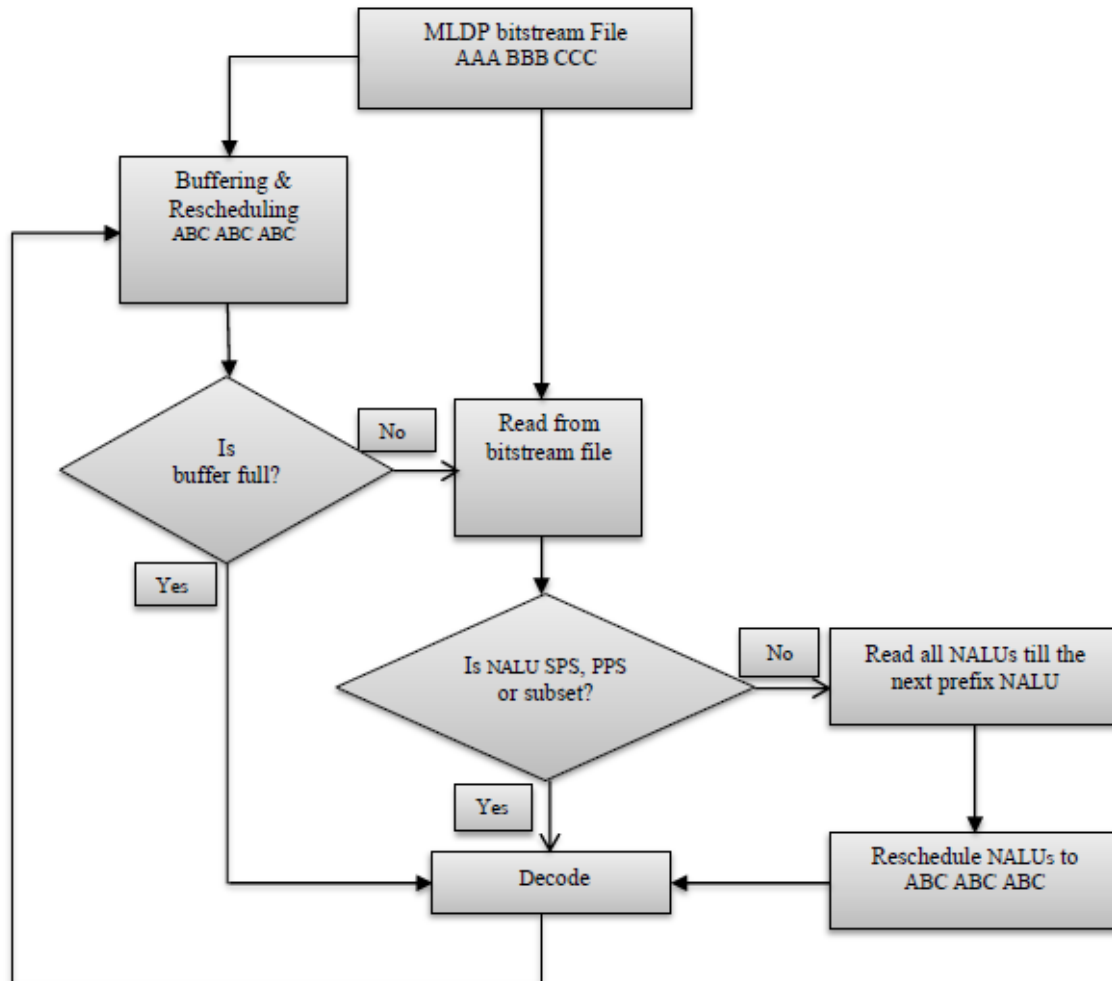


Figure 6. Decoding scheme for erroneous MVV bitstream

So, if only partition A is received correctly then the error concealment algorithm can utilize useful information, such as motion vectors to reconstruct the slice. However, if partition A is lost regardless of whether partition B or/and C is/are received. The frame copy error concealment is invoked by the decoder to replace the missing picture information by a previously received picture in the reference list. If the buffer is empty, then the NAL units are read from the MVV bitstream and the decoder determines whether it is a non-VCL NAL unit or VCL NAL unit. All non-VCL NAL units are sent directly for decoding while the VCL NAL units are all read until the next prefix NAL unit is detected and are rescheduled to the H.264 format before decoding. The whole process is restarted again through a looping system.

3. SIMULATION

To show the performance of 3D MVV bitstream over a wireless error-prone network, a number of coding and transmission experiments and simulations are performed in both JMVC 8.5 reference software and Sirannon network simulator [19]. This section describes the conditions used in the experimental setup.

3.1. Video Encoder Settings

The JMVC 8.5 reference software and simulations were configured as in [18]. All the experiments and simulations conducted in this work were tested on the MERL sequences, Ballroom, Vassar, and Exit. The 4:2:0 Chroma sub-sampling format was considered and a resolution of 640 x 480 pixels. The H.264/MVC codec as part of the standard supports the profile classifications. Our experiments are all based on the Extended Profile (XP) which is intended as video streaming profile. The XP profile has relatively high compression capability and some standard error robustness schemes to the video data losses and server stream switching capability. For simplicity and efficient decoder buffer management in our work, we employed three views and considered the first view to be the base-view and the second and third to be bi-predicted interview and forward predicted view respectively. Quantization parameter (QP) was carefully selected and set to 31 and for each experiment with different GOP, a suitable value for intra-coded frame was also carefully selected and inserted periodically in order to limit the temporal error propagation. Symbol mode is set on Content Adaptive Variable Length Coding (CAVLC) to support the DP in the extended profile, also one slice per NAL unit is considered as part of the H.264/AVC network friendly design [20]. Table 1 summarizes the key parameters used for setting up the JMVC reference software in the experiment.

3.2 Network Simulation Test bed

The Sirannon network simulator is a modular multimedia streamer which supports a wide variety of video formats and streaming protocols for use both in real time video streaming and offline simulation [21]. In this simulation, the offline mode is used. Fig. 7 shows the schematic to introduce packet loss, with different percentage error rate. The multiview coded sequence is read and packetized by `avc-reader` and `avc-packetizer`. The `avc-packetizer` is capable of packetizing the H264 compliant bitstream into packets suitable for real network and the simulated network as defined in RFC 3984. The gilbert classifier component has a random chance of introducing packet loss across the bitstream based on the Gilbert loss model.

Table 1. Key coding parameters setup

MVV test sequence	Ballroom, Exit, and Vassar
Number of Views	3
Frame Size	640x480
Frame rate	25Hz
Number of frames per view	250
Quantization parameter (QP)	31
Group of Pictures (GOP)	4, 8, 12 and 16
Entropy Coding	CAVLC
Intra period coding	Enabled
Bitstream format	Packet oriented bitstream

This is based on a simple concept of the transmission channel as having two states, Good state and Bad state. When the channel is the Good state, all the bits are transmitted correctly, which means that the channel is equal to perfect channel. On the other hand, when the channel is in a bad state, the channel is said to be in a binary symmetrical channel [22]. When these errors are introduced, the damaged stream is unpackitized by *avc-unpacketizer* block back into the original NAL unit format. The resulting coded stream which has lost some of the original frames based on the error rate selected is written to the basic component writer. Statistics component in the tool measures and generates at interval regular information about the passing stream and losses in the buffer. A special block called sink helps to terminate the program gracefully when the last packet of the sequence has passed through the sink.

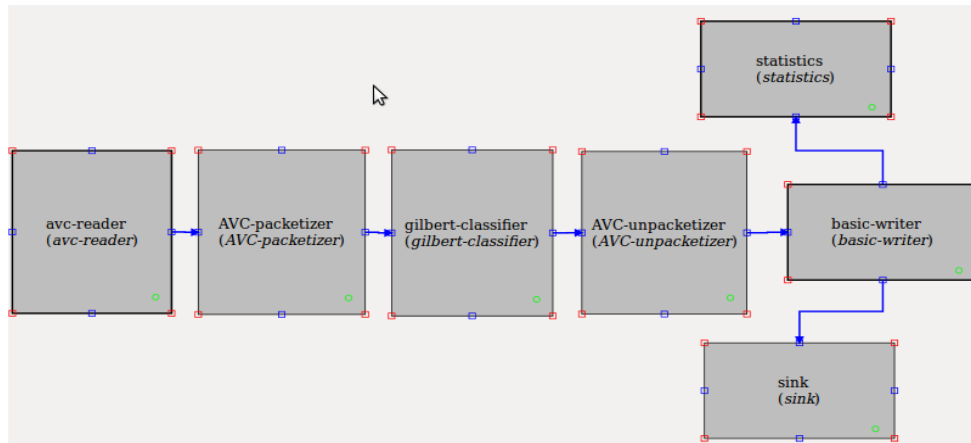


Figure 7. Network simulation test bed

3.3. Experimental Results and Analysis

This section describes the performance evaluation and results of the effects of GOP size on multiview video bitstream over an error prone channel. The values of GOP sizes used in the experiments are 4, 8, 12, and 16 respectively. Also, the error rates used are 0%, 1%, 5%, 10%,

15%, and 20% respectively. For every GOP size and error rate considered, ten different simulations are conducted, and the average results are generated. The perceptual quality of each reconstructed view is measured in terms of peak signal to noise ratio (PSNR) for all the different simulations and error rates used in the experiment. The PSNR values for ballroom and exit sequence are shown in table 2 and table 3 respectively for different loss rates and GOP sizes. The bitrate performance for different GOP sizes for the test sequences can be found in table 4. Fig. 8 and Fig. 9 show the quality performance evaluation for the H.264 DP and the multi-layer DP method for ballroom and exit respectively. The multi-layer DP has demonstrated a better and improved quality performance than the H.264 DP technique for different simulations, especially for higher error rates. Note that, video coding works either as fixed quality and variable bitrate and vice-versa. So in this experiment, various quality levels are examined for constant bitrate as recorded in table 4. The bitrate performance evaluation of the two techniques is reported in Fig. 10 and Fig. 11 for ballroom and exit test sequences respectively. The results demonstrate a very low bit rate cost to implement the H.264 DP technique in the reference software and further illustrates that the multi-layer data partitioning can be implemented with no additional bitrate.

From Fig. 12 and Fig. 13, the objective results of the experiment have revealed that a small number of GOP size means additional number of I-frames in the bitstream. The effect may consume more of bits because of the frequent occurrence of intra frames within the GOP. However, having more I-frames increases the multiview bitstream size. It can have a tendency of reducing the efficiency of the multiview video coding. Different applications can have different GOP requirements such as real time and offline applications each having a different latency or delay requirement [23].

3.3.1. Objective and Subjective analysis

Table 2. Numerical simulation results Ballroom sequence

Ballroom GOP4			Ballroom GOP8	
PLR (%)	H264 DP (dB)	H264 ML (dB)	H264 DP (dB)	H264 ML(dB)
0	35.45	35.45	35.16	35.16
1	34.53	34.93	34.67	34.72
5	28.54	28.90	30.28	27.97
10	24.73	24.37	26.82	24.96
15	21.04	22.93	21.35	21.90
20	18.65	20.04	18.09	19.04
Ballroom GOP 12			Ballroom GOP 16	
PLR (%)	H264 DP ((dB)	H264 ML (dB)	H264 DP ((dB)	H264 ML(dB)
0	34.99	34.99	34.83	34.83
1	34.74	32.83	34.38	33.41
5	30.42	30.10	30.42	31.82
10	24.24	24.22	24.61	25.59
15	20.94	21.63	19.23	22.52
20	18.23	20.09	16.01	19.17

Table 3. Numerical simulation results Exit sequence

Exit GOP 4			Exit GOP 8	
PLR (%)	H264 DP (dB)	H264 ML (dB)	H264 DP (dB)	H264 ML (dB)
0	37.58	37.58	37.36	37.36
1	37.08	35.50	36.82	36.82
5	31.15	33.94	35.52	35.20
10	29.78	27.28	32.43	30.61
15	27.92	27.09	20.32	27.23
20	22.35	21.53	23.02	23.53
Exit GOP 12			Exit GOP 16	
PLR (%)	H264 DP (dB)	H264 ML (dB)	H264 DP (dB)	H264 ML (dB)
0	37.25	37.26	37.11	37.11
1	36.67	34.96	35.87	36.84
5	30.07	33.77	30.75	30.52
10	21.73	29.02	26.37	25.09
15	25.76	22.91	22.17	23.25
20	23.82	24.61	21.11	21.95

Table 4. Bitrate simulation results for different test sequences

Ballroom		Exit		Vassar	
GOP	Bitrate (Kb/s)	GOP	Bitrate (Kb/s)	GOP	Bitrate (Kb/p)
4	1909.69	4	834.36	4	759.05
8	1619.76	8	722.12	8	657.69
12	1527.94	12	700.24	12	691.68
16	1374.75	16	535.23	16	572.54

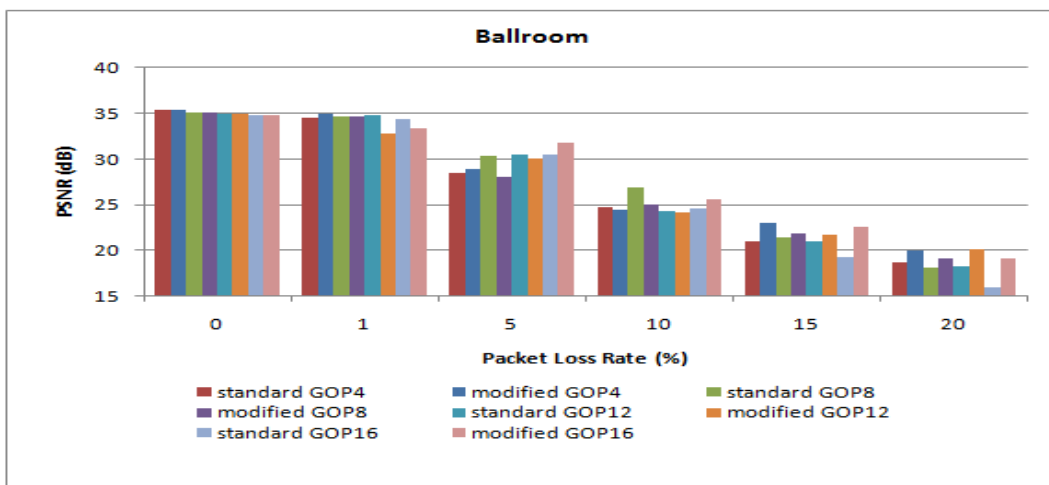


Figure 8. Ballroom quality evaluation with different GOP

From the objective result in Fig. 14, the results obtained illustrate that lower GOP size can slightly give a better perceptual quality for the multi-layer DP technique. This is because low GOP means more intra frames within the GOP with less prediction error which can result in a higher video quality. In video communications over-error prone environment, trade-off between perceptual quality and bitrate consumption is important and necessary [24]. In most cases, applications requiring a high level of quality in an error-prone network can have a higher bitrate in order to make the MVV bitstream more resilient to channel noise and that result in visual quality improvement. [25].

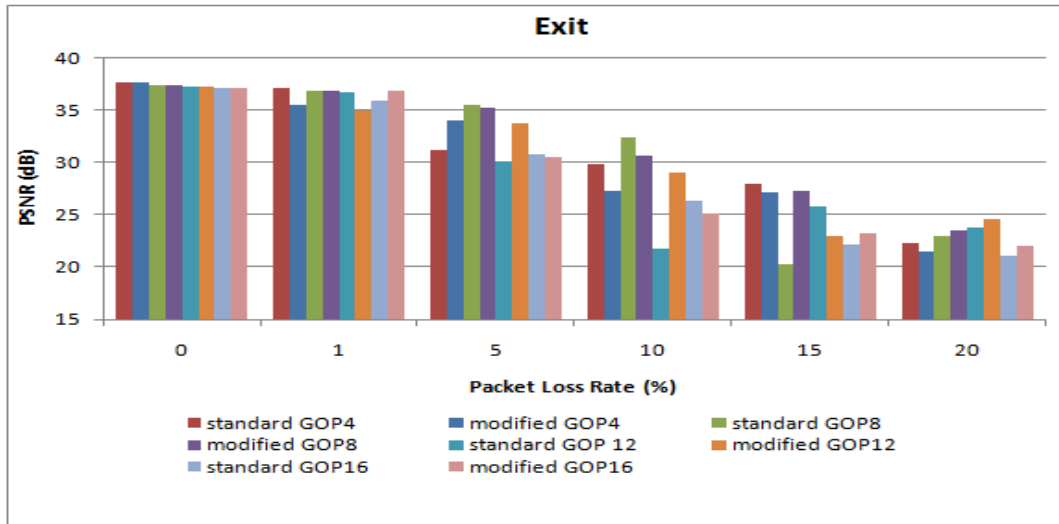


Figure 9. Exit quality evaluations with different GOP

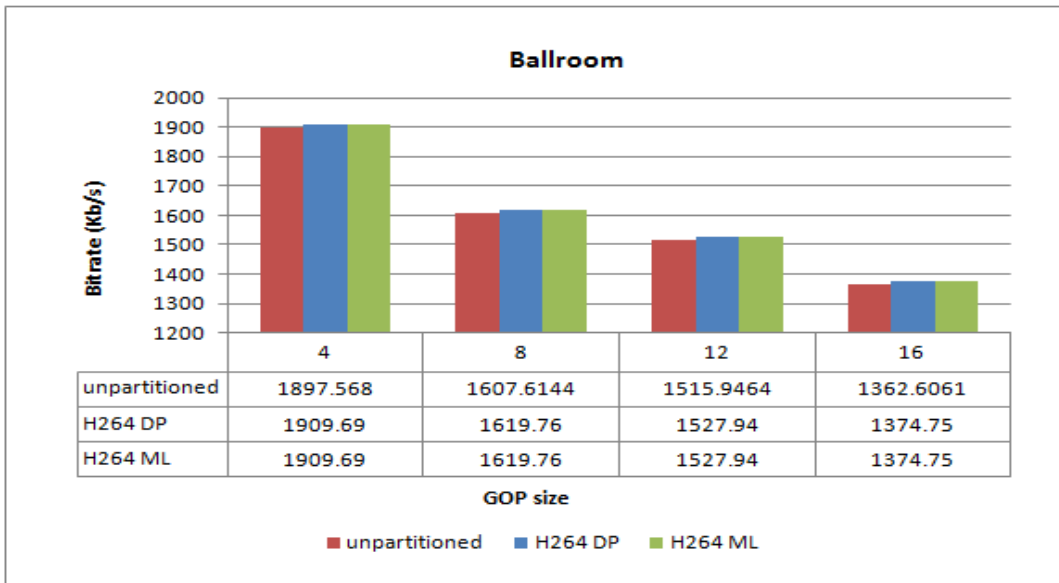


Figure 10. Bitrate performance for different GOP sizes for Exit

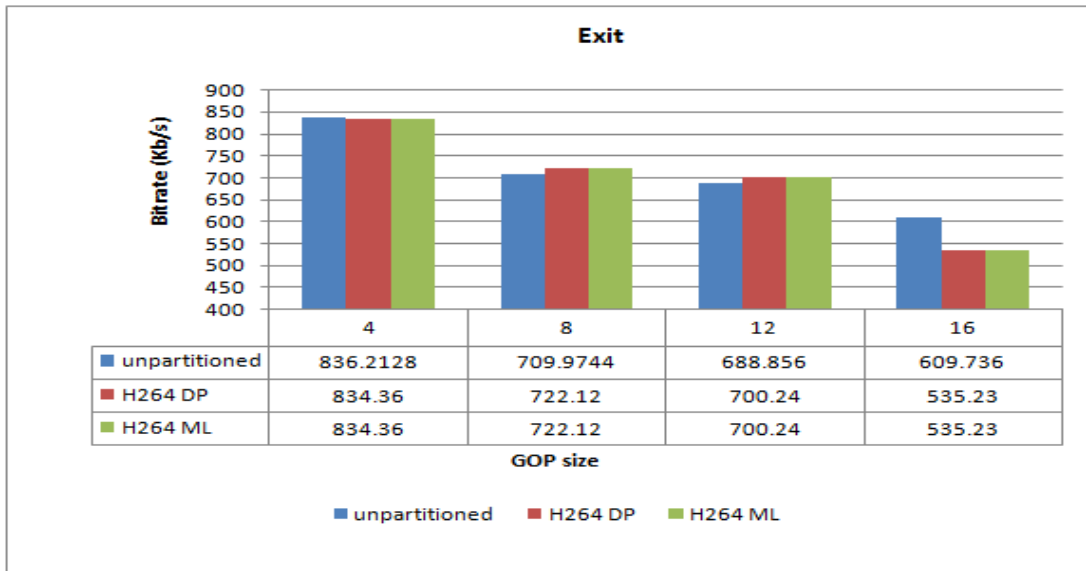


Figure 11. Bitrate performance for different GOP sizes for Exit

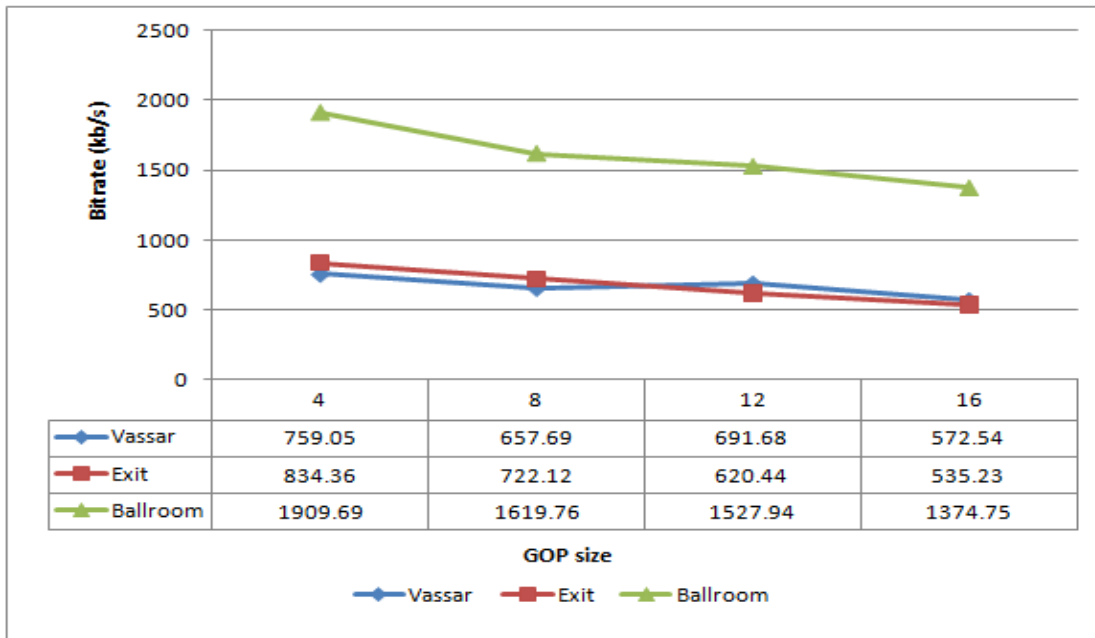


Figure 12. Bitrate performances for different GOP and test sequences

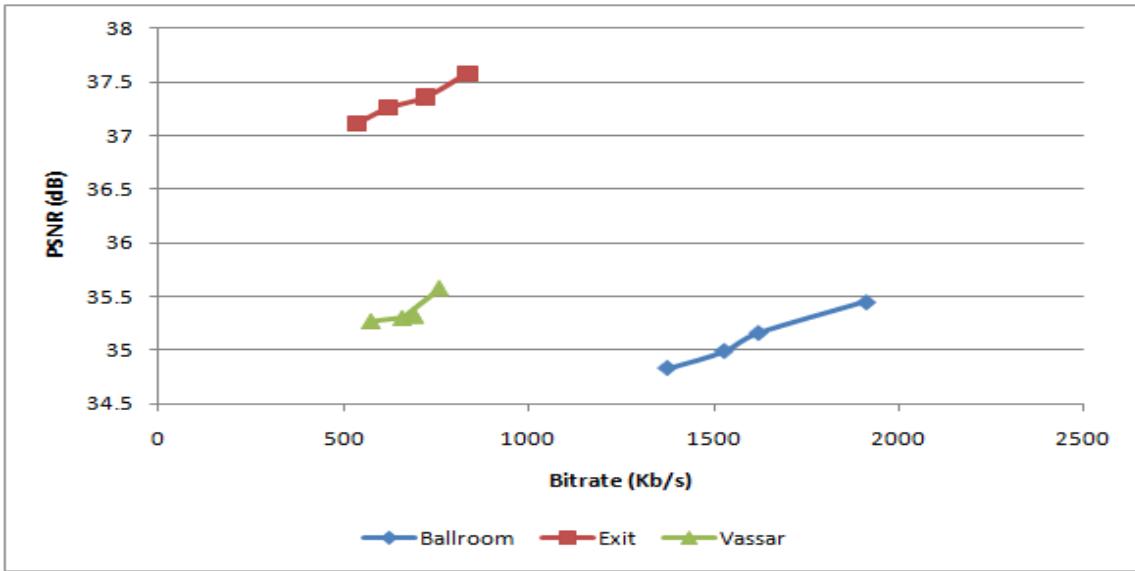


Figure 13. Relationship between quality and bitrate for different test sequences

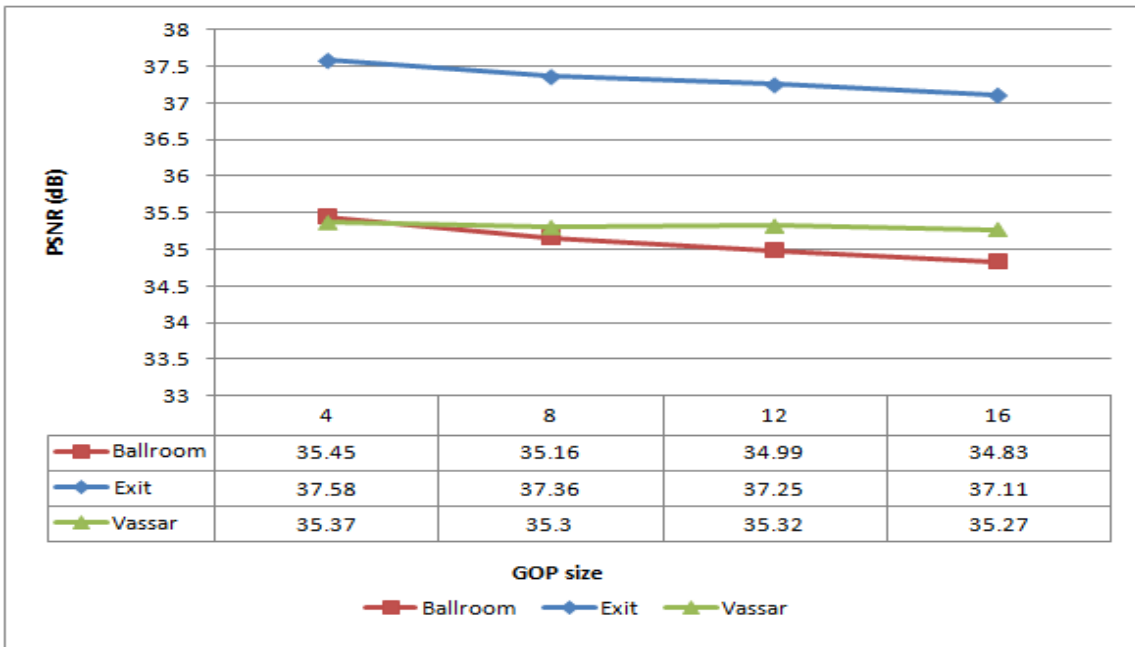


Figure 14. Quality evaluations for different GOP sizes and test sequences

The subjective result is presented for ballroom sequence in Fig. 15 for the three views. It can be observed that Multi-Layer DP technique can improve the perceptual quality performance compared to H.264 DP technique in all the views. The greyscale effect in Multi-layer DP technique is completely removed. When observed closely, the frames in multi-layer DP are not reconstructed with the best quality when compared with the original frames. The reason could possibly be the high error rate used in the network simulations and the limitation of the frame

copy error concealment to recover high losses. At 20% error rate, the multi-layer DP technique could recover most of the lost video information with improved quality compared to H264 DP technique at the same error rate and GOP size. Similarly, in the subjective results of exit test sequence shown in Fig. 16, it can be observed that the multi-layer data partitioning technique can improve the visual quality of the reconstructed video in a better way than the H.264 DP. Frame number 250 of the exit test sequence is selected for comparison and analysis at 20% error rate and GOP of 16. It is important to analyze the effects of error propagation within a GOP of the multi-layer data partitioned bitstream. In hierarchical GOP structure such as the one in multiview video coding, the reference decoder uses the I-frame in the base view and the anchor frames in the non-base view either directly or indirectly as reference frames for all other frames within the GOP.

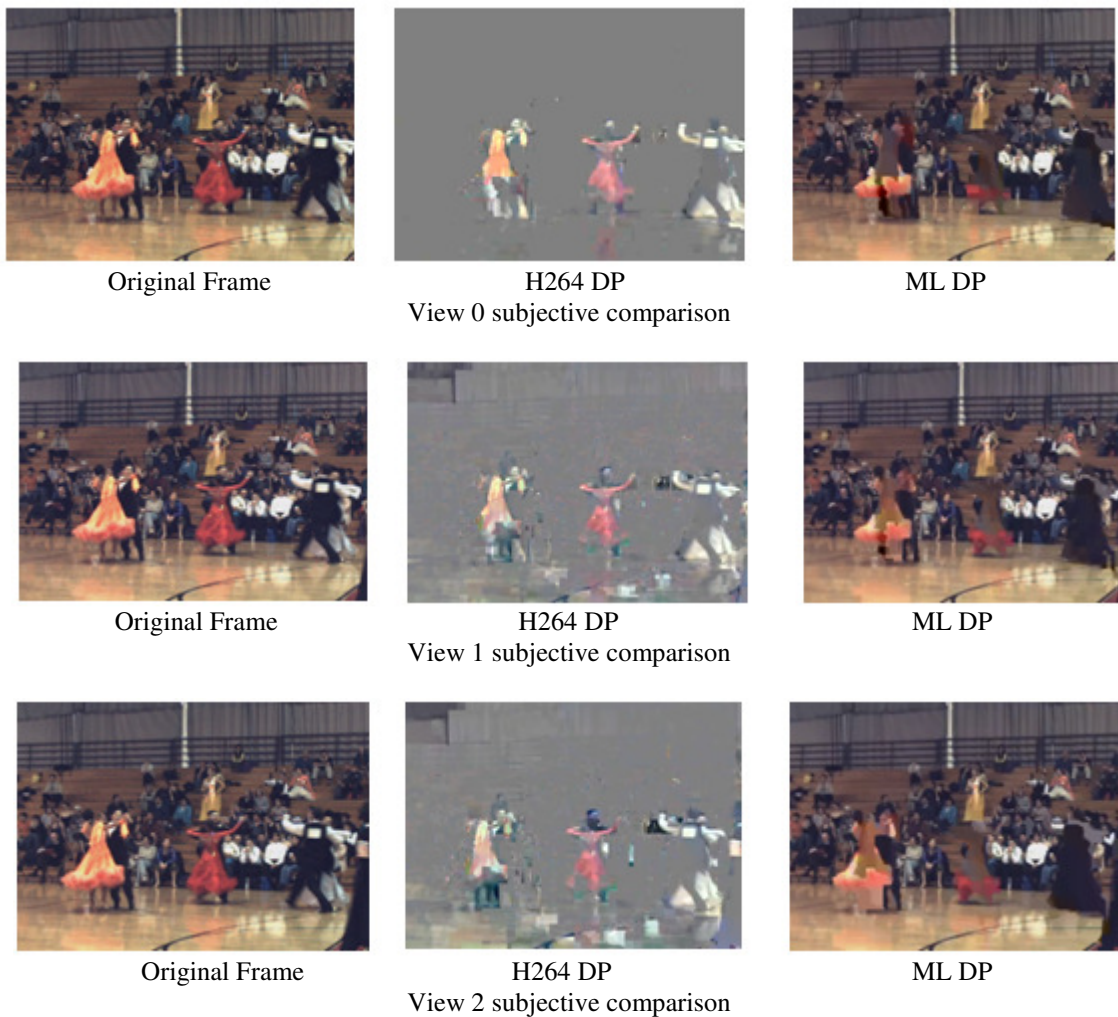


Figure 15. Ballroom subjective comparison for frame 121 at 20% PLR and GOP= 16

If an error occurs in the I-frame of view 1, it can result in artefacts that can continue to propagate throughout the GOP structure. The effect can be experienced in both temporal and interview domains until the next random access point. At this point, the decoder refreshes with the next intra coded frame in view 0 or the anchor frames in either view 1 and 2. It has been noticed that

losses within the I-frame that does not affect the header information such as intra coded MBs coefficient can also propagate errors throughout the GOP. P-frames are coded using motion compensation prediction from previous reference frames. From Fig. 1, anchor frame such as the one in view 2 is forward predicted from the I-frame in view 0, subsequent prediction of other non-anchor frames in both view 2 and view 1 takes reference from their preceding P-frame. Any form of loss in this frame can further propagate errors through the remainder of the GOP until the next refresh frame is received within the multi-layer partitioned bitstream. It can be highlighted that the impact of P-frame or anchor frame of view 2 can be almost as significant as losing an I-frame due to many interdependencies with other frames. Due to the hierarchical nature of MVC bitstream, anchor frame in view 1 that is interview predicted from view 0 and view 2 is used to predict other non-anchor frames temporally within the GOP. So the effect of errors is limited to view 1 only and is less severe than I and P-frames in the multiview video bitstream.

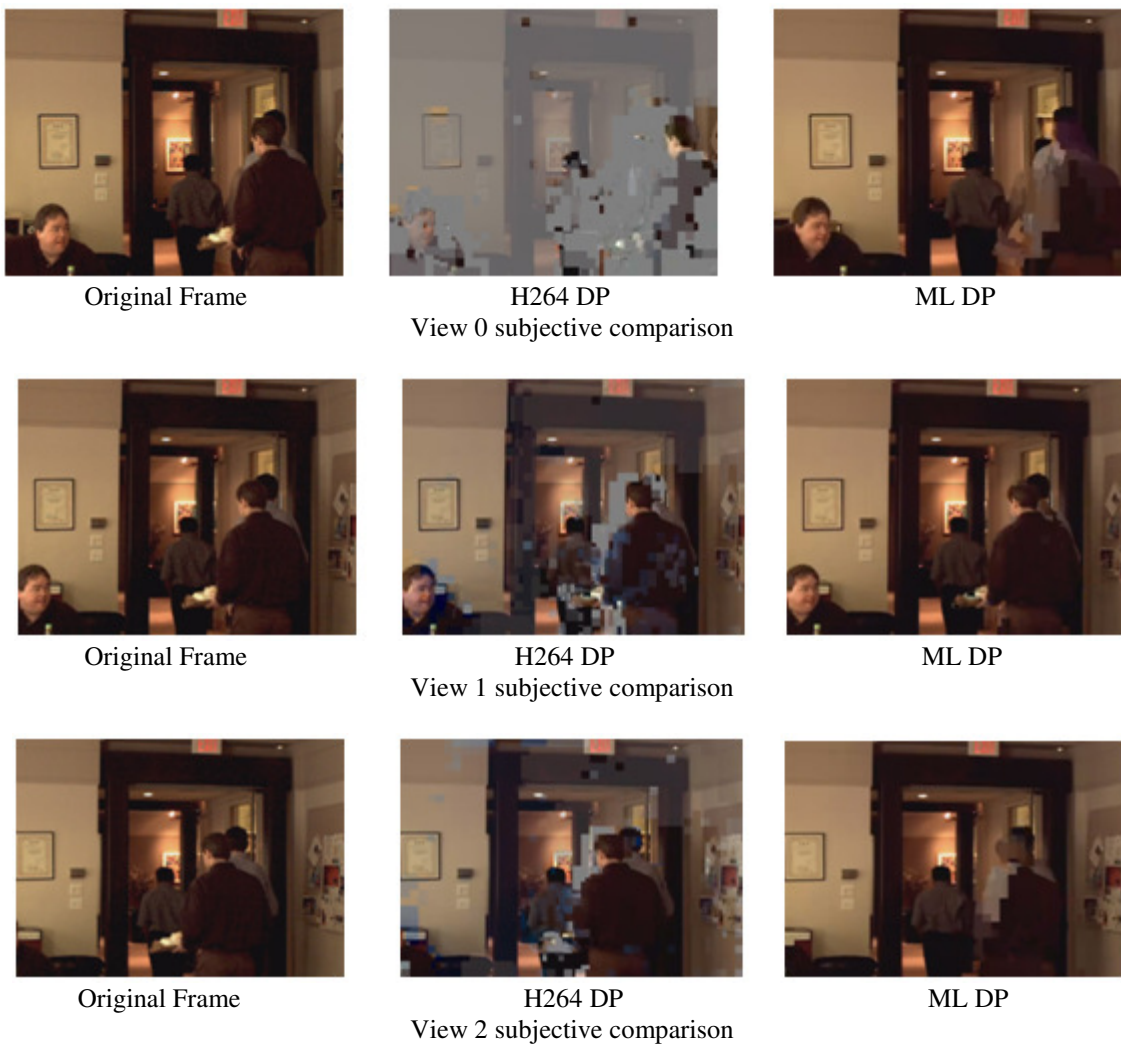


Figure 16. Exit subjective comparison for frame 250 at 20% PLR and GOP =16

4. CONCLUSIONS

The GOP within a video sequence is one of the key coding parameters that determine the video quality perception of the viewer, more importantly, the GOP size and the motion within the sequence. Large GOP size improves the compression efficiency, which can allow more or higher video content to be transmitted for a given bitrate. However, the effects of error propagation or artefacts due to transmission errors in an IP network might be more durable. It is necessary to wisely decide what GOP structure and size to support any application such as streaming or conversational videos. The work in this paper examines the effect of GOP size on erroneous multi-layer data partition bitstream when transmitted over error-prone networks. However, the study in this paper focuses on, and illustrates, the performance of the two algorithms for worst case scenarios. Two different techniques namely H264 DP and multi-layer DP are used to demonstrate this effect. The experimental results illustrate that the Multi-Layer DP technique can improve the visual perception of reconstructed videos for higher error rates within the allowable compression efficiency and bitrate. From the results obtained, we can assume and suggest that multi-layer DP technique can suitably be utilized for delivering multiview video content over bandwidth constraint and high error rate channel at a GOP size of 16. Please note that the work in this paper is not claiming to achieve a remarkable visual quality. We are proposing based on simulated results a different approach that can clearly improve the visual quality of multiview video in a very high error rate channel. Part of our future work is to optimize the multi-layer data partitioning technique by implementing error protection (e.g. forward-error correction) technique. The idea is to protect the multiview data from the high error rate of the channel. The decoder error concealment scheme was necessary because, without it, decoding of the error-prone MVV bitstream would have been impossible. The algorithm is modified to work in the JMVC reference software and be able to handle the multi-layer DP bitstream and conceal losses. From the experimental results obtained, it can be seen that the modified frame copy error concealment has considerably improved the performance of the multi-layer DP method including the JMVC reference decoder. However, there is a need to explore the hybrid error concealment technique that can fully exploit the redundancies between macroblocks in both spatial/temporal and interview direction. It is anticipated that better visual quality can be achieved when these techniques are implemented while considering the cost of bit rate and coding efficiency.

ACKNOWLEDGEMENTS

The authors would like to thank the Petroleum Technology Trust Fund (PTDF) for the research sponsorship.

REFERENCES

- [1] Y. Chen, Y. Wang, K. Ugur, M. M. Hannuksela, J. Lainema and M. Gabbouj, "The emerging MVC standard for 3D video services," *EURASIP Journal on Applied Signal Processing*, vol. 2009, pp. 8, 2009.
- [2] P. A. Akiki and H. W. Maalouf, "A two-stage encoding scheme for holographic data transmission," in *Multimedia and Ubiquitous Engineering (MUE)*, 2011 5th FTRA International Conference on, 2011, pp. 138-142.
- [3] M. Ebian, M. El-Sharkawy and S. El-Ramly, "Enhanced dynamic error concealment algorithm for multiview coding based on lost MBs sizes and adaptively selected candidates MBs," in *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012)*, 2013, pp. 435-443.

- [4] A. Hermans, "H. 264/MPEG-4 Advanced Video Coding," 2012.
- [5] T. Fang and L. Chau, "An error-resilient GOP structure for robust video transmission," *Multimedia, IEEE Transactions on*, vol. 7, pp. 1131-1138, 2005.
- [6] A. Vetro, J. Xin and H. Sun, "Error resilience video transcoding for wireless communications," *Wireless Communications, IEEE*, vol. 12, pp. 14-21, 2005.
- [7] S. Khan, Y. Peng, E. Steinbach, M. Sgroi and W. Kellerer, "Application-driven cross-layer optimization for video streaming over wireless networks," *Communications Magazine, IEEE*, vol. 44, pp. 122-130, 2006.
- [8] B. Zatt, M. Porto, J. Scharcanski and S. Bampi, "Gop structure adaptive to the video content for efficient H. 264/AVC encoding," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 2010, pp. 3053-3056.
- [9] I. E. Richardson, *The H. 264 Advanced Video Compression Standard*. John Wiley & Sons, 2011.
- [10] M. Sun, *Compressed Video Over Networks*. CRC Press, 2000.
- [11] L. Al-Jobouri, M. Fleury and M. Ghanbari, "Protecting H. 264/AVC data-partitioned video streams over broadband WiMAX," *Advances in Multimedia*, vol. 2012, pp. 10, 2012.
- [12] S. Wenger, "H. 264/avc over ip," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, pp. 645-656, 2003.
- [13] A. B. Ibrahim and A. H. Sadka, "Implementation of error resilience technique in multiview video coding," in *IEEE Southwest Symposium on Image Analysis and Interpretation, San Diego, California, 2014*, pp. 1-4.
- [14] Y. Dhondt, S. Mys, K. Vermeirsch and R. Van de Walle, "Constrained inter prediction: Removing dependencies between different data partitions," in *Advanced Concepts for Intelligent Vision Systems, 2007*, pp. 720-731.
- [15] O. Hohlfeld, "Stochastic packet loss model to evaluate QoE impairments," *PIK-Praxis Der Informationsverarbeitung Und Kommunikation*, vol. 32, pp. 53-56, 2009.
- [16] Y. Wang and Q. Zhu, "Error control and concealment for video communication: A review," *Proc IEEE*, vol. 86, pp. 974-997, 1998.
- [17] G. J. Sullivan, P. N. Topiwala and A. Luthra, "The H. 264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions," in *Optical Science and Technology, the SPIE 49th Annual Meeting, 2004*, pp. 454-474.
- [18] I. Rec, "H. 264 & ISO/IEC 14496-10 AVC," *Advanced Video Coding for Generic Audiovisual Services.ITU-T*, 2003.
- [19] A. Rombaut, N. Vercammen, N. Staelens, B. Vermeulen and P. Demeester, "Sirannon: Demonstration Guide," *ACM Multimedia, Beijing, China*, vol. 9, 2009.
- [20] T. Stockhammer and M. Bystrom, "H. 264/AVC data partitioning for mobile video communication," in *Image Processing, 2004. ICIP'04. 2004 International Conference on*, 2004, pp. 545-548.
- [21] N. Staelens, I. Sedano, M. Barkowsky, L. Janowski, K. Brunnstrom and P. Le Callet, "Standardized toolchain and model development for video quality assessment—the mission of the joint effort group in VQEG," in *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, 2011, pp. 61-66.
- [22] C. Jiao, L. Schwiebert and B. Xu, "On modeling the packet error statistics in bursty channels," in *Local Computer Networks, 2002. Proceedings. LCN 2002. 27th Annual IEEE Conference on*, 2002, pp. 534-541.
- [23] D. Wu, Y. T. Hou and Y. Zhang, "Transporting real-time video over the Internet: Challenges and approaches," *Proc IEEE*, vol. 88, pp. 1855-1877, 2000.
- [24] M. Flierl and B. Girod, "Generalized B pictures and the draft H. 264/AVC video-compression standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, pp. 587-597, 2003.
- [25] A. Aggoun, P. Amon, I. Arbel, A. Chernilov, J. Cosmas, G. Garcia, A. Jari, S. Keller, M. Mattavelli and C. Kontopoulos, "Multimedia delivery in the future internet," 2008.

AUTHORS

Abdulkareem Bebeji Ibrahim received the B.ENG. degree in electrical engineering from Bayero University Kano, Nigeria, in 2005, and the MSc. degree in satellite communication and space systems from the University of Sussex, Brighton, United Kingdom, in 2011. He is currently pursuing his PhD. degree in electronic and computer engineering at Brunel University London. His current research interests include error resilience and concealment for 3D multiview video coding and perceptual 3D multiview video quality.



Professor Sadka received the Ph.D. degree in electrical and electronic engineering from Surrey University, Surrey, UK, in 1997. He has nearly 20 years' worth of academic experience and a long track record of scientific leadership in the area of Video Processing and Communications. He is the former Head of the Department of Electronic and Computer Engineering at Brunel University and the Founding Director for the Centre for Media Communications Research.



He has over 200 publications in refereed journals and conferences 3 patents and a specialised book entitled "Compressed Video Communications" published by Wiley in 2002. To date, he has managed to attract circa £4M worth of research grants and contracts and has graduated 20 PhD students. He is widely supported by industry and runs his consultancy company VIDCOM. He is a fellow of the IET, a fellow of the HEA and a senior member of the IEEE.