Performance Evaluation of Error Resilience Techniques in H.264/AVC Standard

Ram Narayan Dubey
Masters in Communication Systems
Dept of ECE, IIT-R, India

Varun Gunnala
Masters in Communication Systems
Dept of ECE, NIT-W, India

Debashis Ghosh
Associate Professor
Dept of ECE, IIT-R, India

ABSTRACT
This paper summarizes some of the very important error resiliency schemes employed in H.264/AVC, like intra placement, slice structuring, effect of increasing the number of slice groups per frame and weighted prediction. In addition, we also focus on some other techniques like bi-prediction and weighted slice structuring which offer better performance than previous methods. The experimental results show that the choice of one or more of the above mentioned error resiliency schemes should consider the practical applications as well as network environment.

General Terms
Error resilience, video coding.

Keywords
H.264/AVC, Slice structuring, Weighted prediction, CABAC, CAVLC.

1. INTRODUCTION
Real-time transmission of video data in wireless and broadband environments is a challenging task, as it requires high compression efficiency and network friendly design. To deal with these two challenges, not only the video needs to be compressed very efficiently but also the compression scheme needs to provide some error resilient features to deal with the high packet loss probability. H.264/AVC utilizes predictive coding to achieve high compression ratio. Predictive coding also makes H.264/AVC bit streams vulnerable to transmission errors, as prediction incurs temporal and spatial propagation of the degradations caused by transmission errors. Due to the delay constraints of real-time video communication applications, transmission errors cannot usually be tackled by reliable communication protocols. Yet, most networks are susceptible to transmission errors. Consequently, error resilience techniques are needed to combat transmission errors in real-time H.264/AVC-based video communication.

The objective of this work is to do a comparative study of some of the important error resiliency techniques like intra placement, slice structuring and weighted prediction. We further evaluate the performance of some other techniques like bi-prediction, weighted slice structuring which offer better performance than previously discussed methods.

The remainder of the paper is organized as follows. In section 2 we discuss about various entropy coding schemes. In section 3 we discuss the various error resilience techniques. Section 4 summarizes the experimental assumptions and the results.

2. ENTROPY CODING SCHEMES
H.264/AVC specifies two alternative methods of entropy coding: a low-complexity technique based on the usage of context-adaptive switched sets of variable length codes, (CAVLC), and the computationally more demanding algorithm of context-based adaptive binary arithmetic coding (CABAC). When using the first entropy-coding configuration, which is intended for lower complexity implementations, the exponential-Golomb code is used for nearly all syntax elements except those of quantized transform coefficients, for which a more sophisticated method called context-adaptive variable length coding (CAVLC) is employed. When using CAVLC [1], the encoder switches between different VLC tables for various syntax elements, depending on the values of the previously transmitted Syntax elements in the same slice. Since the VLC tables are designed to match the conditional probabilities of the context, the entropy coding performance is improved from that of schemes that do not use context-based adaptivity. The entropy coding performance is further improved if the second configuration is used; this is referred to as context-based adaptive binary arithmetic coding (CABAC) [2]. Compared to CAVLC, CABAC can typically provide reductions in bit rate of 10–20 percent for the same objective video quality when coding SDTV/HDTV signals.

3. ERROR RESILIENCE TECHNIQUES
3.1 Intra Placement
Intra placement on the macro-block, slice, or frame level, is used primarily to combat drifting effects. Its main features are [3]

- In H.264 Intra macro-block prediction can be done even from predictively coded Inter macro-blocks. This improves coding efficiency, but degrades re-synchronization property of intra coding. When the Constrained Intra-Prediction Flag on the sequence level is set, it prevents this form of prediction and restores the re-synchronization property of Intra information.
- H.264 has two forms of slices that contain only Intra macro-blocks (1) Intra slices (2) IDR (Instantaneous Decoder Refresh) slices. IDR slices must always form a complete IDR frame i.e. all slices of an IDR frame must be IDR slices, and an IDR slice can only be part of an IDR frame. An IDR frame invalidates all short-term reference memory buffers, and hence, has a stronger re-synchronization property than a frame that contains only Intra slices.

3.2 Bi-Prediction
Prediction can be further improved by the use of bi-prediction at the cost of computational complexity. The B frames are coded based on a forward prediction from a previous I or P frame, as well as a backward prediction from a succeeding I

or P frame where backward prediction requires that the future frames that are to be used for backward prediction be encoded and transmitted first, out of order (Fig. 1). However, due to its high complexity it is generally not considered for low delay applications.

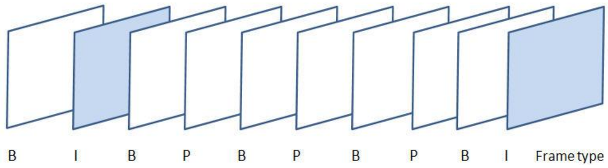


Fig 1: B-frame arrangement in Bi-Prediction

3.3 Slice Structuring: Flexible Macro-block Ordering (FMO)

A slice group is a subset of the macro-blocks (MBs) in a coded frame and may contain one or more slices. Within each slice in a slice group, MBs are coded in raster order. If only one slice group is used per frame, then all macro-blocks in the frame are coded in raster order. Multiple slice group (FMO) makes it possible to map the sequence of coded MBs to the decoded frame in a number of flexible ways. The allocation of macro-blocks is determined by the macro-block to slice group map that indicates which slice group each MB belongs to. The H.264 encoder intelligently groups MBs into a slice whose size is less than (or equal to) the size of the maximum transportation unit (MTU). Prediction beyond the slice boundaries is forbidden to prevent error propagation from intra-frame predictions. The slice structuring strategy thus aims at avoiding error propagation from a corrupted packet to subsequent packets. Fig. 2 to Fig. 7 illustrates the different types of macro-block to slice group maps supported by H.264/AVC. A detailed study of FMO can be found in [4].

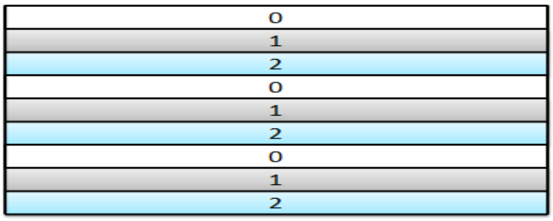


Fig 2. Slice group: Interleaved map (three slice group)

Fig 3. Slice group: Dispersed map (two slice groups per frame)

Fig 4. Slice group: Dispersed map (four slice groups per frame)

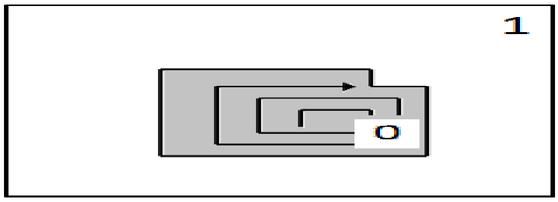


Fig 5. Slice group: Box-out map

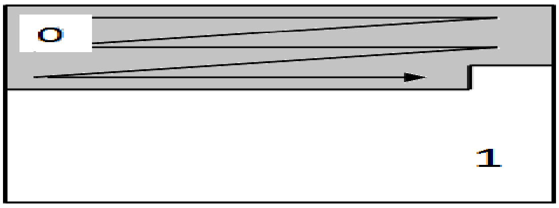


Fig 6. Slice group: Raster map

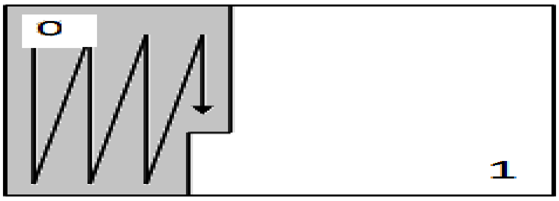


Fig 7. Slice group: Wipe map

The video error concealment schemes performs very well when the lost blocks are arranged in the checker board/scattered blocks fashion (2 slice groups per frame). This is depicted in Fig. 3. Performance may be further improved, if 4 slice groups per frame are used, because, in this case, the samples of a missing slice are surrounded by many samples of correctly decoded slices. Fig. 4 illustrates the slice structuring in case of dispersed FMO, when 4 slice groups per frame were adopted.

3.4 Weighted Prediction

A new innovation in H.264/AVC allows the motion-compensated prediction signal to be weighted and offset by an amount specified by the encoder. This can dramatically improve coding efficiency for scenes containing fades [5], and can be used flexibly for other purposes as well. Weighted prediction is available in the Main and Extended profiles, but not in the Baseline profile.

3.5 Weighted Slice Structuring

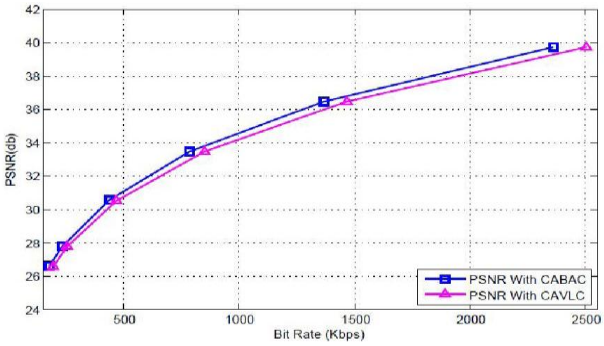
We have combined the weighted-prediction in Extended-profile as an error resilient feature with slice structuring. In normal slice structure, the lost blocks are concealed by their surrounding blocks; while in this case, concealment is done by taking the weights of surrounding blocks.

4. EXPERIMENTAL RESULTS

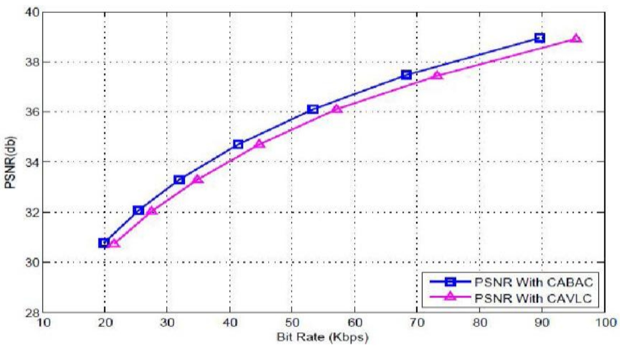
All the results were obtained through computer simulation using H.264/AVC reference software version JM 16.2 [6], for various error resilient video coding features employed by the H.264/AVC, with some necessary computations performed using MATLAB® 7.6.

In video coding, the system performance is measured by the average luminance PSNR. We begin our simulations with coding efficiency of H.264/AVC, which achieves better compression than all other previously existing video coding standards. In our simulations, we used three test video sequences, Bus, City and Foreman of length 150 frames each.

The sequence Bus is in the Common Intermediate Format (CIF, 352 × 288 frame elements, progressive) and the sequences City and Foreman are in the Quarter Common Intermediate Format (QCIF, 176×144 frame elements, progressive). All the test video sequences were taken from a standard source (http://media.xiph.org/video/derf/). Only the first frame was intra-coded and the rest were coded as inter-frame (IPPPP….). CABAC and CAVLC both were used as entropy coding methods. The coding efficiency was measured in terms of average bit rate savings for a constant peak signal to noise ratio (PSNR). Fig. 8 shows the PSNR of the luminance component versus the average bit rate for the two test sequences, out of which the former was encoded at 30 Hz, and the later at 10 Hz. We observe from Fig. 8 that the compression performance is further improved if the CABAC entropy coding configuration is used. Compared to CAVLC, CABAC can typically provide reductions in bit-rate of the order of 10–20 %for the same objective video quality when coding SD/HDTV signals.



(a) Bus CIF @ 30 Hz

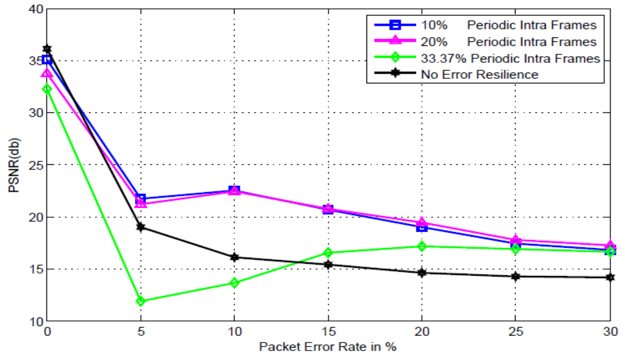


(b) Foreman QCIF @ 10 Hz

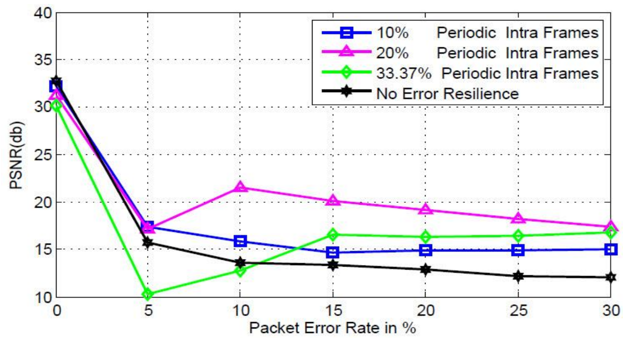
Fig 8. Comparison of CABAC and CAVLC entropy coding schemes

Intra placement on the macro-block, slice, or frame level, was used primarily to combat drifting effects. Simulation was carried out in Baseline profile using two QCIF test video sequences Foreman and City, with a target bit-rate of 56 kbps. The simulation parameters were set as follows: (1) Period of additional Intra frame was chosen as 10, 5 and 3. (2) QP (Quantization Parameter) was kept constant at a value to match the bit rate requirements. (3)10 previous frames were used for inter motion search. (4) No B-slices were used. (5) CAVLC entropy coding was employed. (6) Frame-copy was used as error concealment method.

Fig. 9 shows the PSNR of the luminance (Y) component versus the packet error rate for the two test sequences, under lossy condition, both of which were encoded at 10 Hz. It is observed from Fig. 9 that placing additional intra-frames can reduce the error propagation. However, intra coded frames require a large number of bits and so there is a limitation on the number of intra coded frames per video frame. We cannot place too many intra frames since that reduces compression efficiency. On the other hand, it is not necessary that the performance will be better in all cases when less intra frames are considered. It is clear from Fig. 9 that using 20% additional intra frames provides the best performance and hence their application should always be considered very carefully. Table 1 tabulates the Y-PSNR values at 20% packet error rate for intra placement.



(a) Foreman QCIF @ 10 Hz



(b) City QCIF @ 10 Hz

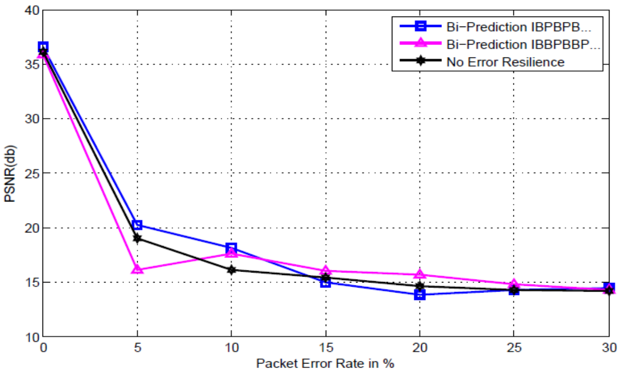
Fig 9. Average Y-PSNR versus packet error rate for intra frame placement

Table 1. Intra Placement

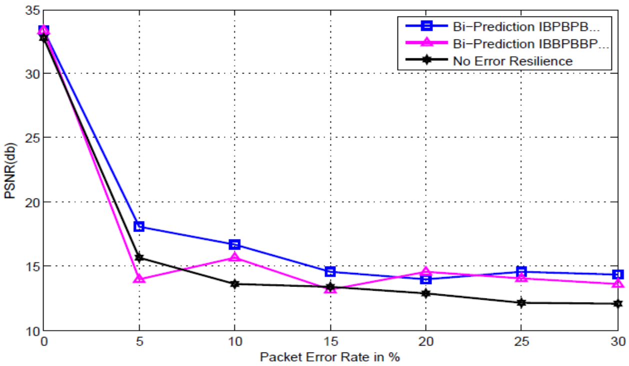
Video sequence used	Error resilience technique used	Y-PSNR (dB)
Foreman QCIF @ 10 Hz	Intra Placement with 10% periodic intra frames	19.01
	Intra Placement with 20% periodic intra frames	19.44
	Intra Placement with 33.37% periodic intra frames	17.14
	No error resilience	14.64
City QCIF @ 10 Hz	Intra Placement with 10% periodic intra frames	14.9
	Intra Placement with 20% periodic intra frames	19.18
	Intra Placement with 33.37% periodic intra frames	16.37
	No error resilience	12.87

Prediction can be further improved by the use of bi-prediction at the cost of computational complexity. We have simulated the performance of bi-prediction under lossy condition. The simulation parameters were set as follows: (1) Only first frame was intra coded (2) QP was kept constant at a value to match the bit rate requirements (3) 10 previous frames were used for inter motion search (4) IBPBPBPB…. and IBBPBBPBBP… patterns of sequence were used (5) CAVLC entropy coding was used (5) Frame-copy was used as error concealment method.

Fig. 10 shows the PSNR of the luminance component versus the packet error rate for the two test sequences, City and Foreman, under lossy condition, in case of Bi-prediction, both of which were encoded at 10 Hz with a target bit rate of 56 kbps. It is observed that pattern IBPBPB… perform better than IBBPBBP… in case of City, while opposite is the case in Foreman. However, H.264/AVC video coding standard does not specify bi-prediction as error resilient feature [3] [7], but as we see from Fig. 10, this method may be used as an error resilient feature, if we can neglect the complexity. Table 2 tabulates the Y-PSNR values at 20% packet error rate for bi-prediction.



(a) Foreman QCIF @ 10 Hz



(b) City QCIF @ 10 Hz

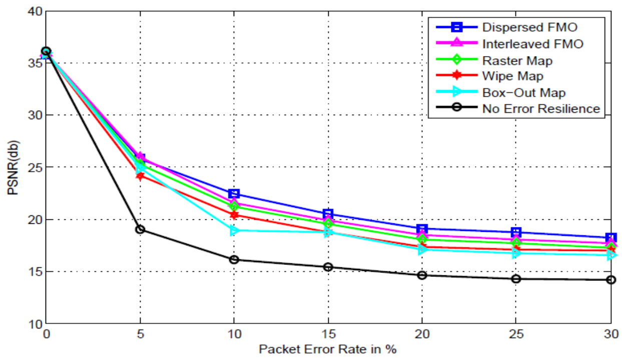
Fig 10. Average Y-PSNR over error rate for bi-prediction

Table 2. Bi-Prediction

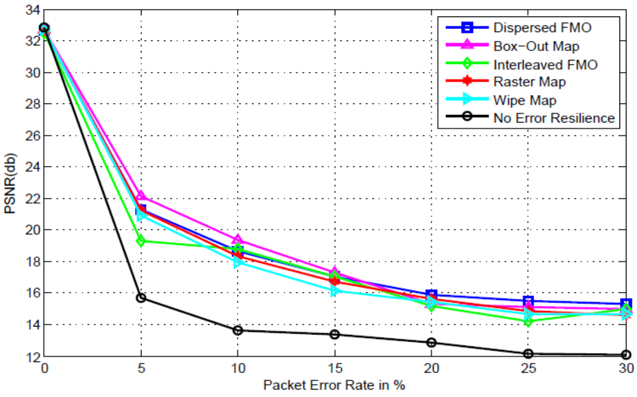
Video sequence used	**Error resilience technique used**	**Y-PSNR (dB)**
Foreman QCIF @ 10 Hz	Bi-Prediction IBPBPB....	13.815
	Bi-Prediction IBBPBBP...	15.645
	No error resilience	14.64
City QCIF @ 10 Hz	Bi-Prediction IBPBPB....	14.01
	Bi-Prediction IBBPBBP...	14.5675
	No error resilience	12.87

FMO, which is available in the Baseline and Extended profile but not in the Main profile, allows assigning MBs to slices in an order other than the scan order. To do so, each MB is statically assigned to a slice group using a macro-block allocation map (MBA map). The slice structuring strategy aims at avoiding error propagation from a corrupted packet to subsequent packets. We have simulated different slice structures supported by H.264/AVC video coding standard [Refer section 3.3]. Firstly, the video sequence was encoded without any slice structure i.e. one packet per frame, then different slice structures were used to encode the test sequence, and the performance was analyzed under lossy condition. Simulation was carried out in Baseline profile using two QCIF test video sequences Foreman and City, with a target bit-rate of 56 kbps. The objective behind the FMO is to scatter possible errors to the whole frame as equally as possible to avoid error accumulation in a limited region. Fig. 11 shows the PSNR of the luminance component versus the packet error rate for the two test sequences, City and Foreman, under lossy condition; in case of different slice structures where 2 slice groups per frame were taken for simulation. Both sequences were encoded at 10 Hz with a target bit rate of 56 kbps.

In case of slice structuring, every lost MB has several spatial neighbors that belong to the other slice. Hence, an error-concealment mechanism has sufficient information for efficient concealment. It is observed from the Fig. 11 that dispersed or checker board structure (2 slice group per frame i.e., macro-blocks with odd addresses in slice group 1, with even addresses in slice group 2) performs better than other slice structures with increase in loss rate. Table 3 tabulates the Y-PSNR values at 20% packet error rate for FMO.



(a) Foreman QCIF @ 10 Hz



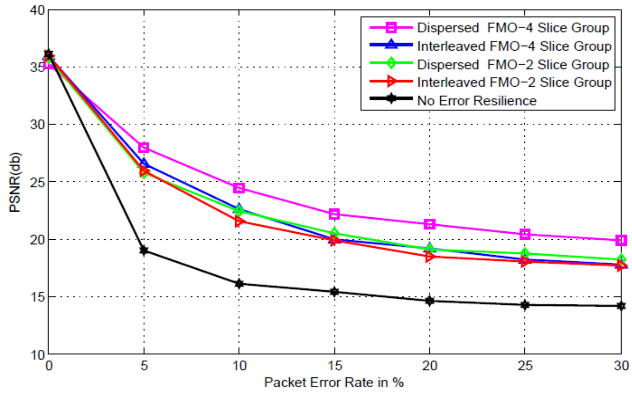
(b) City QCIF @ 10 Hz

Fig 11. Average Y-PSNR over error rate for different slice structures

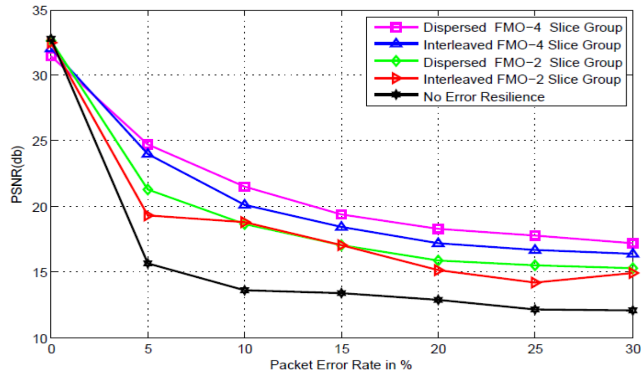
Table 3. Slice Structuring: FMO (2 slice group)

Video sequence used	Error resilience technique used	Y-PSNR (dB)
Foreman QCIF @ 10 Hz	Dispersed FMO	19.12
	Interleaved FMO	18.464
	Raster map	17.997
	Wipe map	17.309
	Box-out map	17.05
	No error resilience	14.64
City QCIF @ 10 Hz	Dispersed FMO	15.86
	Interleaved FMO	15.336
	Raster map	15.173
	Wipe map	15.609
	Box-out map	15.446
	No error resilience	12.87

Performance may be further improved, if 4 slice groups per frame are used, because, in this case, the samples of a missing slice are surrounded by many samples of correctly decoded slices. Fig. 4 illustrates the slice structuring in case of dispersed FMO, when 4 slice groups per frame were adopted in simulation. Fig. 12 shows the PSNR of the luminance component versus the packet error rate for the two test sequences, City and Foreman, under lossy condition, in case of dispersed and interleaved slice structures where 4 slice groups per frame were taken for simulation. Both sequences were encoded at 10 Hz with a target bit rate of 56 kbps. It is observed from Fig. 12 that the performance of structure with 4 slice groups outperforms the performance of structure with 2 slice groups. Table 4 tabulates the Y-PSNR values at 20% packet error rate for the 2 and 4 slice group structures.



(a) Foreman QCIF @ 10 Hz



(b) City QCIF @ 10 Hz

Fig 12. Comparison of 4 slice group frame with 2 slice group frame

Table 4. 2 and 4 Slice Group Structures

Video sequence used	Error resilience technique used	Y-PSNR (dB)
Foreman QCIF @ 10 Hz	Dispersed FMO-4 slice group	21.2375
	Interleaved FMO-4 slice group	19.146
	Dispersed FMO-2 slice group	19.12
	Interleaved FMO-2 slice group	18.464
	No error resilience	14.64
City QCIF @ 10 Hz	Dispersed FMO-4 slice group	18.27
	Interleaved FMO-4 slice group	17.20
	Dispersed FMO-2 slice group	15.86
	Interleaved FMO-2 slice group	15.173
	No error resilience	12.87

Weighted prediction is available in the Main and Extended profiles, but not in the Baseline profile. We have simulated the weighted prediction in Main profile using CAVLC. Fig. 13 shows the performance of weighted prediction from which it is clear that weighted prediction enhances PSNR. Table 5 tabulates the Y-PSNR values at 20% packet error rate, for with and without weighted prediction.

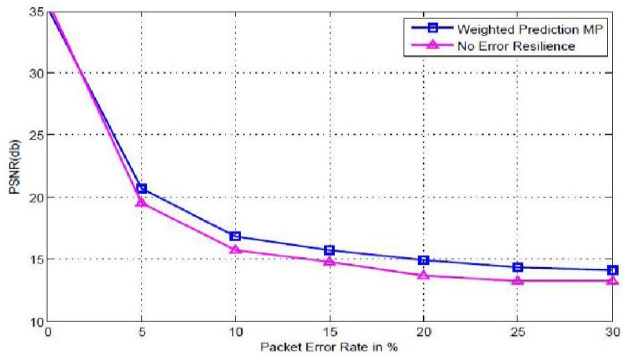


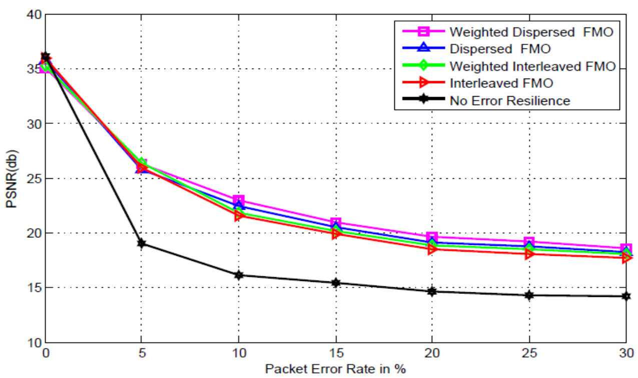
Fig 13. Performance evaluation of weighted prediction

Table 5. Weighted Prediction

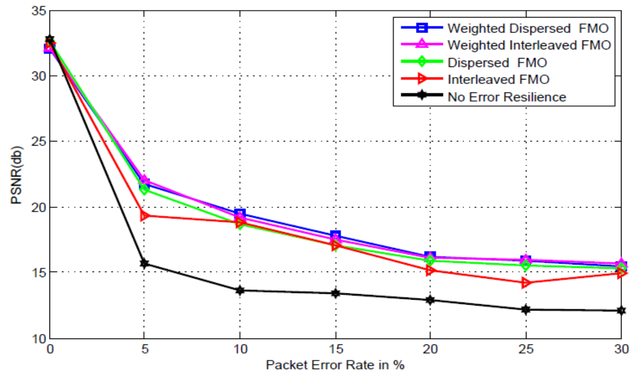
PSNR	Weighted prediction	No error resilience
Y-PSNR (dB)	14.90	13.67

We have combined the weighted-prediction in Extended-profile as an error resilient feature with slice structuring. In normal slice structure, the lost blocks are concealed by their surrounding blocks; while in this case, concealment is done by taking the weights of surrounding blocks. Our main aim was to check the performance of weighted-prediction within different slice structures. It is observed from the Fig. 14 that the weighted prediction concealment scheme gives improvement in PSNR. Table 6 tabulates the Y-PSNR values at 20% packet error rate for weighted slice structuring.

In our experiment, we also tested the performance of frame copying technique for error concealment to combat frame loss during video transmission. The two QCIF video sequences "Foreman" and "City" at 10 frames/sec were used for the purpose. The sequences were coded as IPPPP….. with a period of I-frame reset 15. In this simulation, a P-frame is dropped in every GOP. The dropped frame is then concealed by frame copying. The corresponding PSNR values are then calculated and compared to the no error case. Table 7 presents the average PSNR performances over the erroneous frames defined as the frames that are corrupted due to frame losses.



(a) Foreman QCIF @ 10 Hz



(b) City QCIF @ 10 Hz

Fig 14. Performance evaluation of weighted slice structuring

Table 6. Weighted Slice Structuring

Video sequence used	Error resilience technique used	Y-PSNR (dB)
Foreman QCIF @ 10 Hz	Weighted dispersed FMO (2 slice group)	19.58
	Weighted interleaved FMO (2 slice group)	19.12
	Dispersed FMO (2 slice group)	18.825
	Interleaved FMO (2 slice group)	18.464
	No error resilience	14.64
City QCIF @ 10 Hz	Weighted dispersed FMO (2 slice group)	16.20
	Weighted interleaved FMO (2 slice group)	16.078
	Dispersed FMO (2 slice group)	15.86
	Interleaved FMO (2 slice group)	15.173
	No error resilience	12.87

Table 7: Frame Copying

Sequences	PSNR (dB)	
	No Error	Frame Copying
Foreman QCIF @ 10 Hz	36.0	31.5
City QCIF @ 10 Hz	32.5	29.0

5. CONCLUSIONS

In this paper, various error resilient video coding schemes suitable for video transmission have been discussed. Based on the results obtained, following conclusions are drawn. The coding tools of H.264/AVC when used in an optimized mode allow for bit saving of about 50% compared to previous video coding standards like MPEG-4 and MPEG-2 for a wide range of bit rates and resolutions. However, these savings come at the price of an increased computational cost. The use of I-frame is undoubtedly the most powerful tool to stop error propagation and recover corrupted frames. Since the I-frames are independently coded without using temporal prediction, they reset the prediction process. The coding efficiency is therefore lower than the prediction frame. Moreover, transmitting an I-frame requires a large bandwidth and incurs large bit rate variation, which may not be acceptable in many low delay applications. Prediction may be improved by the use of bi-prediction at the cost of increased computational complexity. However, due to its high complexity it is generally not considered for low delay applications. Slice structured coding reduces packet loss probability and the visual degradation from packet losses, especially in combination with decoder error concealment methods. The price of the use of slice structuring offers somewhat lower coding efficiency (because of the broken in-frame prediction mechanisms between non-neighboring MBs) and, in highly optimized environments, a somewhat higher delay. The use of weighted prediction with slice structuring allows the motion-compensated prediction signal to be weighted which in turn improves coding efficiency. Finally, the choice of one or more of the above-mentioned error resiliency schemes should consider the practical applications as well as network environment.

6. FUTURE WORK

The work presented in this paper is limited to equal error protection. However, in video coding, some syntax elements in the bit-stream are more important than others. So, video bit-streams can be partitioned to segments of different priorities according to their impact on subjective quality. In view of this, the work may be extended to priority partitioning, which enables unequal error protection according to the importance of syntax elements.

7. REFERENCES

[1] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra," Overview of the H.264/AVC video coding standard" *IEEE Transaction on Circuit and Systems forVideo Technology*, vol. 13, no. 7, pp. 560-576, 2003.

[2] D. Marpe, T. Wiegand and G. J. Sullivan, "The H.264/MPEG-4 advanced video coding standard and its applications," *IEEE Communications Magazine*, vol. 94, no. 8, pp. 134-143, 2006.

[3] S. Wenger, "H.264/AVC over IP," *IEEE Transaction on Circuit and Systems for Video Technology*, vol. 13, pp. 645–656, 2003.

[4] Sood. A, Chilamkurti N.K and Ben Soh, "Study and Analysis of an Error Resilient Technique in H.264 Video Using Flexible Macroblock Grouping," IEEE Wicom, International Conference on Digital Object Identifier, Sept 2006.

[5] T. Stockhammer and M. M. Hannuksela, "H.264/AVC video for wireless transmission," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 6-13, 2005.

[6] H.264/AVC reference software available online at http://iphome.hhi.de/suehring/

[7] Sunil Kumar, Liyang Xu, Mrinal K. Mandal, and Sethuraman Panchanathan, "Error resiliency schemes in H.264/AVC standard," *Journal of Visual Communication & Image Representation*, vol. 17, no. 2 pp. 425-450, 2006.