# Error-Resilient Performance of Dirac Video Codec Over Packet-Erasure Channel

M. Tun, K. K. Loo, *Member, IEEE*, and J. Cosmas, *Member, IEEE*

*Abstract*—Video transmission over the wireless or wired network requires error-resilient mechanism since compressed video bitstreams are sensitive to transmission errors because of the use of predictive coding and variable length coding. This paper investigates the performance of a simple and low complexity error-resilient coding scheme which combines source and channel coding to protect compressed bitstream of wavelet-based Dirac video codec in the packet-erasure channel. By partitioning the wavelet transform coefficients of the motion-compensated residual frame into groups and independently processing each group using arithmetic and Forward Error Correction (FEC) coding, Dirac could achieves the robustness to transmission errors by giving the video quality which is gracefully decreasing over a range of packet loss rates up to 30% when compared with conventional FEC only methods. Simulation results also show that the proposed scheme using multiple partitions can achieve up to 10 dB PSNR gain over its existing un-partitioned format. This paper also investigates the error-resilient performance of the proposed scheme in comparison with H.264 over packet-erasure channel.

*Index Terms*—Coefficient partitioning, Dirac, error-resilient coding, H.264, wavelets.

## I. INTRODUCTION

COMPRESSION techniques for digital video have given rise to applications such as storage, broadcast over digital cable, internet broadcasting, satellite and terrestrial digital TV, video-conferencing and video-telephony. This has led the development of international standards, namely H.261 and H.263 under International Telecommunication Union Telecommunication Standardization Sector (ITU-T) and MPEG-1, MPEG-2 and MPEG-4 under International Organization for Standardization/International Electrotechnical Commission (ISO/IEC). The latest H.264 (a.k.a. MPEG-4 Part 10) standard which is being developed by the Joint Video Team (JVT) of the ITU-T Video Coding Expert Group (VCEG) and MPEG is aimed to elaborate an open standard that is not application-specific and that perform significantly better than the existing standards in terms of compression, network adaptation and error robustness. In the near future, H.264 will gain wide acceptance on many applications especially on Internet broadcasting. However, the usage of H.264 incurs royalty fees [1] which may not be cost effective for non-profit and public content owners such as public

service broadcasters, archive institutes, etc., for deployment of Internet-based services. Whilst these costs are manageable initially, these could become prohibitive if the services scaling up to millions of users, or if new services are deployed which were not envisaged in the original license agreements.

As an alternative, a royalty-free general-purpose video codec called Dirac [2] is designed, which is aimed at a wide range of applications from storage of video content to streaming video in view to address the above demands. Being "open technology", Dirac is an attractive option as it allows content owners to distribute contents without royalty-fees in anyway. However, the Dirac current alpha releases have only been optimized for storage purposes and still there is no error-resilient encoding mechanism. This paper investigates the performance of a simple and low complexity error-resilient coding scheme which combines source and channel coding to protect compressed bitstream of wavelet-based Dirac video codec in the packet-erasure channel, i.e. Internet. This paper also investigates the error-resilient performance of the proposed scheme in comparison with H.264 over packet-erasure channel.

The organization of this paper is as follows. Section II discusses the related work and the work of this paper. Sections III and IV present the Dirac coding architecture and the wavelet co-efficients partitioning, respectively. Sections V and VI present the structure of the proposed error-resilient encoding and decoding, respectively. Finally, Sections VII and VIII present the comprehensive result analyses and the conclusion respectively.

## II. RELATED WORK

Several techniques have been developed over the last decades to make video transmission over a wireless or wired channel resilient to errors. One approach is to transmit the video sequence into several bitstreams, called descriptions. In this method, a video sequence will be encoded into two or more bitstreams or descriptions and transmitted over different channels. The descriptions can be generated independently or correlatively. When all of the descriptions are correctly received, the decoder can reconstruct the video with best quality. If any of the description is lost during transmission, the decoder can still reconstruct the video with a lower, but acceptable quality. In [3], a new generation scheme using side information was introduced which is based upon the multiple reference frames. A novel technique of Multiple Description Scalar Quantization for Fine Granularity Scalability (MDSQ-FGS) is presented in [4]. In Multiple Description Scalable Coding (MDSC), the advantages of scalable coding was combined with Multiple Description in order to adjust the amount of redundancy and the bitrate allocated to each description at transmission time [5].

Another approach, called coefficient partitioning makes video transmission resilient to channel errors by partitioning the wavelet coefficients into groups and independently processing each group. Thus, a bit error in one group does not affect the others, allowing more uncorrupted information to reach the decoder. This method was first reported by Creusere [6] for use with the EZW algorithm. Block based coefficient partitioning method is presented in [7] where they partition each subband into an equal number of coefficient blocks. Each coefficient block in a subband carries information about some localized region in the original frames. The components are then formed by grouping from each subband, equal number of coefficient blocks that correspond to different spatial regions of the source.

Some consider protecting the transmitted bitstreams against packet losses by applying an unequal amount of Forward Error Correction (FEC) to different data fragments according to the importance of the data [8], [9]. However, this technique has the disadvantage of still being vulnerable to packet erasures or channel errors that occur early in the transmission, either of which can cause a total collapse of the decoding process. To overcome this problem, combined source and channel coding has been considered in most cases where one of the coefficient partitioning methods is used as source coding and combined together with FEC to achieve double level of protection from transmission error [10], [11].

In Pearlman's work [12], the wavelet transform coefficients is first broken into a number of spatio-temporal tree blocks according to [6], and the 3-D SPIHT algorithm is modified to work independently with these blocks. They then apply Kim's method [13], [14] of rate compatible punctured convolutional (RCPC) channel coding to every packet to protect the data. It is interesting to note that the scheme proposed in [15] could be used inline with any error-resilient coding method to alleviate the effect of error propagation by adding some periodic macroblocks in every fifth inter-frames.

In this work, the wavelet transform coefficients partitioning method of progressive image [6], [8]–[14] is extended to work in 2D wavelet transformed *motion-compensated residual* video frames in Dirac and used as the source coding. The source coding is combined with FEC, either RCPC [16] or Turbo Coding (TC) [17] as channel coding to protect the compressed bitstreams of Dirac to achieve robustness to transmission errors in packet-erasure channel. In this way, the decoder can still reconstruct the corrupted blocks if the reference frame and its motion vector are correctly received. Error-resilient transmission for the packet-erasure channel can be achieved by using the bitwise interleaver at the output of the channel encoder. In this work, transform coefficients coding algorithms such as EZW, ZTE, SPIHT, etc., are not considered since all of these are heavily patented and Dirac has no intention to include any patented algorithm in the codec architecture [2]. Moreover, these algorithms do not perform very well in applying to the *motion-compensated residual frames* since most of the coefficients in these frames have already been transformed to zeros.

## III. Dirac Video Codec

The Dirac design is that of a conventional hybrid motion-compensated architecture which based around fundamental coding algorithms, i.e. hierarchical motion estimation,
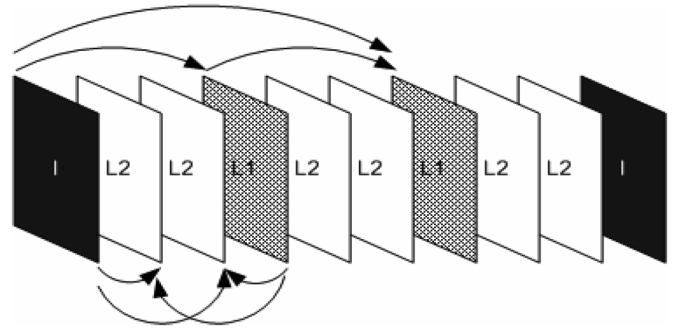


Fig. 1.  Prediction of L1 and L2 frame.

overlapped-block motion compensation (OBMC), discrete wavelet transform (DWT), rate-distortion optimization (RDO) quantization and entropy coding [2]. It is aimed to be competitive with other state-of-the-art video coding standards, and the current performance is two-fold of MPEG-2 and slightly less than H.264 [19] even in the Alpha development stage. However, the performance is not the only factor driving the Dirac design but it is also intended to be simple, modular and most importantly not to infringe any patented algorithms, hence it is offered as open technology.

Dirac uses OBMC to avoid block-edge artifacts which are sensitive to wavelet transform. First the motion-compensated residual frames are wavelet-transformed using separable wavelet filters and divided into subbands which are then quantized using RDO quantizers. The quantized data is entropy coded using an Arithmetic encoder. Dirac supports any frame dimensions (QCIF to HDTV) and common chroma formats (luma only, 4:4:4, 4:2:2, 4:2:0) by means of frame padding. The padding ensures that the wavelet transform can be applied properly. Frame padding also allows for any size blocks to be used for motion estimation, even if they do not evenly fit into the picture dimensions.

Dirac defines three frame types. Intra frames (I-frames) are coded independently without reference to other frames in the sequence. Level 1 frames ($L_1$ frames) and Level 2 frames ($L_2$ frames) are both inter frames, which are coded with reference to other previously (and/or future) coded frames. The definition of the $L_1$ and $L_2$ frames are the same with $P$ and $B$ frames in H.264. The encoder operates with standard Group of Picture (GOP) modes whereby the number of $L_1$ frames between I-frames, and the separation between $L_1$ frames, can be specified depends on the application. A prediction method for frame coding using a standard GOP structure is shown in Fig. 1. In this figure, the number of $L_1$ frames between I-frames is 2 and the $L_1$ frame separation is 3.

## IV. Wavelet Coefficient Partitioning

The wavelet coefficient partitioning is to divide the wavelet coefficients at the output of the DWT process of the Dirac codec into $S$ groups and then quantize and code each of them independently so that $S$ different bitstreams are created [6]. By coding the wavelet coefficients with multiple, independent bitstreams, any single bit error truncates only one of the $S$ bitstreams while the others are still correctly received. Therefore,
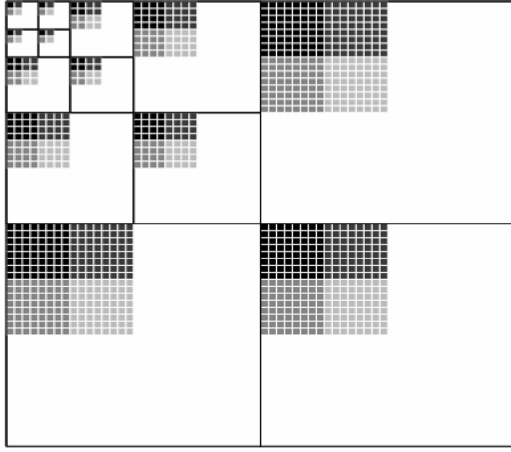
Fig. 2. Wavelet Coefficient partitioning for $S = 4$, with four levels wavelet transform.



Fig. 3. Structure of error-resilient robust wavelet coefficient partitioning and encoding procedure.

the wavelet coefficients represented by the corrupted bitstreams are reconstructed at reduced accuracy, while those represented by the error-free streams are reconstructed at the full encoder accuracy. The partitioning method used here is the extension of [6], in which it is applied to the *motion-compensated residual frames* instead of the intra coded frames for the image transmission in [6] and 3D wavelet transformed frames in [12]. In this way, the quality of the reconstructed frames particularly at high packet loss rate becomes much better than the original scheme in [6] and [12] especially when the motion vector data and reference frames are correctly received. It is because the corrupted data can simply be replaced with the shifted version of the data from the reference frame pointed by the motion vector.

Fig. 2 graphically illustrates this wavelet coefficient partitioning for $S = 4$ bitstreams for four levels wavelet decompositions. In this figure, each coefficient with the same shade of grey maps the same group. Note that, $S$ can be increased by powers of four. If the image is of size $X \times Y$ and $L$ levels of wavelet decomposition are used, then the maximum number of independent bitstreams allowed is:

$$S = (X \times Y)/4^L \qquad (1)$$

## V. ERROR-RESILIENT ENCODING

Fig. 3 shows the block diagram of error-resilient encoding procedure of the Dirac. The DWT output of Dirac encoder is divided into $S$ sub-frames according to the wavelet coefficient partitioning method shown in the Section IV. These sub-frames are then processed independently by employing the RDO quantization and arithmetic encoding before entering multiplexer. In the multiplexer, all the independent parallel bitstreams are combined to obtain a serial stream starting from bitstream 1, followed by bitstream 2 and so on until bitstream $S$ is reached. The output serial bitstream is then protected by channel coding using rate 2/3, 1/2, 1/3 and 1/4 RCPC or rate 1/2 TC.

Turbo encoder is the parallel concatenation of two recursive systematic convolutional encoders having generator polynomials $g_1 = 31$, $g_2 = 27$ with memory $m = 4$. Puncturing is performed at the output of the encoder by taking only odd parity
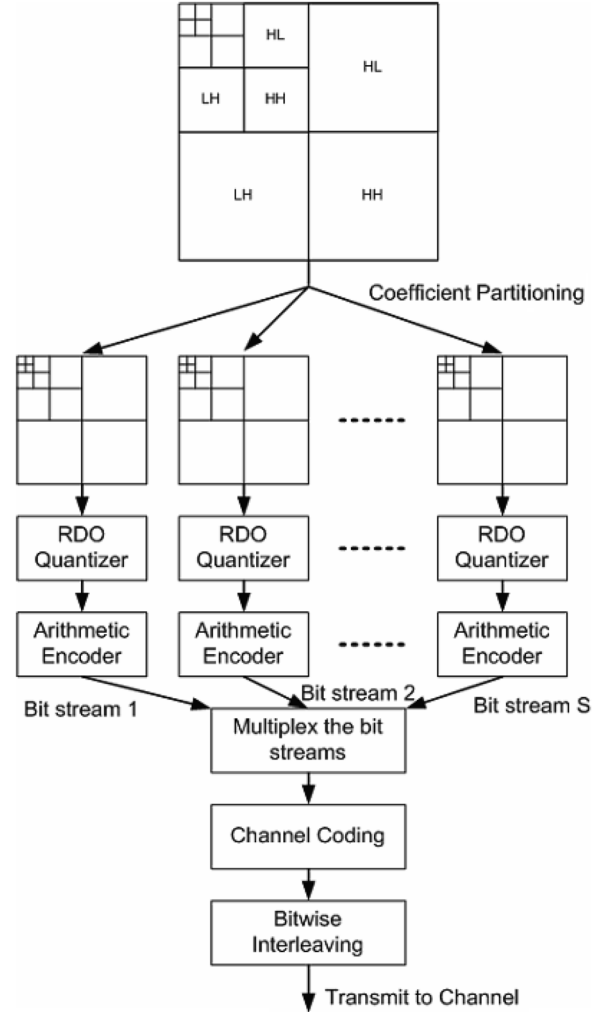
bit and even parity bit from the upper and lower convolutional encoder output correspondingly. The encoder interleaver is a pseudorandom interleaver with any size; this case, it is 200 bits. In the decoder, symbol-by-symbol MAP algorithm is employed with the 6 turbo decoding iterations.

The generator matrix and puncturing tables of the RCPC encoder are given in the Appendix. The memory of the associated convolutional encoder, $m$ is 4 and punctured periodically with period, $P = 8$. At the receiver, the Viterbi algorithm is used to decode the received data.

In order to incorporate channel coding to the output bitstream, the bitstreams is first break into equal length segments of $N$ bits. In our case $N = 200$ bits. The checksum bits, $c = 16$ of the Cyclic Redundancy Code (CRC) is generated for each segment of $N$ bits and then appended to each segment. Next $m$ bits, where $m$ is the memory size of the convolutional coder, are padded at the end of each $N + c$ bit segment to flush the memory of the encoder, i.e. to terminate the trellis at original stage. Finally, each segment of the $N + c + m$ bits is passed through the rate $R$ of the channel encoder. The generator polynomial used for the CRC is $g(x) = X^{16} + X^{14} + X^{12} + X^{11} + X^8 + X^5 + X^4 + X^2 + 1$ as in [12].
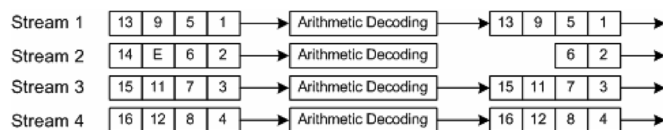
Fig. 4. Example of decoding failure at packet number 10.

Bitwise interleaver is placed at the channel encoder output before the construction of packet. The role of the bitwise interleaver is to distribute the series of information bits into several different locations so that a packet lost in the packet-erasure channel does not affect the error correcting capability of the channel decoder, i.e. to avoid the formation of error burst. For example, the bitwise interleaver length is set to 100 times the length of the packet, where packet length is equal to $R^{-1} \times (N + c + m)$ bits. That means a packet loss in a packet-erasure channel does not mean that the whole packet is lost instead only 1/100 of a packet is lost. In the receiver, the channel decoder can effectively correct those errors since the possibility of error burst formation is eliminated by using the bitwise interleaving.

## VI. ERROR-RESILIENT DECODING

In packet-erasure channel, there will be no bit errors occurred inside each packet except the loss of whole packet because of network congestion. In the event of packet loss, all-zero data packet is created at the decoder to replace the lost one and undergo a bitwise deinterleaving process.

The channel decoding and demultiplexing are carried out repetitively in order to obtain $S$ number of sub bitstreams. The channel decoder, i.e. RCPC or TC decoder, normally attempts to correct the errors; but if the errors are still persisting the erroneous packet is marked by ERROR_CODE. Received packet is checked whether it is erroneous by using CRC. Arithmetic decoder will stop decoding for this packet and the remaining packets of this bitstream once the ERROR_CODE is encountered. Up until this point, there will be some clean packets already decoded and lose only the remaining packets. The decoder is then continues to decode packets from other bitstreams.

In contrast, if single bitstream is transmitted without partitioning, any single bit error in the middle of the bitstream will affect the whole remaining bitstream. Therefore, by coding the wavelet coefficients with multiple and independent bitstreams, any single bit error affects only one of the $S$ bitstreams, while the others are received unaffected.

In Fig. 4, if error is occurred in packet number 10, this packet and the rest of this bitstream are discarded. After decoding, the un-partitioned format has only 9 clean packets while in the partitioned format still retain 14 clean packets. Obviously, the partitioned format could deliver more clean packets than in the un-partitioned since it just stops decoding at the point of first error occurrence. A better error-resilient performance can be achieved using maximum possible number of bitstreams which can be calculated by using (1).

## VII. RESULTS

The performance of the proposed Dirac error-resilient scheme is tested with two sequences: Canal-vertical in CIF and Squirrel in SD576; they are chosen to envisage in providing two

TABLE I
THE COMPRESSED VIDEO FILE'S SIZES AND THEIR CORRESPONDING AVERAGE
PSNR VALUES FOR DIFFERENT NUMBER OF PARTITIONS

| | | Num. of Y Partitions | Num. of UV Partitions | Compressed File Size (% of YUV) | Avg. PSNR (dB) |
|---|---|---|---|---|---|
| Un-partitioned | CIF | 1 | 1 | 1.62 | 32.53 |
| | SD576 | 1 | 1 | 2.60 | 32.23 |
| Partitioned | | 6 | 3 | 1.78 | 32.28 |
| | CIF | 33 | 33 | 3.14 | 32.17 |
| | | 99 | 99 | 6.30 | 32.14 |
| | SD576 | 18 | 6 | 2.72 | 32.00 |

viewing experience to Internet users. Both of these sequences can be downloaded from the Dirac project website [2]. For both test sequences, a default GOP structure of Dirac is used, i.e. 36 for CIF and 12 for SD576 where $L_1$ frame separation is 3, and the number of $L_1$ frames between *I-frame*s is 11 and 3 correspondingly. The number of $L_2$ frames can be calculated by using the (2) as follow:

$$Num\_L_2 = (Num\_L_1 + 1) \times L_1\_Sep \qquad (2)$$

The goal is of this section is to analyse the Dirac error-resilient capability and so there has been no attempt to conceal error (error concealment) at the decoder. The Dirac header is considered received successfully from the channel, i.e. no packet loss is introduced to the header. The distortion is measured by the peak signal-to-noise ratio (PSNR). All the PSNR values are averaged over 10 independent runs.

Table I lists the number of partitions used in the simulation for each video component, the compressed video file sizes and their corresponding average PSNR values. It is also possible to set the number of partitions for Y and UV components independently, i.e. different number of partitions for luma and chroma components. The compressed video file size shown there is in percentage of the original uncompressed YUV size.

The use of coefficient partitioning incurs compression inefficiency direct-proportion to the number of partitions. It is because the arithmetic encoder needs to restart as many times as the number of partitions to encode one frame in the multiple partitions format. This is the major trade off between the compression efficiency and error-resilient performance.

The compression inefficiency become prominent in employing the multiple partitioning to the smaller frame size video sequence such as CIF format since there are less coefficients in each partition causing less statistical information providing to the context based Arithmetic encoder. But on the other hand, larger frame size video such as SD576 does not suffer much; the loss is only 0.12 dB in 18_6 partitioned because of having higher number of coefficients in each partition.
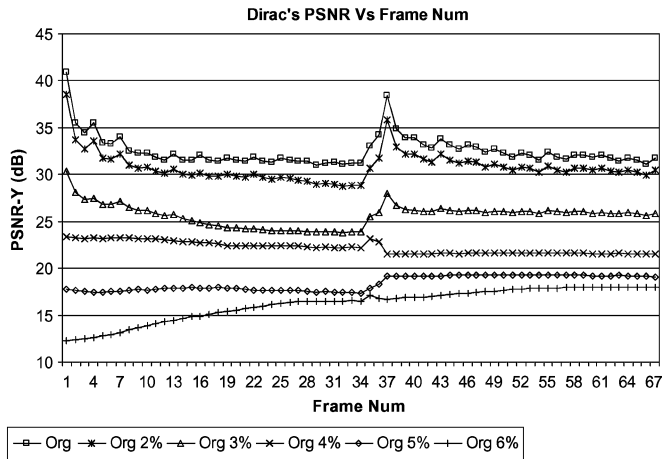
Fig. 5. PSNR Performance comparisons between different percentages of packet loss for un-partitioned format with rate 1/2 RCPC.
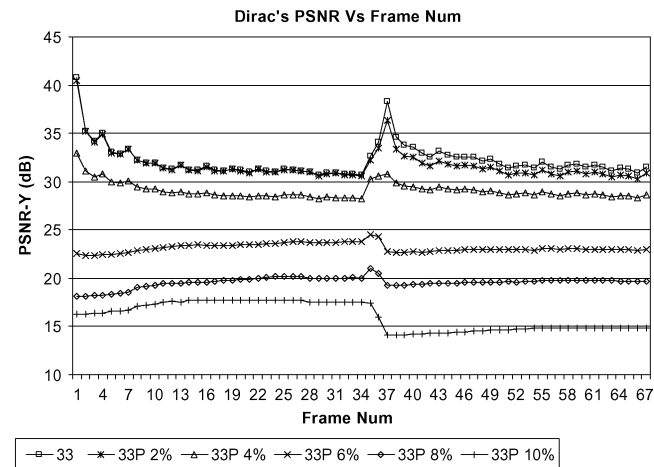


Fig. 6. PSNR Performance comparisons between different percentages of packet loss for 33-partitioned with rate 1/2 RCPC.
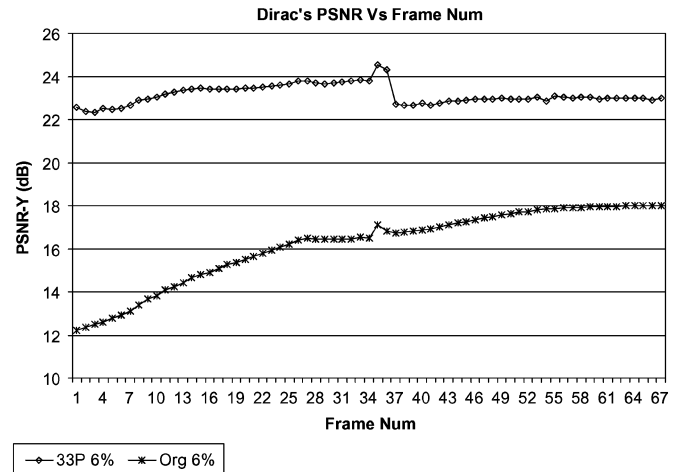


Fig. 7. PSNR Performance comparisons between 33-partitioned and un-partitioned (original) formats for 6% packet loss with rate 1/2 RCPC.
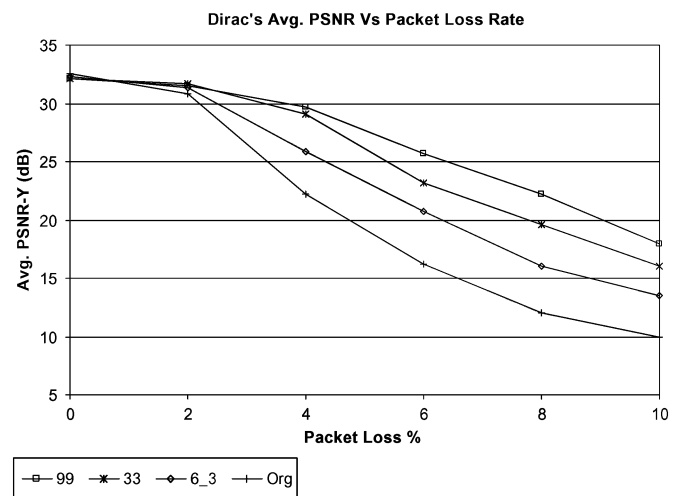


Fig. 8. Average PSNR Performance comparisons between original, 6_3, 33 and 99-partitioned formats for the packet loss from 0 to 10% with rate 1/2 RCPC.

## A. Dirac Error-Resilient Scheme With RCPC

Figs. 5 and 6 show the PSNR performance comparison between different percentages of packet losses for both original (un-partitioned) and 33-partitioned with rate 1/2 RCPC channel coder for Canal-vertical sequence in CIF format. There is no bit error at the RCPC decoder output for the 1% packet loss in both cases, i.e. with the use of bitwise interleaver at the encoder; the RCPC decoder can effectively correct the resulting error pattern at the output of the bitwise de-interleaver. From these figures, the PSNR curve of 2% packet loss is just below the error free curve with a slight performance degradation which is the result of having a few bit errors in the received sequence after RCPC decoding.

Fig. 7 shows the PSNR performance comparison between two formats with the same percentage of packet loss. It is clear that the 33-partitioned achieves at least 5 dB gains over un-partitioned case when the percentage of packet loss is 6%.

Fig. 8 shows the average PSNR performance comparison between un-partitioned and three types of multiple partitioning formats as shown in Table I for the packet loss from 1% to 10%. From this figure, the PSNR gain tends to increase gradually starting from the 2% packet loss and the maximum PSNR

gain achieved for 6_3, 33 and 99 partitioned formats are 4 dB, 8 dB and 10 dB respectively at the 8% packet loss. A better error-resilient performance can be achieved if higher number of partitions is used at the expense of lower compression efficiency.

Fig. 9 shows the average PSNR performance comparison between un-coded, rate 2/3, rate 1/2, rate 1/3 and rate 1/4 of 33-partitioned. It is interesting to note that the performance of rate 2/3 encoding achieves a few dB gains over the un-coded one for the packet loss less than 4% and cross over occurred after that. This is because the error correcting capability of the rate 2/3 RCPC decoder is relatively low and cannot correct the errors effectively when the packet loss rate increases. At this point, because of the usage of bitwise interleaving in rate 2/3 case, packet errors are spread over the interleaving length making the PSNR performance even lower than the un-coded case.

On the other hand, rate 1/2, 1/3 and 1/4 offer better error correcting capabilities and achieve much higher PSNR gain than the un-coded. The coding gain over un-coded case is around 4 dB, 17 dB and 20 dB for the rate 1/2, 1/3 and 1/4 respectively at the 10% packet loss. From the Fig. 9 again, it is clear that
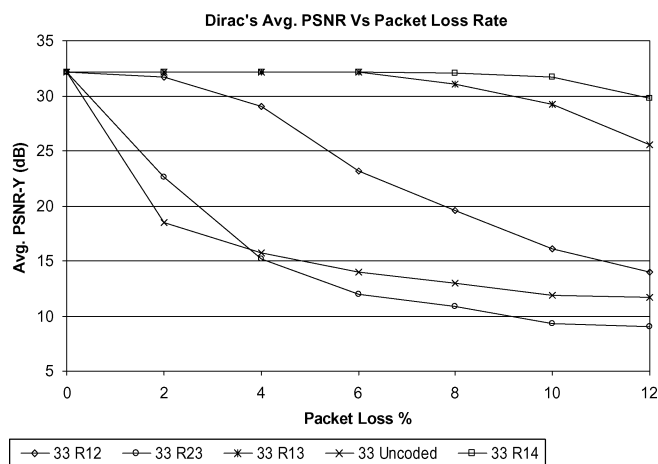
Fig. 9. Average PSNR Performance comparisons between un-coded, rate 2/3, rate 1/2, rate 1/3, and rate 1/4 of 33-partitioned formats for the packet loss from 0 to 12% with RCPC.
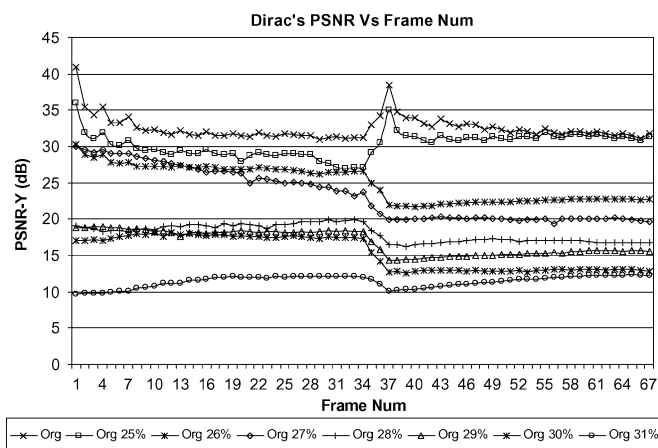


Fig. 11. PSNR Performance comparisons between different percentages of packet loss for 33-partitioned format with rate 1/2 TC.



Fig. 10. PSNR Performance comparisons between different percentages of packet loss for un-partitioned format with rate 1/2 TC.



Fig. 12. PSNR Performance comparisons between 33-partitioned and un-partitioned (original) formats for 30% packet loss with rate 1/2 TC.

there are no losses in terms of PSNR performance in the rates 1/3 and 1/4 encoding; according to the simulation results, there are no bit errors at the output of the RCPC decoder from 1% to 6% packet loss at these encoding rates. So, it is safe to conclude that encoder rates 1/3 and 1/4 are able to be used to protect the header layer of the Dirac's compressed bitstream for less than 6% packet loss. Nevertheless, the performance will be improved enormously by using the error-resilient scheme with TC instead of RCPC as discussed next.

### B. Dirac Error-Resilient Scheme With Turbo Codes

Figs. 10 and 11 show the PSNR performance comparison between different percentage of packet losses for both original (un-partitioned) and 33-partitioned formats with rate 1/2 TC. Clearly, both cases can sustain more than 30% packet loss with the use of TC. There is no error at the decoder output for less than 25% packet loss. This shows that the combined effect of the bitwise interleaver and channel encoder is much more efficient with the use of powerful channel coding mechanism.

Fig. 12 shows the PSNR performance comparison between two partitioned formats with the same percentage of packet loss.
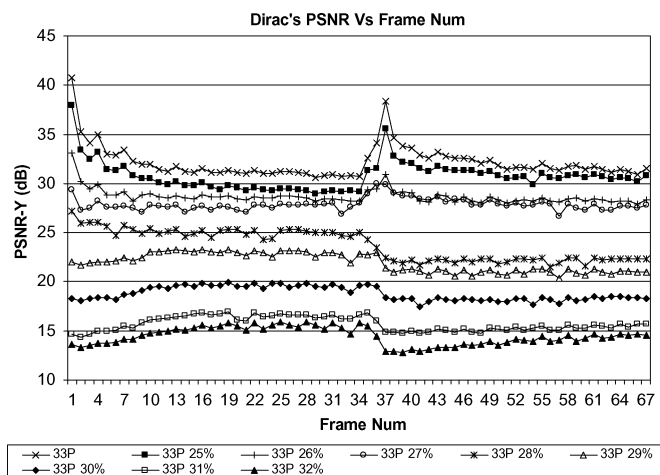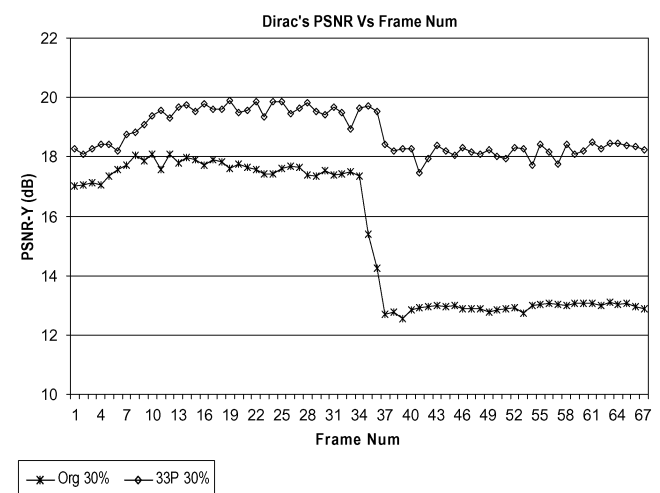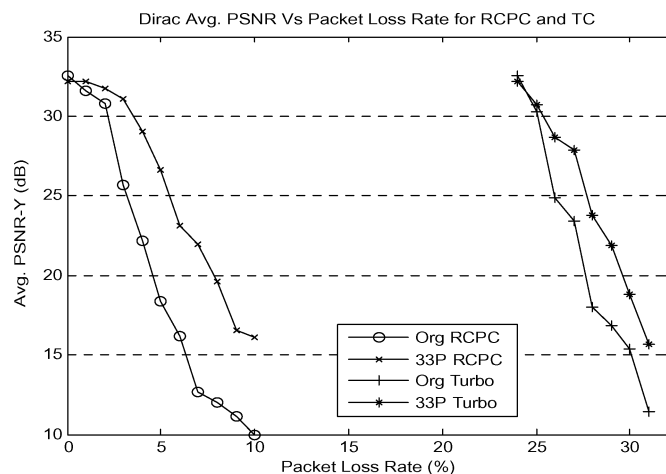


Fig. 13. Average PSNR Performance comparisons between 33-partitioned and un-partitioned (original) formats with Rate 1/2 RCPC coding and TC.

It is clear that the 33-partitioned also achieve at least 5 dB gains over the un-partitioned format in TC case as well. This result
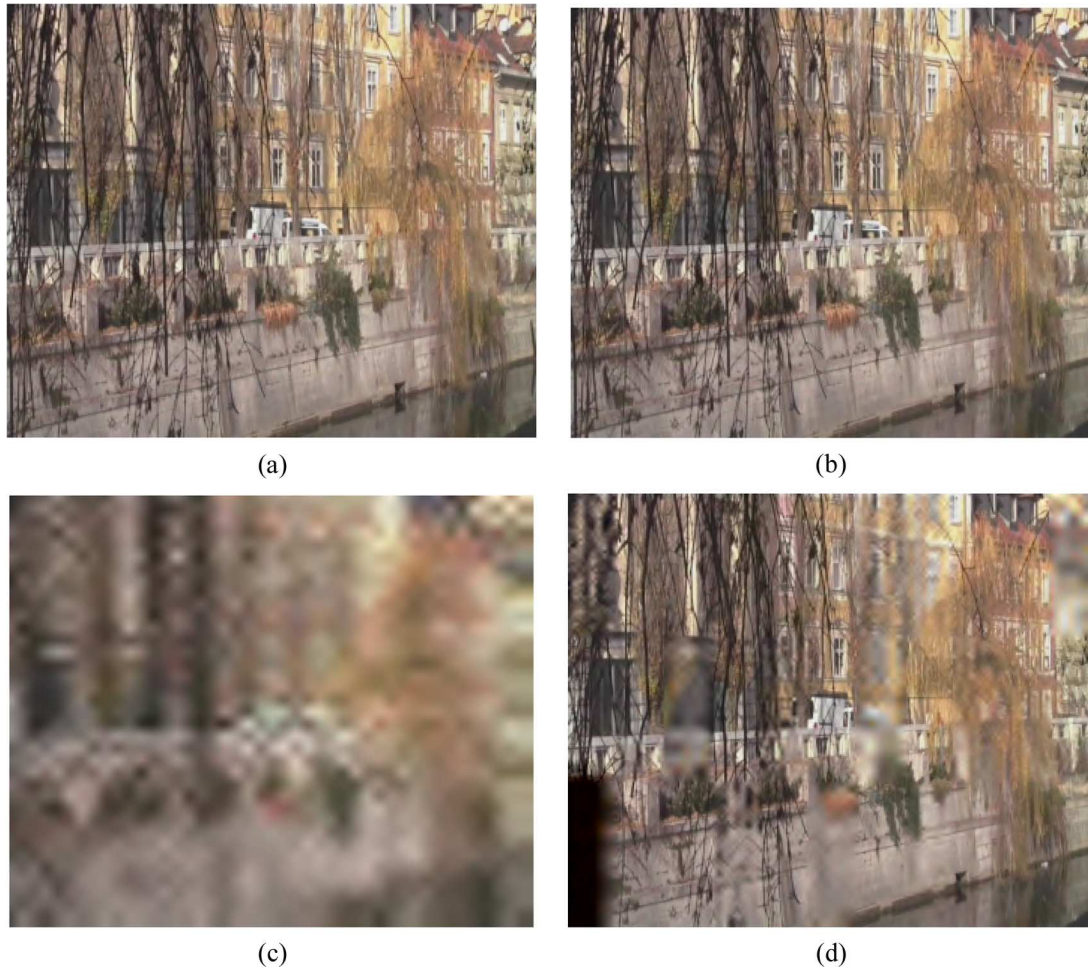
Fig. 14.  4:2:0, CIF format Canal-vertical sequence (frame 37, I-frame) with rate 1/2 RCPC: (a) un-partitioned, 2% packet loss; (b) 33 partitioned, 2% packet loss; (c) un-partitioned, 6% packet loss; (d) 33 partitioned, 6% packet loss.

is identical with the RCPC coding result from Fig. 7. It is because both of these cases use the same number of partitions even though they have different type of channel coding. The 5 dB gain is regarded as the coefficient partitioning gain.

Fig. 13 shows the average PSNR performance comparison between the un-partitioned and 33-partitioned formats with rate 1/2 RCPC and the rate 1/2 TC for the packet loss from 1% to 31%. From this figure, the PSNR gain of RCPC tends to increase gradually starting from the 2% packet loss and the maximum PSNR gain achieved for 33-partitioned is 8 dB at the 8% packet loss. On the other hand, the maximum PSNR gain of TC is approximately 6 dB over the un-partitioned format at the percentage of packet loss around 28%. A better error-resilient performance can be achieved if higher number of partitions is used at the expense of lower compression efficiency. But in both channel coding types, the average PSNR gain is approximately 5 dB over un-partitioned format, which can be seen clearly in Figs. 7 and 12 for RCPC and TC respectively. As expected, the TC can protect the transmitted packet sequence much better then RCPC with the expense of higher decoding complexity and iteration delay at the receiver. Nevertheless, the advent of network-on-chip and high throughput turbo decoders enables the application of turbo codes on packet network [20], [21].

### C. Video Quality of the Dirac Error-Resilient Scheme

Figs. 14(a)–(d) show the frame number 37 (I-frame) for un-partitioned and the 33-partitioned formats with 2% and 6% packet loss. The corresponding PSNR values for Figs. 14(a)–(d) are 38.45, 38.34, 19.45 and 22.22 dB respectively. A vertical black patch at the lower left corner of the Fig. 14(d) is the result of the loss of the whole partition. This happened when the bit error occurred in the lowest subband (DC subband) of a particular sub-frame or partition since the whole remaining bitstream of this sub-frame starting from the error location needs to be discarded. It is occurred only in the I-frame coding since there is no reference frame in order to compensate this error.

Figs. 15(a) and (b) show the frame number 20 of un-partitioned and the 33-partitioned formats with rate 1/2 TC for 30% packet loss. The corresponding PSNR values for the Figs. 15(a) and (b) are 18.92 and 24.15 dB, respectively. From the figures, it can be seen that the proposed scheme offers very good reconstructions for all packet loss rates compared to the un-partitioned; for example, the 6% packet loss in the Figs. 14(c) and (d) where the proposed scheme provides an excellent protection while the reconstructed quality of the un-partitioned is completely corrupted even though the PSNR
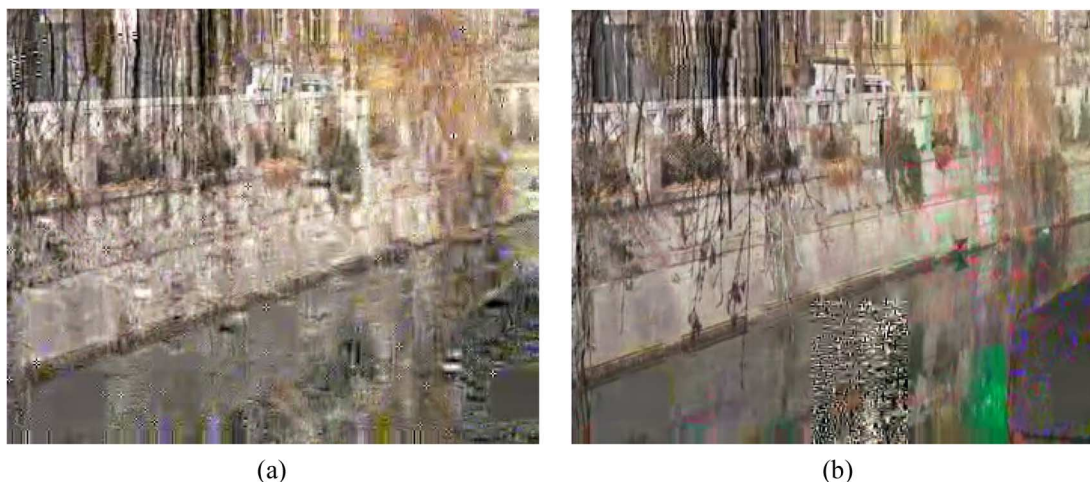
Fig. 15.   4:2:0, CIF format Canal-vertical sequence (frame 20) with rate 1/2 TC: (a) un-partitioned, 30% packet loss; (b) 33-partitioned, 30% packet loss.
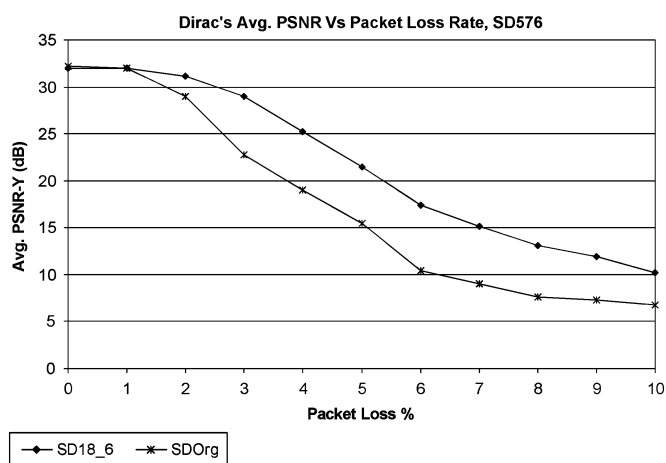


Fig. 16.   Average PSNR performance comparisons between un-partitioned and 18_6 partitioned formats of SD576 Squirrel sequence for the packet loss from 0 to 10% with rate 1/2 RCPC.
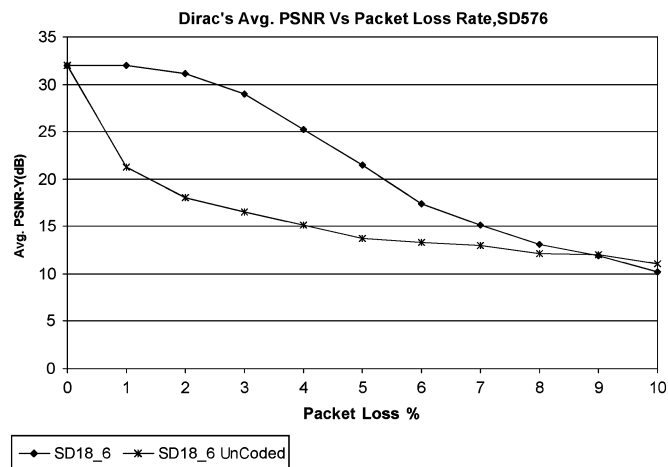


Fig. 17.   Average PSNR performance comparisons between un-coded and coded with rate 1/2 RCPC for 18_6 partitioned formats of SD576 Squirrel sequence, the packet loss from 0 to 10%.

difference is only 2.77 dB. The reconstruction quality of the Fig. 15(b) is also much better than Fig. 15(a) in TC.

Fig. 16 shows the average PSNR performance comparison between un-partitioned and 18_6 partitioned formats of Squirrel sequence for the packet loss from 1% to 10%. From this figure, it can be seen that the PSNR gain tends to increase gradually starting from the 2% packet loss and the maximum PSNR gain achieved is 7 dB at the 6% packet loss.

Fig. 17 shows the average PSNR performance comparison between un-coded and rate 1/2 RCPC encoding of Squirrel sequence with 18_6 partitioned formats. The coding gain of rate 1/2 RCPC encoding over un-coded case is around 13 dB at the 2% packet loss.

Fig. 18(a) shows the example of the losses of the some sub-frames in frame number 37 (I-frame) of the Canal-vertical sequence in CIF format with 33-partitioned. Applying the appropriate error concealment method can easily conceal these types of error patches. Fig. 18(b) shows the results of the application of the error concealment method to the video sequences in Fig. 18(a). The method is implemented by replacing the

corrupted coefficients with the average values of the coefficients of the previous and next partitions. But unfortunately, this method is not applicable to the un-partitioned format since there is only one partition in one video frame. Fig. 18(c) shows the example of the losses of the some sub-frames in frame number 12 (I-frame) of the Squirrel sequence in SD576 format with 18_6 partitioned and the result of the application of error concealment method is shown in Fig. 18(d). It can be seen from the figures that the subjective qualities of the resulting video become much better after applying even simple error concealment scheme. So, it is interesting to note that the application of the error concealment becomes easier with the proposed scheme since the error are localized in one partition and the loss data can be estimated either from the adjacent partitions or the adjacent frames.

### D. Dirac Error-Resilient Scheme Without FEC

This section analyzes the performance of Dirac error-resilient scheme with wavelet coefficient partitioning in comparison to H.264 error-resilient with flexible macroblock ordering (FMO). To facilitate this experiment, H.264 is encoding in RTP packet
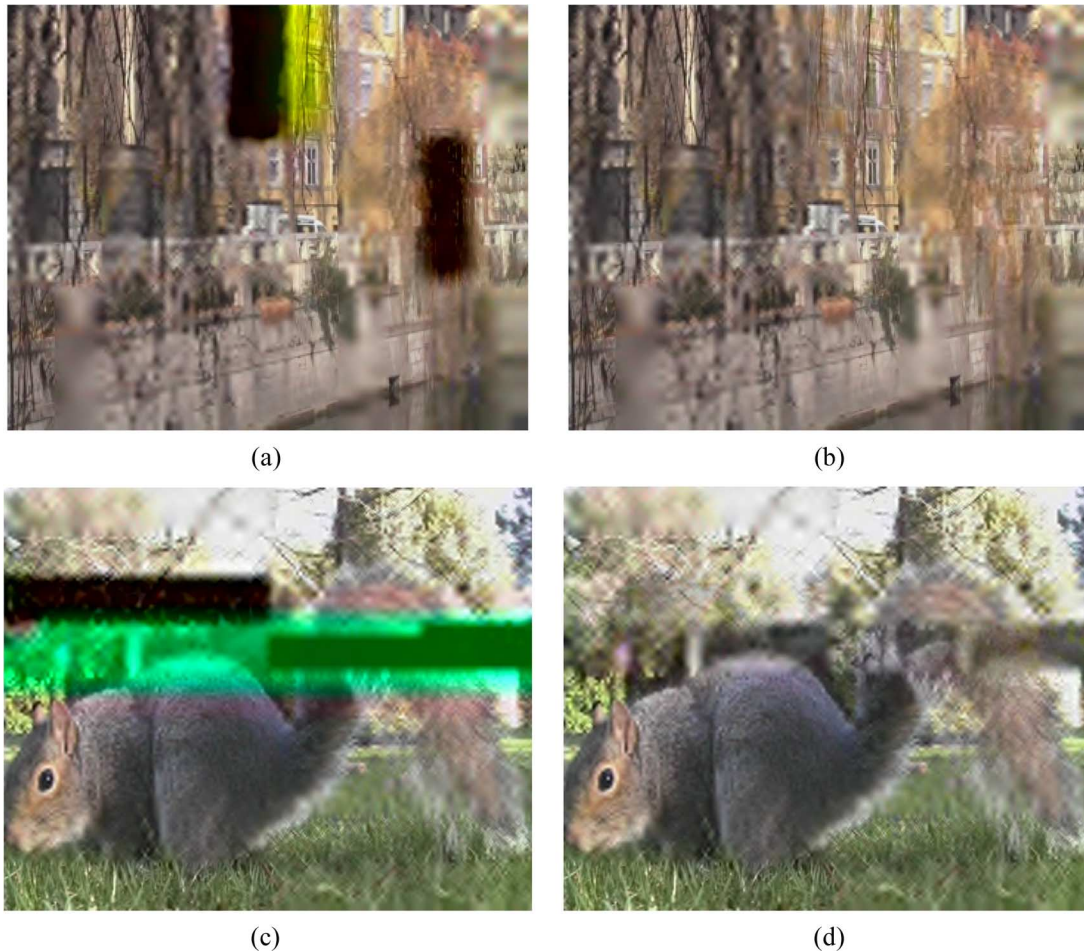
Fig. 18. Application of simple error concealment method: (a) frame number 37 (I-frame) of 33-partitioned format, CIF; (b) after applying error concealment method to Fig. 18(a); (c) frame number 12 (I-frame) of 18_6 partitioned format, SD576; (d) after applying error concealment method to Fig. 18(c).

mode and packet length is set to 30 bytes, which is approximately the same value as in the Dirac. H.264 rate control is used to regulate the bitrate at 370 kbps matching the bitrate generated by the Dirac encoding with 33-partitioned without FEC. In this comparison, both Dirac and H.264 are not using FEC and error concealment since the focus is to evaluate their error-resilient capability. The results of the two codecs are shown in Fig. 19.

As expected, Dirac with 33-partitioned cannot survive even a small packet loss and its average PSNR falls very sharply while the H.264's PSNR decreases gradually with the increase in packet loss. This shows that the Dirac source coding (data partitioning) only is not enough for the error-resilient transmission of wavelet compressed bitstream and the requirement of FEC becomes crucial in this case. Due to the nature of wavelet transform, the coefficients from the lowest subbands are sensitive to channel error and even a few corrupted bits in these subbands, error become prominent on the reconstructed frame because of the magnifying effect that happens in the inverse wavelet transform process. The only way to protect these data is to use FEC for lower subband data and/or all subbands data in a frame in order to survive up to 30% packet loss with good video output (refer to Figs. 13 and 15). From the figure, it is interesting to see that the PSNR of Dirac is higher than H.264 in error free condition at 0% packet loss. It is because H.264 requires more bits in order to achieve error-resilient transmission, thus lower



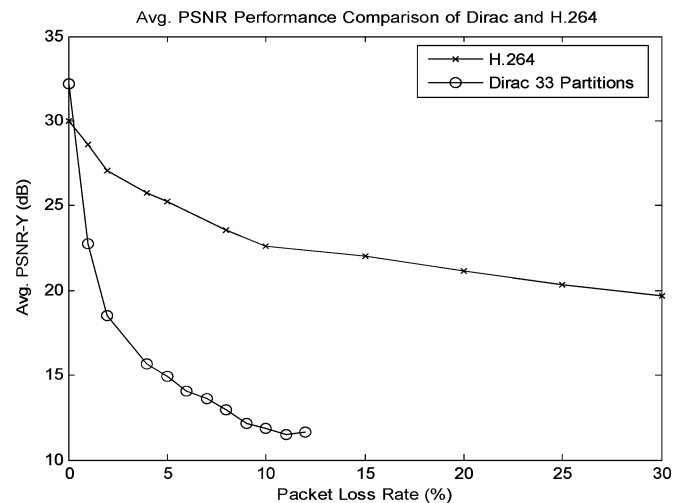Fig. 19. Average PSNR performance comparisons between H.264 and un-coded Dirac with 33-partitioned formats of CIF Canal-vertical sequence.

PSNR than Dirac since both codecs are having same bitrate of 370 kbps. The result also suggests that both Dirac and H.264 require to use FEC to protect the compressed bitstream in the packet-erasure channel in order to maintain video quality at least 25 dB.

## VIII. Conclusion

This paper has proposed a possible error-resilient scheme for Dirac video codec and has investigated its performance in the packet-erasure channel. The proposed scheme with 33-partitioned is recommended since it offers good compression efficiency and acceptable PSNR gain. It could achieve at least 5 dB over the un-partitioned format with/without FEC. A better error-resilient performance can be achieved if the maximum possible number of partitions 99 is used at the expense of lower compression efficiency. Even though the coefficient partitioning method could achieve high PSNR gain over the un-partitioned format, the compressed bitstream still requires protection from channel error by FEC. This is due to the nature of wavelet transform where the lower subband data is rather sensitive to error; as the result, the average PSNR of the unprotected bitstream fall very sharply even for a small packet loss as shown in Fig. 19. It is crucial to protect the sensitive wavelet compressed data at least to the lower subbands in order to survive in high packet loss environment. The simulation results suggested that with the use of less powerful FEC e.g. RCPC, Dirac could only operate in a congested network up to 10% packet loss. Instead, using a more powerful FEC not limited to Turbo codes, but also turbo product code (TPC), low density parity check code (LDPC) etc., Dirac could operate in a congested network up to 30% packet loss. This paper concludes that the coefficient partitioning process itself does not introduce much complexity to the encoder, and the use of FEC as a whole offer a simple and effective error-resilient scheme for Dirac to combat the harsh packet-erasure channel/network. Furthermore, the advent of network-on-chip with embedded high throughput FEC [20]–[22] provides the mean in applying Dirac to transport large volume of compressed video files to end-users through networked media applications e.g. on-demand video, IPTV etc.

## Appendix

Generator Matrix for Convolutional Encoder

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Puncturing Matrices used for Puncturing the output of the Convolutional Encoder

Rate 2/3

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rate 1/2

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rate 1/3

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rate 1/4

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

## Acknowledgment

## References

[1] "H.264 Video Coding License Fees," [Online]. Available: http://www.vialicensing.com/products/AVCH264VC/license.terms.html

[2] EBU Technical Review, T. Borer and T. Davies, "Dirac—Video compression using open technology," [Online]. Available: http://www.ebu.ch/en/technical/trev/trev_303-borer.pdf

[3] J. Zheng and L. P. Chau, "Multiple description coding using multiple reference frame for robust video transmission," in *IEEE Int. Symp. Circuits and Systems, ISCAS 2005*, May 2005, vol. 4, pp. 4006–4009.

[4] C. M. Chen, Y. C. Chen, and C. M. Chen, "Multiple description motion compensation video coding for MPEG-4 FGS over Lossy packet networks," in *IEEE Int. Conf. Multimedia and Expo, ICME2004*, Jun. 2004, vol. 2, pp. 1143–1146.

[5] M. Schaar and D. S. Turaga, "Multiple description scalable coding using wavelet-based motion compensated temporal filtering," in *IEEE Int. Conf. Image Processing, ICIP2003*, Sept. 2003, vol. 2, pp. 489–492.

[6] C. D. Creusere, "A new method of robust image compression based on the embedded zerotree wavelet algorithm," *IEEE Trans. Image Processing*, vol. 6, pp. 1436–1442, Oct. 1997.

[7] W. Tan and A. Zakhor, "Resilient compression of video for transmission over the internet," in *32nd Asilomar Conf. Signals, Systems and Computers*, Nov. 1998, vol. 1, pp. 243–247.

[8] P. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Processing Lett.*, vol. 4, pp. 189–191, July 1997.

[9] W. R. Heinzelman and M. B. R. Talluri, "Unequal error protection of MPEG-4 compressed video," in *Int. Conf. Image Processing, ICIP99*, Oct. 1999, vol. 2, pp. 530–534.

[10] N. V. Boulgouris, N. Thomos, and M. G. Strintzis, "Transmission of images over noisy channels using error-resilient wavelet coding and forward error correction," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 12, pp. 1170–1181, Dec. 2003.

[11] J. Kim, R. M. Mersereau, and Y. Altunbasak, "Error-resilient image and video transmission over the internet using unequal error protection," *IEEE Trans. Image Processing*, vol. 12, no. 2, pp. 121–131, Feb. 2003.

[12] S. Cho and W. A. Pearlman, "A full-featured, error-resilient, scalable wavelet video codec based on the set partitioning in hierarchical trees (SPIHT) algorithm," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 157–171, Mar. 2002.

[13] B.-J. Kim, Z. Xiong, W. A. Pearlman, and Y. S. Kim, "Progressive video coding for noisy channels," *J. Vis. Commun. Image Repres.*, vol. 10, pp. 173–185, 1999.

[14] Z. Xiong, B.-J. Kim, and W. A. Pearlman, "Progressive video coding for noisy channels," in *Proc. IEEE Int. Conf. Image Processing (ICIP '98)*, Oct. 1998, vol. 1, pp. 334–337.

[15] J. Zheng and L. P. Chau, "Error-resilient coding of H.264 based on periodic Macroblock," *IEEE Trans. Broadcasting*, vol. 52, no. 2, pp. 223–229, Jun. 2006.

[16] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC Codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, Apr. 1988.

[17] C. Berrou, A. Glavieux, and P. Thitimasjshima, "Near Shannon limit error correcting coding and decoding: Turbo codes," in *Proc. of IEEE Int. Conf. on Commun.*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[18] "H264/AVC JM12.2 Reference Software," [Online]. Available: http://iphome.hhi.de/suehring/tml/

[19] K. Onthriar, K. K. Loo, and Z. Xue, "Performance comparison of emerging Dirac video codec with H.264/AVC," in *Int. Conf. on Digital Telecommunications (ICDT'06)*, 2006, p. 22.

[20] C. Neeb, M. J. Thul, and N. Wehn, "Network-on-chip-centric approach to interleaving in high throughput channel decoders," in *IEEE Int. Sym. on Circuits and Systems (ISCAS'05)*, May. 2005, pp. 1766–1769.

[21] A. Giulietti, B. Bougard, V. Derudder, S. Dupont, J.-W. Weijers, and L. Van der Perre, "A 80 Mb/s low-power scalable turbo codec core," in *Proc. IEEE Custom Integrated Circuits Conference*, 2002, pp. 389–392.

[22] T. Theocharides, G. Link, N. Vijaykrishnan, and M. J. Irwin, "Implementing LDPC decoding on network on chip," in *Proc. of VLSI Design*, Jan. 2005, pp. 134–137.

**Kok-Keong Loo** (M'01) a.k.a. Jonathan Loo received his MSc degree (Distinction) in Electronics at University of Hertfordshire, UK in 1998 and PhD degree in Electronics and Communication at the same university in 2003. After completing his PhD, he works as a lecturer in multimedia communications at Brunel University, UK. He is also a course director for MSc Digital Signal Processing. Besides that, he currently serves as principle investigator for a joint project between Brunel University and British Broadcasting Corp (BBC) on the Dirac research and development. He also serves as research engineer for the IST-FP6 PLUTO project. His current research interests include visual media processing and transmission, digital/wireless signal processing, software defined radio, and digital video broadcasting and networks.

**Myo Tun** received his BEng in Electronics from Yangon Institute of Technology, Yangon, Myanmar and M.Eng in Telecommunication from Asian Institute of Technology, Bangkok, Thailand in 1996 and 2000 respectively. Currently, he is a PhD candidate in the School of Engineering and Design, Brunel University, UK. His research interests include video compression and transmission, error control coding, multi-carrier CDMA, wireless communication.

**John Cosmas** (M'86) received his BEng in Electronic Engineering at Liverpool University, UK in 1978 and a PhD in Image Processing at Imperial College, University of London, UK in 1986. He worked for five years in industry first with Tube Investments and then with Fairchild Camera and Instruments. After completing his PhD, he worked for 13 years as a lecturer in digital systems and telecommunications at Queen Mary College, University of London. Since 1999, he has worked for Brunel University, first as a reader and then in 2002 as a Professor of multimedia systems. His current research interests are concerned with multimedia broadcast communications systems, which evolved from his longstanding interests in video codecs (MPEG-1/2/4) and mobile communication systems.

Prof. Cosmas currently serves as Associate Editor for IEEE TRANS. BROADCASTING. He also contributes to the Digital TV Group's "Mobile Applications" sub-group and to the Digital Video Broadcast's Technical Module: Converged Broadcast and Mobile Services (CBMS).