

**INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO**

ISO/IEC JTC1/SC29/WG11

MPEG2007/M14468

April 2007, San Jose, USA

Title **Performance of a Distributed Video Codec behaviors in Presence of Transmission Errors**

Status **Informative**

Authors **Riccardo Leonardi, Claudia Tonoli, Pierangelo Migliorati, Nicola Adami (University of Brescia)**
Contacts: firstname.lastname@ing.unibs.it

Abstract

Distributed Video Coding (DVC) is one of the most important and active research fields in video coding. The basic idea underlying DVC is to exploit the temporal correlation among frames directly in the decoding phase. The main properties of a distributed video coding system is that the computational load could in principle be shifted towards the decoder, with respect to a traditional video coding system.

Anyway, the distributed coding approach has other interesting properties. In particular, one of the most promising benefits derived by the use of DVC is its natural error resilience to channel errors. Nevertheless, very few results on the actual error resilience properties of distributed video coding systems have been presented in literature. In this contribution we present a detailed analysis of the error resilience properties of a video coding system based on Stanford architecture. We analyze the behavior of such codec in presence of channel error, first focusing on the effect of such errors on the different parts of the encoded stream, and then making a preliminary comparison with H264.

1. INTRODUCTION

In recent years Distributed Source Coding (DSC) has become one of the most important research fields. The Distributed Source Coding paradigm has been developed following the guideline given by two fundamental results of Information Theory: the Slepian-Wolf theorem (1972) and the Wyner-Ziv (1975) theorem.

The main idea underlying DSC is that the correlation between two sources can be exploited only at the decoder, with ideally no performance loss with respect to the traditional encoding, that basically exploits the correlation also in the encoding phase. The direct consequence of the application of this principle is that the computational load is shifted toward the decoder, and the encoder is characterized by a very low computational complexity.

Specifically, in the video coding application of DSC, namely the DVC (Distributed Video Coding, DVC), neither motion estimation, nor other complex operation are required at the encoder. On the other hand, the decoder is more complex than in traditional coding, because the complex operation, computationally analogous to motion estimation, must be performed in the decoding phase.

Because of this shift of computational load, DVC is particularly well suited for system requiring low power and low cost encoding devices, for example video-surveillance, biomedical applications, 3D reconstruction applications.

The low complexity of encoding is the first direct consequence of the exploitation of correlation in the decoding phase. Anyway, from the fact that no prediction-based encoding is performed another important property should come: distributed coding system should have an inner robustness to transmission errors. Nevertheless, few studies about error resilience properties of distributed video codecs have been presented in the literature. In this work an analysis of the error resilience performances of a mono-camera distributed video codec is presented.

We considered a video codec based mainly on the DVC scheme proposed in [AZG-02] and further developed within the DISCOVER project.

We simulated the transmission of the encoded stream over a noisy channel to evaluate the error resilience performance of the codec.

The contribution is organized as follows. In section 2 we briefly introduce the architecture of the considered codec and the underlying main ideas. In section 3 the performed tests are explained in detail. We first analyze the effect of errors on

different parts of the encoded stream, then we present a preliminary comparison with H264 performances. Concluding remarks are given in the final section.

2. THE CONSIDERED VIDEO CODEC ARCHITECTURE

In this section we introduce the Wyner-Ziv coding principles and the architecture of the monoview distributed video codec we have considered in our tests. We briefly describe the main features of the considered coding scheme, mainly taking into account those which are responsible for the error resilience performances, and refer the reader to the references for an in depth description of the video coding system.

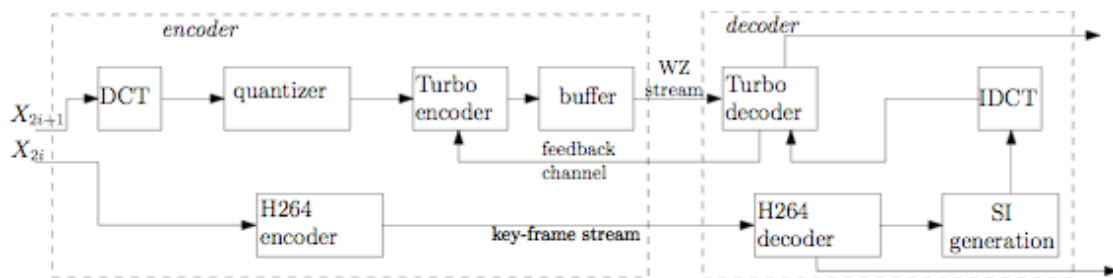


Figure 2.1.1 - Architecture of the considered codec.

The considered distributed codec is based on a Wyner-Ziv (WZ) coding scheme: a source X is encoded under the assumption that the decoder has access to some information that the encoder does not know and that is related to X . Based on this information, the decoder can perform a first prediction of X , referred to as Side Information (SI).

The main idea underlying the WZ coding is in the mechanism of correction of the SI in order to recover the original information X : in the encoding phase, the parity bits of a systematic channel code are computed on the original information and are sent to the decoder; the decoder joins this parity bits with the SI and performs the decoding of the channel code.

The monoview DVC scheme considered in our tests is based on an application of this idea. As shown in the block diagram shown in Figure 1.1.1, each frame is encoded independently, as temporal inter-frame correlation is going to be exploited at the decoder side.

The frame sequence is divided in two groups: the odd indexed frames are Wyner-Ziv coded, while the even indexed frames, referred to as the key-frames, are traditionally encoded, using H264 in intra mode, and are used for the Side Information generation.

In more detail, the WZ frames are first transformed, with a block based DCT, and quantized with proper quantization matrices applied to the DCT coefficients. The quantized coefficients are then encoded, bit-plane by bit-plane, with the use of a turbo code. In particular, each bit-plane is fed to a turbo code with rate $1/3$; while information bits are discarded, the parity bits are stored in a buffer. The encoded

stream is therefore composed of two different parts: the H264 intra coded key-frame stream, and the WZ parity bits stream.

As stated before, the decoding of a WZ frame is performed by first generating the Side Information (SI). In order to do that, a motion compensated interpolation between the two adjacent key-frames is carried out. A turbo decoding process is then applied to correct these estimations, using the parity bits received from the encoder, bit-plane by bit-plane.

An important remark about this architecture must be done, in order to better understand the codec behavior in presence of channel errors: parity bits are not sent all at once, but are iteratively requested by the decoder, using the feedback channel, until the estimated error probability on the decoded bit-plane reaches a given threshold. In this way, the decoder can approach a given target quality in the reconstructed frame, while the impact of errors reflects in the total bit-rate sent by the encoder.

3. ERROR RESILIENCE PERFORMANCE

Error resilience is probably one of the most important properties of distributed video coding systems. We have tested the error resilience properties of the DISCOVER codec studying its behavior on an error-prone transmission channel. In particular, the DISCOVER encoded bit-stream is composed of two different parts: the key-frames and the Wyner-Ziv frames.

First of all we investigate the effect of errors on each part independently, supposing first that only WZ frames are affected by transmission error, and then that WZ frames are received correctly, while key-frames are affected by channel errors.

We have then compared the performance of the DISCOVER codec with respect to an H264 codec, constrained to the GOP2 encoding mode.

In our test we have always assumed the feedback channel to be error-free.

3.1 Effects of errors on WZ bit-stream

First of all, we analyze the effect of channel errors on the WZ part of the encoded stream. We have simulated the transmission of the WZ bit-stream at different Bit Error Rates (BER), and supposed that the key-frame were received correctly, as described in Figure 3.1.1. We plotted the PSNR curve, considering 8 different working points. We have considered two different error conditions: the random packet loss and the random bit errors.

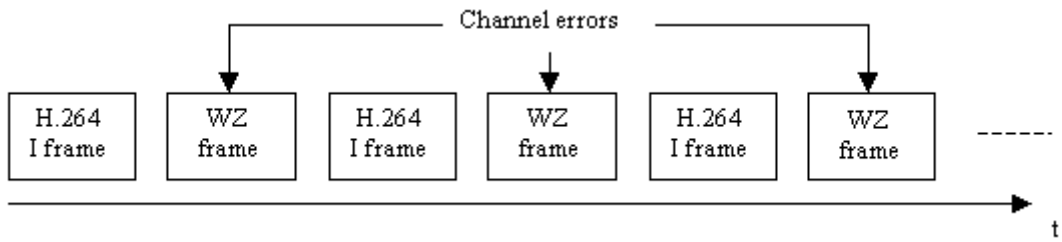


Figure 3.1.1 - Error corruption of WZ frames.

Random packet loss

First of all, we consider the transmission of the WZ stream over a packet network, characterized by a given Packet Loss Rate (PLR). We suppose that the decoder knows exactly which packets are lost, and asks the encoder for them to be retransmitted. In this case, the rate increment depends only on the PLR and it is not affected by the packet size. The plots in Figures 3.1.2, 3.1.3, 3.1.4 shows the PSNR versus the rate, for *Coastguard*, *Foreman* and *Hallmonitor*, respectively, at different PLR.

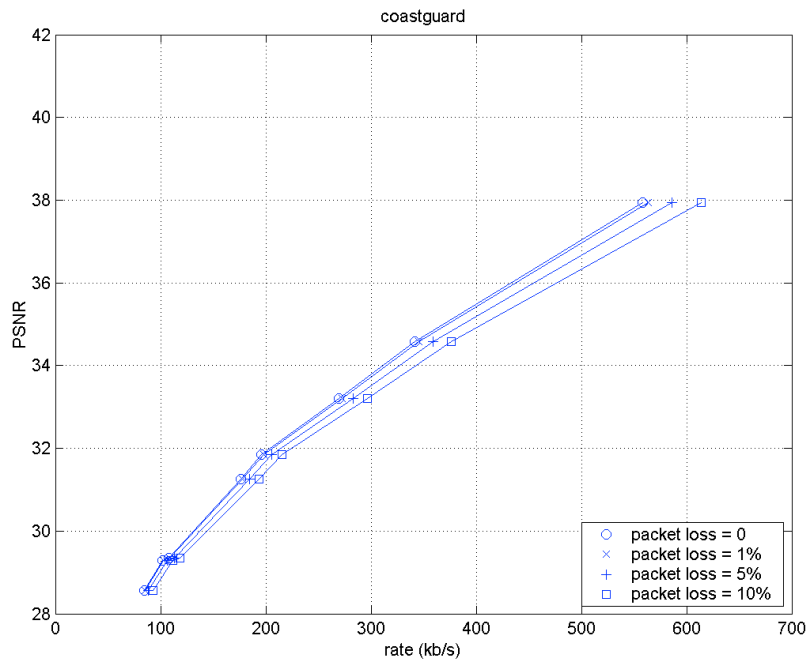


Figure 3.1.2 - WZ packet loss (Coastguard sequence).

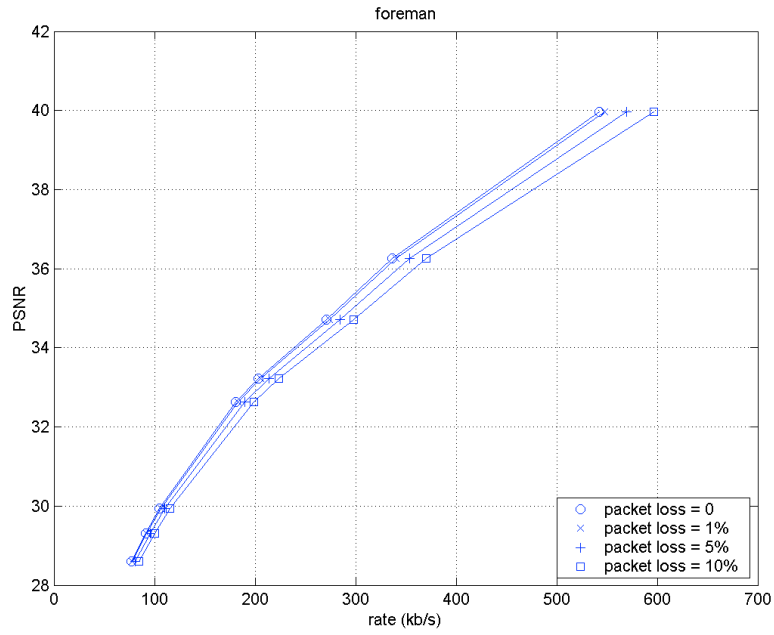


Figure 3.1.3 - WZ packet loss (Foreman sequence).

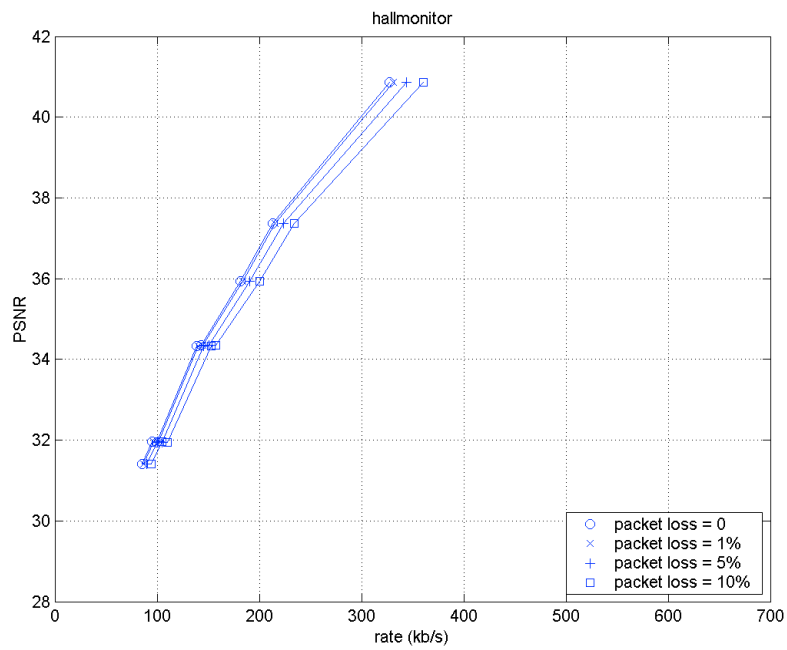


Figure 3.1.4 - WZ packet loss (Hallmonitor sequence).

Bit per bit errors

We consider now the case of the presence of random bit errors on the WZ stream.

The simulation results are shown in Figures 3.1.5, 3.1.6, 3.1.7, for the *Coastguard*, *Foreman*, and *Hallmonitor* sequences respectively.

In Figures 3.1.5 (a), 3.1.6 (a), 3.1.7 (a) the PSNR plots are computed taking into account only the reconstructed WZ frames, while in Figures 3.1.5 (b), 3.1.6 (b), 3.1.7 (b) the average PSNR on the whole reconstructed sequence (WZ frames and key-frames are considered). Of course we don't consider the key-frames alone, because they are supposed to be received correctly, and then their rate-distortion curve is not affected by the channel BER.

The plots show that the DISCOVER codec can compensate quite well channel errors on WZ bits. In fact, at a BER of 10^{-3} the performance loss is very low in each one of the considered working points, for all the test sequences. At higher BER (10^{-2} , $5 \cdot 10^{-3}$) the performance loss is still smaller than 0.5 dB at low bit-rates, and it only increases slightly at higher bit-rates.

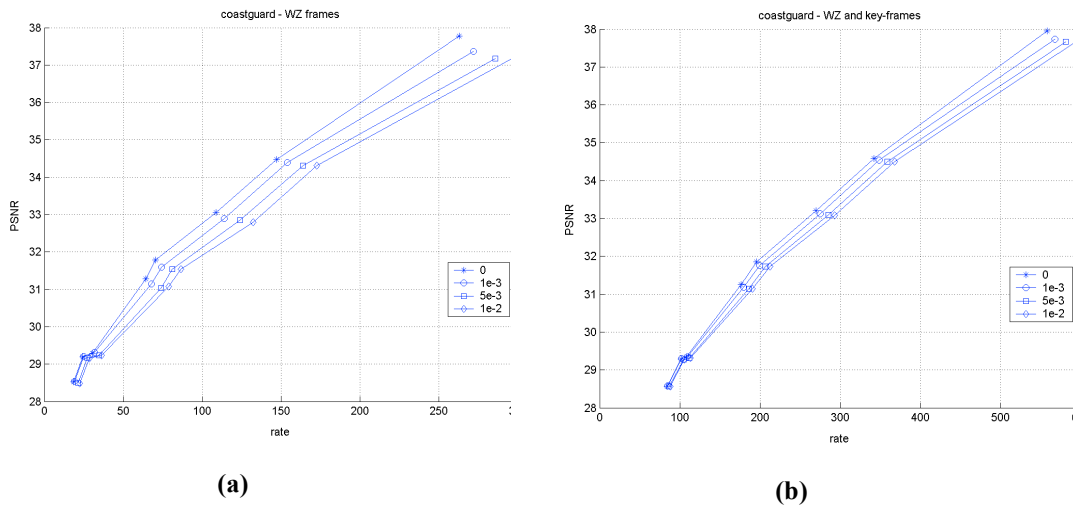


Figure 3.1.5 - Bit errors on WZ frames (Coastguard sequence).

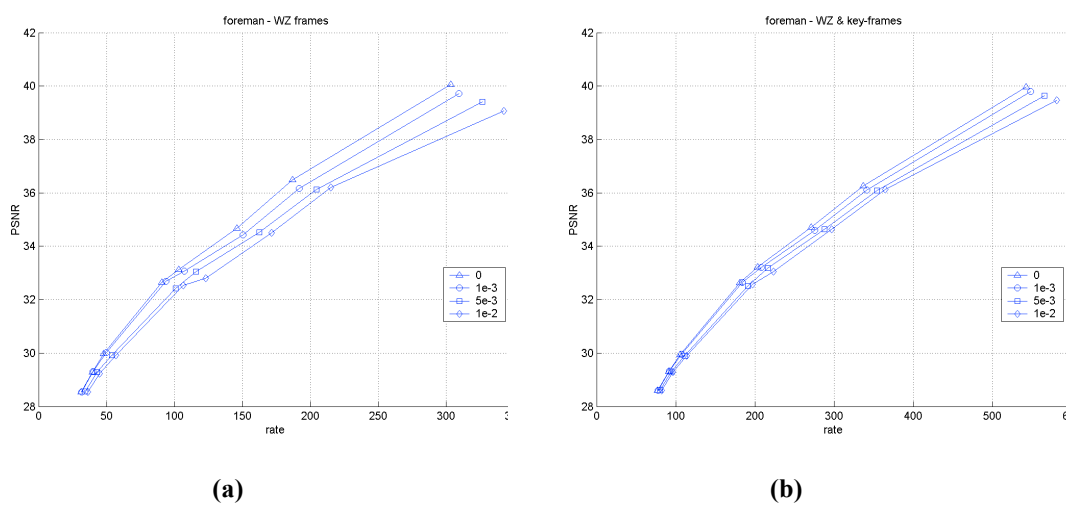


Figure 3.1.6 – Bit errors on WZ frames (Foreman sequence).

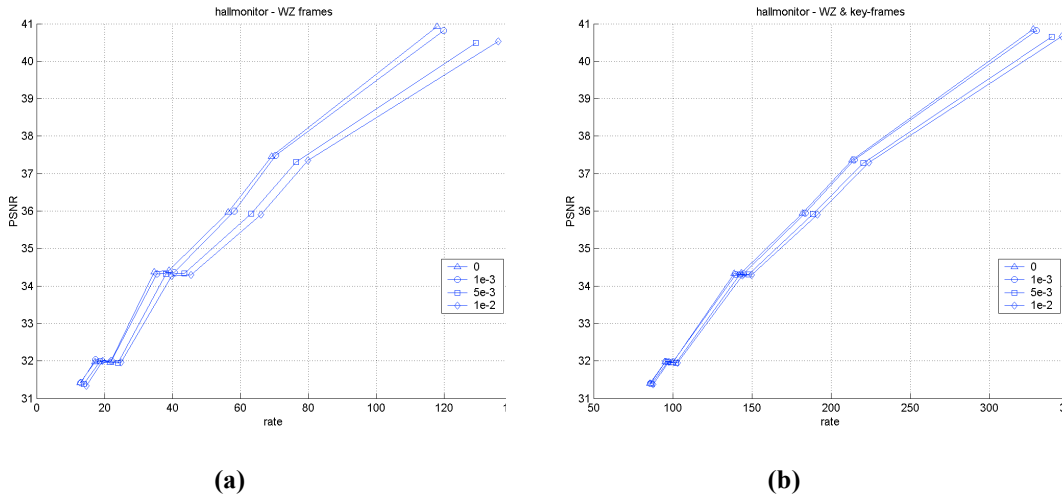


Figure 3.1.7 - Bit errors on WZ frames (Hallmonitor sequence).

3.2 Effects of errors on key-frames

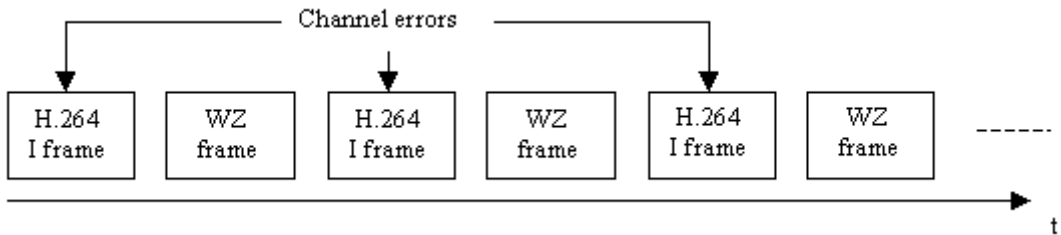


Figure 3.2.1 - Error corruption of key-frames.

We have then evaluated the effect of transmission errors on the key-frames bit-stream. As the key-frames are encoded in H264 intra mode, than the study of the distributed codec performance on error prone channels necessarily involves the error resilience features of the H264 standard codec.

H264 has various error concealment strategies, mainly based on the flexible reordering of the macro-blocks inside the slices at the encoder (Flexible Macro-Block Reordering, FMO), and on the estimation of lost slices at the decoder. The H264 implementation used in our test is JM11.0. We enabled the error concealment at the decoder, which is necessarily performed in intra mode, because the frames are intra-coded.

To simulate the transmission of the key-frames over an error-prone channel, we introduced random errors on the bit-stream, and dropped those packets that have resulted corrupted by the presence of at least one bit error.

First of all, we have analyzed the behavior of the codec considering jointly the plots in Figures 3.2.2 (a) and (b), depicting, respectively, the rate and the PSNR frame by frame, for the first 149 frames of the *Foreman* sequence, encoded at an average PSNR

of 34.5 dB, at the 5th working point. The key-frames that are affected by packet losses are characterized by a PSNR much lower than the correctly received ones (e.g., frame n. 36). Let us consider the encoded rate of the WZ frames that are adjacent to a corrupted key-frame (e.g., frames n. 35 and 37): we can notice that such frames require more bits than they do in the case of no transmission errors. Anyway their quality is much better than that of the corrupted key-frames: this means that the decoder can react positively to a bad side information by asking more parity bits to the encoder, using the feedback channel.

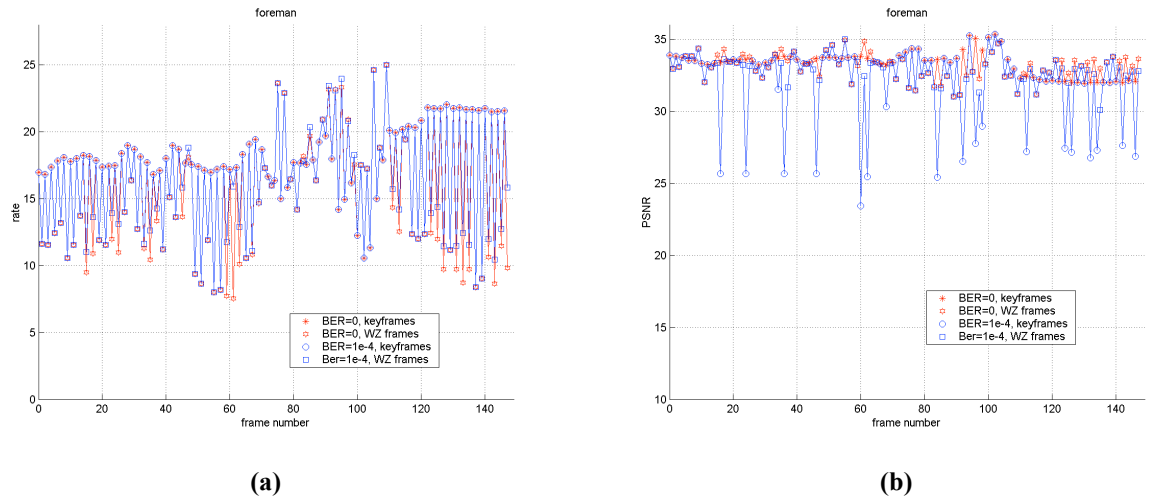
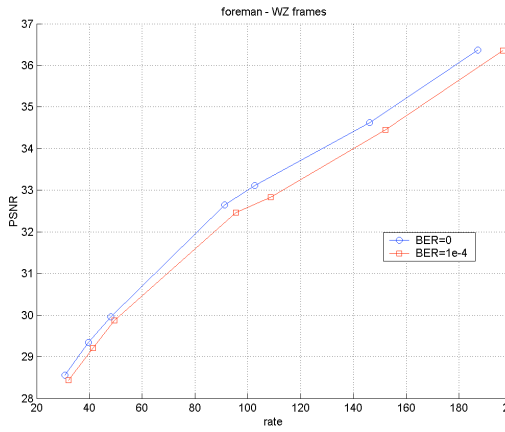


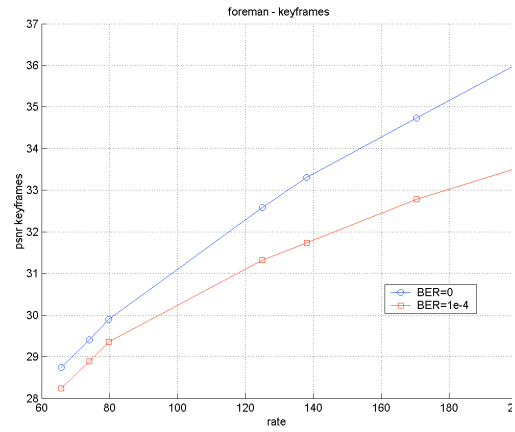
Figure 3.2.2 - Rate (a) and PSNR (b) (Foreman sequence).

In Figures 3.2.3 (a), 3.2.4 (a), 3.2.5 (a) the rate-distortion plot for a BER of 10^{-4} , corresponding to a packet loss rate of about 4%, is shown, for the WZ frame of the *Foreman*, *Hallmonitor* and *Coastguard* sequences, respectively. We can notice that at low bit-rates the performance loss is about 1 dB, while at higher bit-rates it increases, especially for the *Coastguard* and *Hallmonitor* sequences.

The plots in Figure 3.2.3 (b), 3.2.4 (b), 3.2.5 (b) show the PSNR versus the rate for all the working points of the *Foreman*, *Hallmonitor* and *Coastguard* sequences respectively. These figures show that the PSNR loss of the corrupted key-frames at high bit-rates is much lower than the one at low bit-rates. This means that the feedback mechanism of the encoder can compensate quite well the presence of low quality side information, especially a low bit rates.

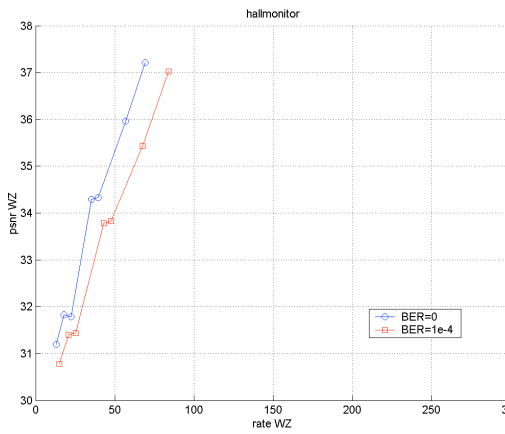


(a)

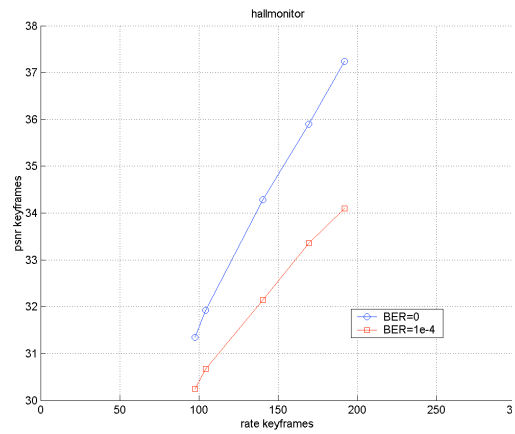


(b)

Figure 3.2.3 - Errors on key-frame bit-stream – reconstructed WZ frames (a) and key-frames (b) (Foreman sequence).



(a)



(b)

Figure 3.2.4 - Errors on key-frame bit-stream – reconstructed WZ frames (a) and key-frames (b) (Hallmonitor sequence).

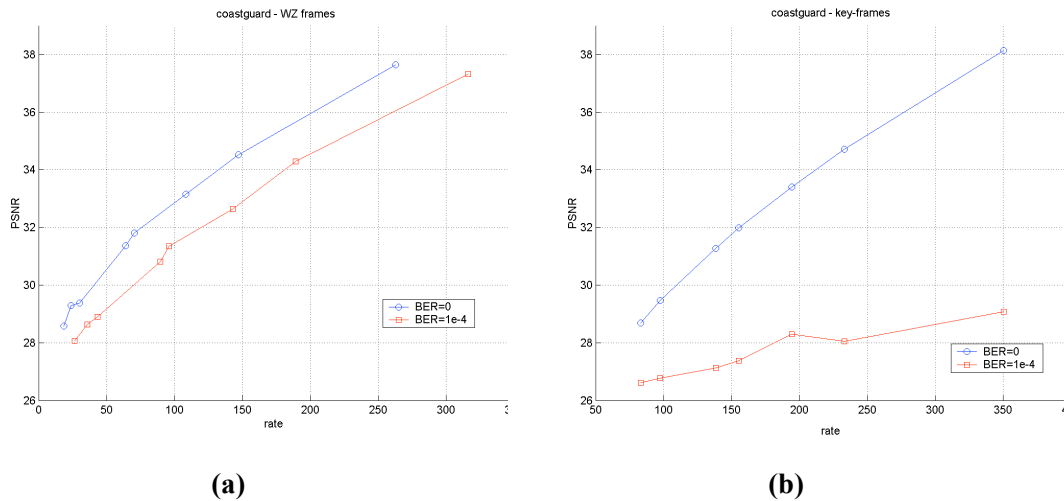


Figure 3.2.5 – Errors on key-frame bit-stream – reconstructed WZ frames (a) and key-frames (b) (Coastguard sequence).

We have also evaluated the behavior in presence of bad quality side information, i.e., when no concealment is applied and lost blocks are filled with a uniform grey level.

In Figures 3.2.6 (a) and 3.2.6 (b) we consider the PSNR versus rate for the reconstructed WZ frames and key-frames respectively, for the *Hallmonitor* sequence. These results show the importance of having good reconstructed key-frames, especially at high bit-rates. In fact, the feedback mechanism can partially compensate the bad quality of the side information, but the overall reconstructed quality is still several dBs lower than the original.

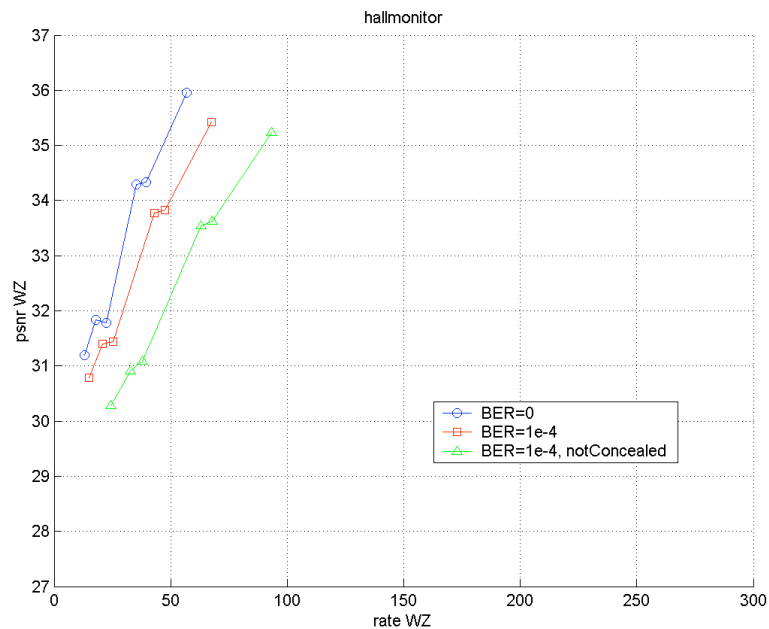


Figure 3.2.6 (a) - Errors on key-frames not concealed - PSNR of WZ frames.

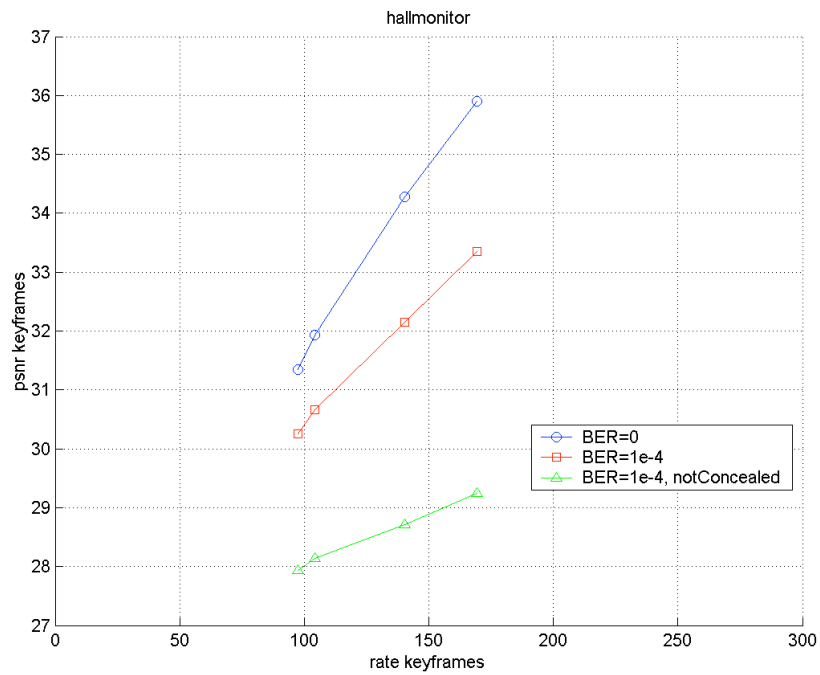


Figure 3.2.6 (b) - Errors on key-frames not concealed - reconstructed key-frames.

3.3 Corruption of both parts of the bit-stream

We have considered also the case of corruption of both parts of the bit-stream, simulating the transmission of the whole encoded stream over an error prone channel. We have considered bit errors on the WZ stream, as described previously, and the dropping of corrupted packets for the key-frame stream, as previously described.

The simulation results are shown in Figures 3.3.1, 3.3.2, 3.3.3.

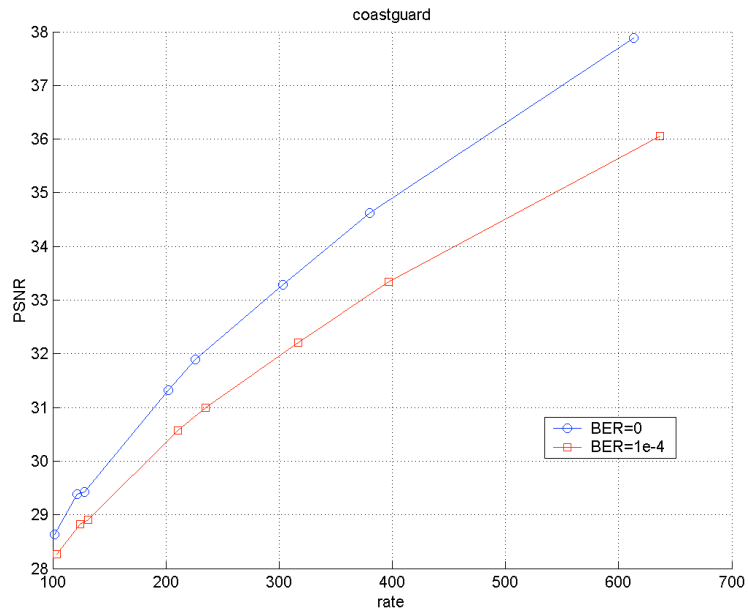


Figure 3.3.1 - Errors on both parts of the stream (Coastguard sequence).

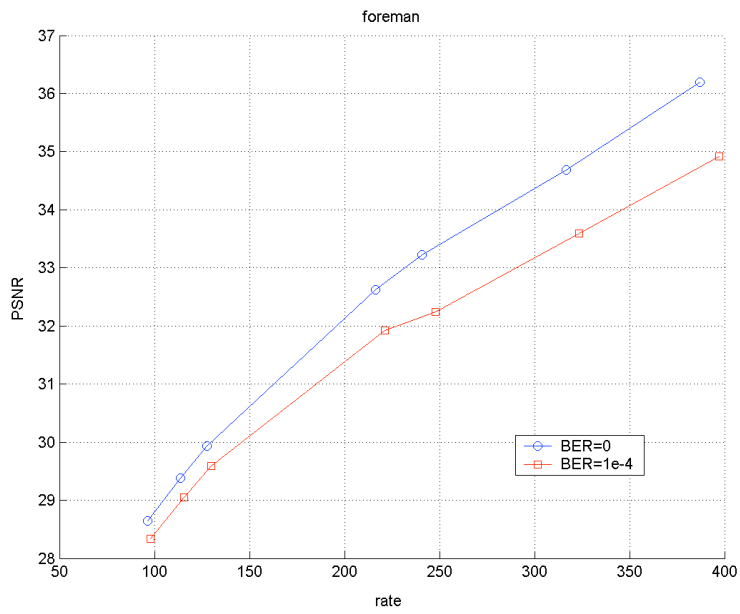


Figure 3.3.2 - Errors on both parts of the stream (Foreman sequence).

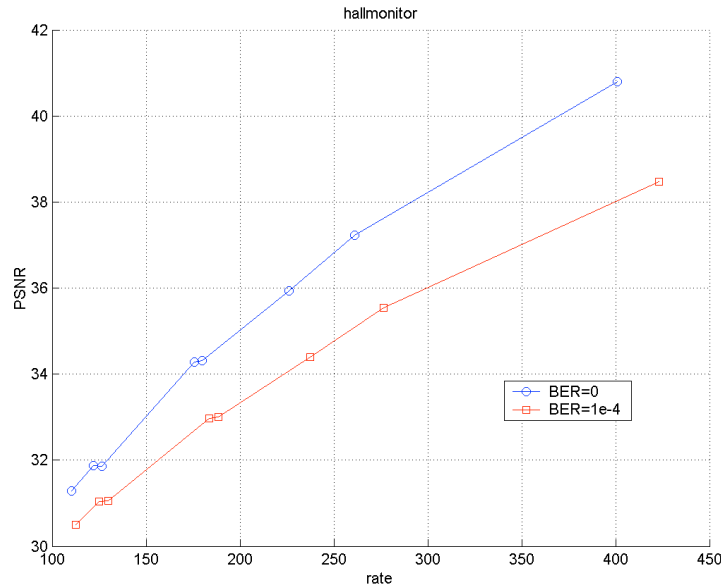


Figure 3.3.3 - Errors on both parts of the stream (Hallmonitor sequence).

3.4 Comparison with H.264-AVC

We have also evaluated the error resilience performance of the DISCOVER codec, comparing the rate-distortion curve in presence of channel errors with those of an implementation of the H264 standard, namely the JM11. For this test we have used a GOP2 encoded in IBIBIBI mode, which is a working condition comparable with that of the DISCOVER codec. We have used the same concealment method already used for the evaluation of the error resilience performance in presence of errors on the key-frames only; we enabled the FMO features and we turned on the intra concealment at the JM decoder.

In Figure 3.4.1 is reported the comparison between DISCOVER and H264 GOP2, for the *Foreman* sequence.

As we can see, at higher bit-rates the performance is about 3 dB under H264, while at the lower bit-rates the DISCOVER performance loss, with respect to H264 is about 1dB. The results for the *Coastguard* sequence, plotted in Figure 3.4.2, show that in this case at each working point the DISCOVER curve is less than 1 dB lower that the H264 curve at each considered point.

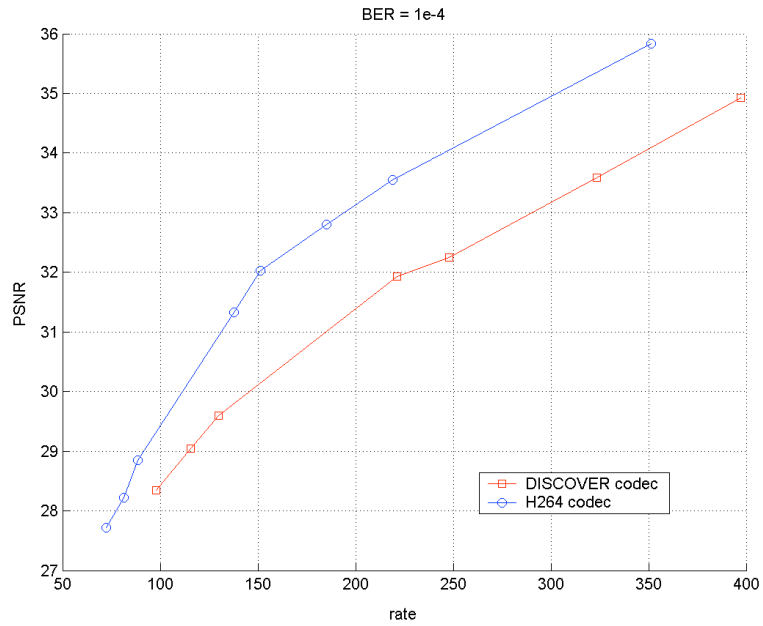


Figure 3.4.1 - Comparison with H264 GOP2 (Foreman sequence).

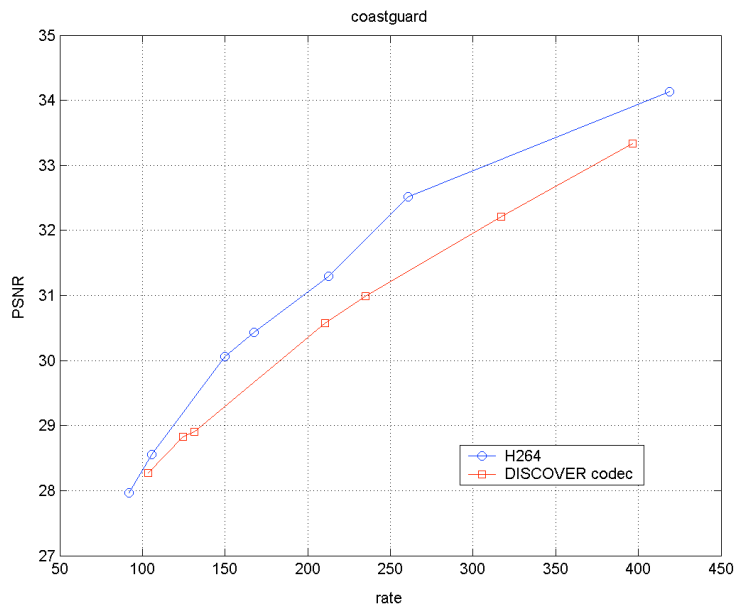


Figure 3.4.2 - Comparison with H264, GOP2 (Coastguard sequence).

Conclusion

In this contribution we have presented an analysis of the behavior of a distributed mono-view video codec in presence of channel errors. First of all, we have considered separately the effect of errors on the two components of the encoded bit-stream: the WZ bit-stream and the key-frame bit-stream. We have simulated the transmission over a noisy channel of the WZ part of the bit-stream: the obtained results show that this component is characterized by a good robustness to transmission errors, especially at low encoding rates. We have then considered the effect of errors on the key-frame part of the bit-stream, showing that the codec can also compensate low

quality side information by asking more parity bits for reconstructing good quality WZ frames. Finally, we have presented a comparison between the error resilience performance of the considered distributed codec and the performance of H264, in a GOP2 IBIBI mode. In this case the performance of the H264 codec remain superior than the performance of the DVC codec, but likely the situation could be inverted for GOP size greater than 2 since the temporal prediction loop cannot recover from channel errors. Undergoing research is devoted to a performance comparison under such conditions.

Acknowledgments

The work presented in this contribution was developed within DISCOVER (Distributed Coding for Video Services), a European Commission Future and Emerging Technologies (FET) project, funded by the European Commission. DISCOVER software started from the IST-WZ software developed at the Image Group from Instituto Superior Tecnico (IST) of Lisbon by Catarina Brites, Joao Ascenso and Fernando Pereira.

References

[ABP-06] J. Ascenso, C. Brites, and F. Pereira, "Content Adaptive Wyner-Ziv Video Coding Driven by Motion Activity", *IEEE ICIP*, Atlanta, USA, October 2006.

[ARSG-04] A. Aaron, S. Rane, E. Setton and B. Girod, "Transform-Domain Wyner-Ziv Codec for Video", Visual Communications and Image Processing Conference, San Jose, CA, USA, January 2004.

[AZG-02] A. Aaron, R. Zhang, and B. Girod, "Wyner-ziv coding for motion video", Asilomar Conference on Signals, Systems and Computers, November 2002.

[BAP-06] C. Brites, J. Ascenso, and F. Pereira, "Improving Transform Domain Wyner-Ziv Video Coding Performance", *IEEE ICASSP*, Toulouse, France, May 2006.

[GARR-05] B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero, "Distributed Video Coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71-83, Jan. 2005.

[H264-03] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264 AVC Video Coding Standard", *IEEE Transaction on circuits and systems for video technology*, July 2003.

[H264-06] Sunil Kumar, Liyang Xu, Mrinal K. Mandal and Sethuraman Panchanathan, "Error Resiliency Schemes in H.264/AVC Standard", *Elsevier J. of Visual Communication and Image Representation*, Special issue on Emerging H.264/AVC Video Coding Standard, Vol. 17, April 2006.

[SJA-04] Anshul Sehgal, Ashish Jagmohan, and Narendra Ahuja, "Wyner-Ziv Coding of Video: An Error-Resilient Compression Framework", *IEEE Transaction on Multimedia*, vol.6, no.2, April 2004.

[SW-73] J. Slepian and J. Wolf, "Noiseless Coding of Correlated Information Sources", *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471-480, July 1973.

[WZ-76] A.D. Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. Information Theory*, vol. 22, no. 1, pp. 1-10, Jan. 1976.