

# A DISTRIBUTED VIDEO CODING APPROACH FOR MULTIPLE DESCRIPTION VIDEO CODING OF STEREO SEQUENCES

*S. Milani and G. Calvagno*

Dept. of Information Engineering, University of Padova, Italy  
e-mail: {simone.milani, calvagno}@dei.unipd.it

## ABSTRACT

Recent works have proposed the combination of Multiple Description Coding (MDC) schemes with the novel Distributed Video Coding (DVC) paradigm in order to enable robust video transmission over lossy channels. The paper presents a Multiple Description Distributed Video Coder (MDDVC) that combines a polyphase-subsampling MDC approach with a novel hybrid Distributed Source Coding technique that processes the video signal both in the pixel and in the transform domain. The proposed approach is applied to the transmission of stereo video sequences in order to avoid an exceeding visual quality degradation in presence of losses. Tests on simulated lossy channels show that at high loss probabilities the PSNR value of the reconstructed video sequence improves with respect to its traditional MDC counterpart.

## 1. INTRODUCTION

The transmission of video sequences over wireless networks presents several issues due to the presence of delays, packet losses and bandwidth limitations. Their drawbacks are significantly amplified whenever transmitting stereo or multiview video sequences since the visual quality of each view has to be preserved together with the quality of the disparity information [1]. In order to overcome these problems, different robust video coding strategies have been proposed in the recent literature. Among these, Multiple Description Coding (MDC) schemes [2] have proved to be significantly more efficient with respect to traditional single description coding (SDC) architectures for wireless networks, where channel conditions are time-varying and the bursty nature of packet losses can lead to a significant distortion of the reconstructed sequence. In relation to the transmission of stereo sequences, the MDC schemes have been recently employed generating different descriptions for each view or depthmap. Each description is associated to other descriptions from the other views in order to create separate independently-decodable streams that permit reconstructing the multi-view sequence at a lower quality.

In addition, several video coding architectures based on the information-theoretic concept of Distributed Source Coding (named Distributed Video Coding or DVC schemes) have

been proposed during the last years [3]. DVC approaches permit an independent coding of the data at the transmitter, and a predictive decoding at the receiver using any predictor among a set of possible candidates. In this way, the coded bit stream results more robust to losses with respect to traditional coding methods since decoding can be performed successfully provided that at least one of the candidate predictors is available. These schemes have also been adopted for the transmission of multi-view video sequences, since the DVC solution enabled the coded bit stream both with robustness and flexibility as it is possible to switch from one view to another more easily [4].

Recent works have been focusing on combining these two techniques in Multiple Description Distributed Video Coders (MDDVC) since coding the prediction residual using DVC permits reducing the drifting problem in traditional MDC [5]. The paper proposes a novel MDDVC scheme that adopts a hybrid DVC unit that processes the video signal both in the spatial and in the transform domain. More precisely, the correlation between blocks is estimated in the spatial domain, but the generated syndromes are compressed via a transform coding algorithm. Experimental results will show that the proposed approach permits improving the PSNR of the sequence reconstructed by the original MDC scheme up to 1 dB. Moreover, the proposed scheme proves to be extremely effective when applied to the transmission of stereo sequences. The DVC approach embedded in the coding architecture permits a successful decoding of the coded sequence exploiting the correlated information coming from the other view.

In the following, Section 2 provides an overview of different DVC and MDC solutions adopted for the robust transmission of stereoscopic sequences. Then, Section 3 presents the implemented codec, while Section 4 describes the Distributed Source Coding (DSC) scheme that have been adopted to characterize the prediction residual signal. Section 5 presents how the proposed coder is employed for the transmission of stereo video sequences over wireless channels affected by packet losses. Experimental results in Section 6 show how the proposed DSC approach improves the quality of the coded sequences at high loss rates. Conclusions are drawn in Section 7.

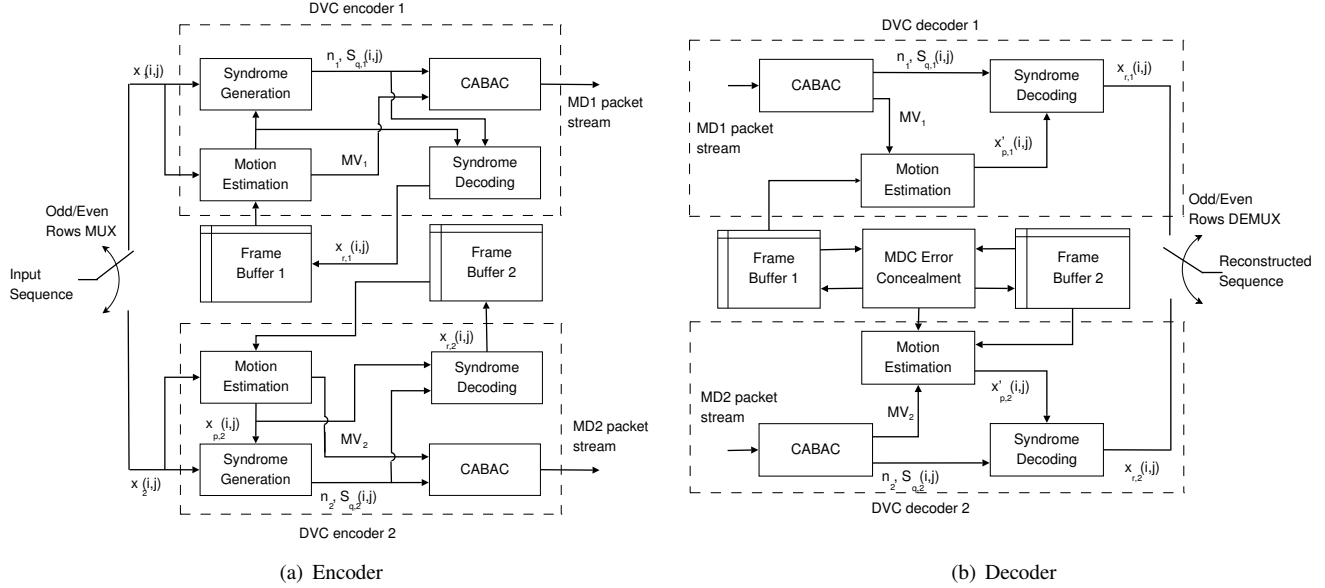


Fig. 1. Block diagrams for the proposed encoder and decoder for a single view of the stereoscopic video sequence.

## 2. PREVIOUS WORKS ON MDC AND DVC SCHEMES FOR STEREOSCOPIC VIDEO TRANSMISSION

The adoption of MDC schemes for the robust transmission of stereo video sequences has been recently exploited in different ways. In [1] the authors proposed an MDC scheme for stereoscopic video transmission based on spatial upsampling/downsampling. In the same paper the authors presented an MDC based on a temporal polyphase subsampling of the odd and even frames from different views, which is adopted in [6] to transmit a single view together with the relative depth map. Concurrently with MDC solution, several DVC schemes have been proposed for multiview video transmission. In [4] Yeo and Ramchandran presented a robust coder that can use different frames from different views to achieve a successful decoding of the transmitted data. In [7] a robust DVC coding scheme for 3DTV applications was proposed. Unfortunately, the computational complexity required at the decoder is significantly higher with respect to MDC schemes since the decoding process has to perform a motion-search on a huge set of candidates.

Many MDC schemes have been recently improved by including a DVC coding unit within their architecture. In [8] Crave *et al.* proposed an MDDVC approach based on motion compensation temporal filtering. In this case, the adopted DVC solution is able to mitigate the inefficiency of the original MDC scheme reducing the amount of redundancy introduced in the data stream. The solution proposed in [9] applies a Wyner-Ziv coding technique to a multiple description scalar quantizer, while the approach by Fan *et al.* [10] associates to each description some hash information, which is transmitted

with the packet stream of the other description. In addition, the approach presented in [11] by Wang *et al.* adopts a zero-padding of DCT coefficients to generate the different descriptions, which are coded by a SPIHT-based Slepian-Wolf coder.

The improved results obtained with the MDDVC approaches suggest that the application of these techniques for the transmission of stereoscopic and multi-view video sequences is an interesting field of research.

## 3. THE PROPOSED MDC ARCHITECTURE FOR SINGLE VIEWS

Many MDDVC schemes that have been presented in literature are derived from previous traditional MDC counterparts replacing the traditional temporally-predictive video coder with a Wyner-Ziv Video coding scheme (see [5] as an example). Similarly, the proposed coder relies on a traditional polyphase-subsampling MDC architecture and adopts a novel DVC video coding paradigm to code each description separately.

Figure 1 shows the block diagrams of the employed encoder and decoder for a single view of the stereo sequence. The input sequence is split into two descriptions sampling odd and even rows of pixels, which are sent to two separate H.264/AVC-based DVC coders. Each coder has been designed reusing the block structure of H.264/AVC, where the residual coding adopts a DSC-based algorithm in place of the traditional solution based on the Displaced Frame Difference (DFD).

Each  $4 \times 4$  block of each description is approximated by the Motion Estimation unit generating a predictor block identified by a Motion Vector. The residual signal is then coded

using a Distributed Source Coding paradigm (described in Section 4) creating two packet streams that are independently sent to the receiver. Each packet stream is decoded using two separate DVC decoders, which are helped by a Multiple Description Error Concealment unit that estimates the lost information. More precisely, in case one description is missing, it is possible to recover the lost rows interpolating the rows of the other description. In this way, the decoder obtains a degraded version of the coded description, which permits blocking the distortion propagation to the following frames. In fact, the generated Motion Vectors for the two description are quite close, and therefore, it is possible to assume that the motion information is completely recovered whenever at least one description is available. In case the concealed image proves to be sufficiently correlated with the lost data, the motion compensated blocks result close to the ones at the encoder, and the transmitted syndromes permits decoding correctly the video signal. The following section will provide more details about the syndrome generation and decoding units.

#### 4. A DISTRIBUTED SOURCE CODING APPROACH FOR CODING THE PREDICTION RESIDUALS

In the technical literature DVC approaches can be divided into two main groups: schemes that process the video signal in the pixel domain (see [12] as an example) and schemes that operate in the transform domain (like the coder by Aaron *et al.* [13] and the PRISM coder by Puri *et al.* [14]). The former require a lower computational complexity and prove to be effective when no feedback channel is available, but they have a limited compression efficiency. The latter provide higher compression ratios, but the needed computational effort is considerably higher. In our work, we adopted a hybrid pixel-transform domain DVC scheme that permits reducing the decoding complexity and improves the robustness of the video stream to packet losses with high compression gains.

Given the current  $4 \times 4$  block  $\mathbf{x}_m$  of pixels for description MD $m$  ( $m = 1, 2$ ) and its predictor  $\mathbf{x}_{p,m}$ , for each pixel  $x_m(i, j)$  of block  $\mathbf{x}_m$  at position  $(i, j)$ ,  $i, j = 0, \dots, 3$ , we compute the number of bits  $n_m(i, j)$  as

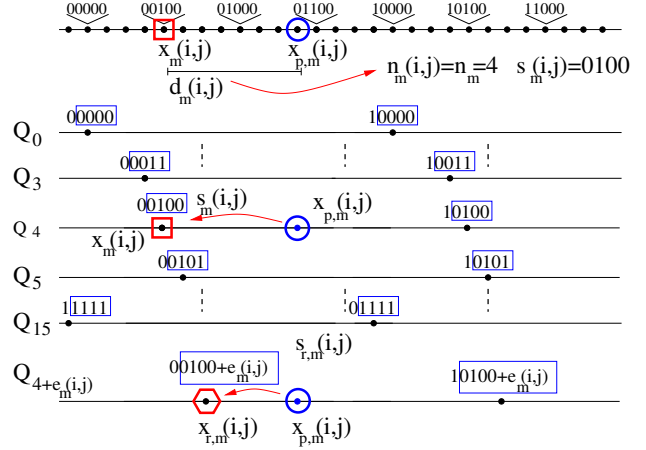
$$n_m(i, j) = \begin{cases} \lfloor \log_2(d_m(i, j)) \rfloor + 2 & \text{if } d_m(i, j) < \delta \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where  $d_m(i, j) = |x_m(i, j) - x_{p,m}(i, j)|$  and  $\delta$  is a threshold value depending on the Quantization Parameter (QP) chosen for the current block (in our setting, we have set  $\delta = \Delta/12$  where  $\Delta$  is the quantization step associated to the current QP). Then, the coding unit computes the maximum value

$$n_m = \max_{i,j=0,\dots,3} n_m(i, j) \quad (2)$$

within the current block and, in case  $n_m$  is higher than 0, it generates a block of syndromes  $s_m(i, j)$  via the following equation

$$s_m(i, j) = x_m(i, j) \& (2^{n_m} - 1) \quad (3)$$



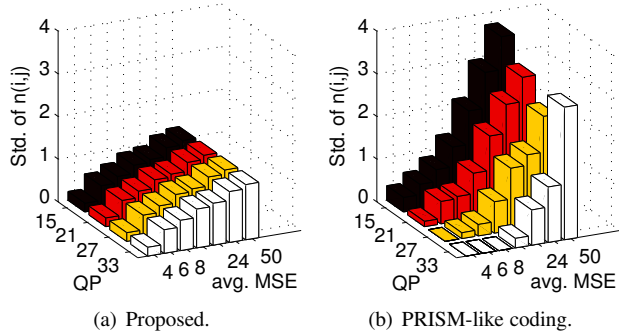
**Fig. 2.** Example of syndrome generation and decoding. Note that using the original syndrome  $s_m(i, j)$  leads to a lossless reconstruction of  $x_m(i, j)$ . The proposed scheme transforms and quantizes  $s_m(i, j)$ , and as a consequence, the final quantizer is  $Q_{4+e_m(i, j)}$  in place of  $Q_4$ .

where the symbol  $\&$  denotes a bitwise AND operator. In this way, the  $n_m$  least significant bits of each pixel are selected to generate the block  $s_m$ . The value  $s_m(i, j)$  is called “syndrome” (as in [14]), while the parameter  $n_m$  is the number of syndrome bits. The block  $s_m$  is then transformed via the  $4 \times 4$  H.264/AVC integer DCT into the block  $\mathbf{S}_m$ , which is quantized into the block  $\mathbf{S}_{q,m}$ , dequantized, and inversely-transformed into the block  $\mathbf{s}_{r,m}$ . The syndrome  $s_{r,m}(i, j)$  is a lossy version of the original syndrome  $s_m(i, j)$  such that  $s_{r,m}(i, j) = s_m(i, j) + e_m(i, j)$ , where  $e_m(i, j)$  is the distortion introduced by the quantization in the transform domain. Each lossy<sup>1</sup> syndrome  $s_{r,m}(i, j)$  identifies a different quantizer  $Q_{s_{r,m}(i, j)}$  with quantization step  $2^{n_m}$  and offset  $s_{r,m}(i, j)$  such that the reconstruction levels can be expressed as  $s_{r,m}(i, j) + k 2^{n_m}$ ,  $k \in \mathbb{Z}$ .

Given the predictor block  $\mathbf{x}_{p,m}$ , it is possible to reconstruct the coded pixel  $x_{r,m}(i, j) = x_m(i, j) + e_m(i, j)$  by quantizing  $x_{p,m}(i, j)$  using the quantizer characteristics associated to  $s_{r,m}(i, j)$  (an example is shown Figure 2). Note that it is also possible to reconstruct the signal  $x_{r,m}(i, j)$  using a different predictor  $x'_{p,m}(i, j) \neq x_{p,m}(i, j)$  provided that the correlation between  $\mathbf{x}_m$  and  $\mathbf{x}'_{p,m}$  is the same or higher (i.e. replacing  $d_m(i, j)$  with the difference  $d'_m(i, j) = |x_m(i, j) - x'_{p,m}(i, j)|$  in eq. (1) brings to the same value  $n_m(i, j)$  or lower). The higher is  $n_m$ , the more robust is the stream since a larger number of possible candidate predictors can be used in a successful decoding. However, the bit rate increases as  $n_m$  increases since the energy of the syndrome signal to be coded is higher.<sup>2</sup>

<sup>1</sup>Here the term “lossy” in relation to the syndromes refers to the lossy coding scheme adopted to characterize them.

<sup>2</sup>Note that for  $n_m = 8$  we have Intra coding for uncompressed video signals sampled with 8 bits/pixel.



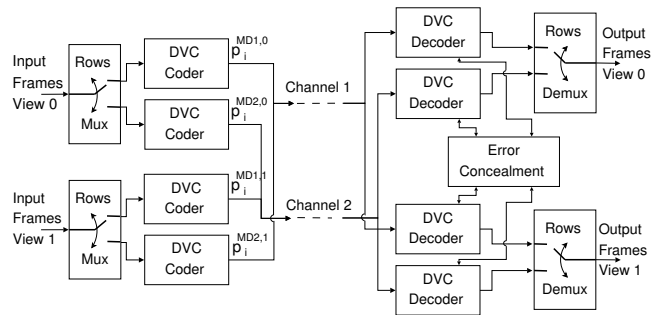
**Fig. 3.** Standard deviation (Std.) for the number of syndrome bits coded for the proposed approach and the PRISM-like approaches (sequence `foreman`). Note that each standard deviation value is reported as a function of the Mean Square Error (MSE) of the current block (quantized into 7 classes for the sake of simplicity) and of the Quantization Parameter (QP).

The decoding scheme of the proposed DVC approach is similar to the scheme adopted in the PRISM coder (see [14]), made exception for the fact that it codes syndromes with a lossy technique and operates in the pixel domain. However, this difference permits reducing the computational complexity at the decoder since the motion search is performed in the pixel domain and the inverse quantization and transform are not strictly related to syndrome decoding.

Moreover, modeling the correlation of transform coefficients is not easy since the variances of coefficient distributions significantly change depending on the spatial frequencies, and therefore, the robustness and the compression efficiency of the scheme can be seriously degraded by a wrong modelization. As an example, the PRISM-like coder in [15] achieves high compression ratios comparable to those of the H.264/AVC coding standard, but specifies different syndrome bits  $n_m(i, j)$  for each coefficient and requires modifying the context structure of the arithmetic coder since the statistics of the coded data is completely changed. In the proposed approach, the modelling of the correlation results easier since the correlation level does not vary significantly for the different pixels within a block, while the temporal correlation of transform coefficients strictly depends on the spatial frequency. As an example, in Fig. 3 we report the standard deviation of parameter  $n(i, j)$  as function of the Mean Square Error (MSE) and the adopted Quantization Parameter (QP) for the proposed approach and for the PRISM-like approach in [15].

It is possible to notice that the variance of the number of syndrome bits  $n(i, j)$  is much lower for the proposed approach, making the classification task easier and the coded stream more robust at the same time.

In the proposed coder, we simplify the characterization of the correlation between  $\mathbf{x}_m$  and  $\mathbf{x}_{p,m}$  since  $d_m(i, j)$  does not significantly vary depending on  $(i, j)$ . Moreover we are



**Fig. 4.** MDDVC scheme for stereoscopic video transmission

able to easily reuse the coding contexts and several blocks of the H.264/AVC coder [16] since the coded signal  $S_{q,m}(i, j)$  presents a statistics similar to the Intra coefficients.

The adoption of a DSC coding technique in place of a traditional residual coding scheme, like that of H.264/AVC, reduces the compression efficiency (as syndromes need more bits to be coded), but improves the robustness of the coding scheme to losses. Experimental results in Section 6 show that the DSC solution proves to be extremely effective at high packet loss rates, which correspond to the channel conditions where MDC provides a better video quality to the end-user with respect to traditional FEC techniques, e.g. the adoption of channel codes [17]. However, for sequences with high vertical correlation the traditional MDC scheme provides a better performance with respect to the DSC-based scheme since the lost information can be estimated quite well by the error concealment algorithm. The same result can be noticed at low loss rates since the introduced distortion is quite small. In these cases, the error drifting is limited, and the lower source coding distortion introduced by traditional MDC schemes permits a better estimate of the lost parts. As a consequence, it is possible to design an effective adaptive strategy that employs both the DSC unit and the traditional residual coding methods.

## 5. AN MDDVC SCHEME FOR STEREO VIDEO TRANSMISSION

The proposed MDDVC system can be effectively applied to the robust transmission of stereo video sequences since the spatial correlation existing between odd and even lines, together with the correlation between different views, can be exploited to estimate the lost signal. In our approach, we apply the MDDVC scheme separately to each view in the encoding phase, while the decoding of all the descriptions for all the views is performed jointly at the decoder. More precisely, the MDDVC coder generates for the description  $MDd$  of the view  $v$  the packets  $p_i^{MDd,v}$ , where  $i$  denotes the packet length and a possible impairment between the data available for the different views, we merge the original 4 description

into 2, concatenating  $p_i^{\text{MD},v}$  in the packets  $p_i^{\text{cd}}$  according to the following rule:

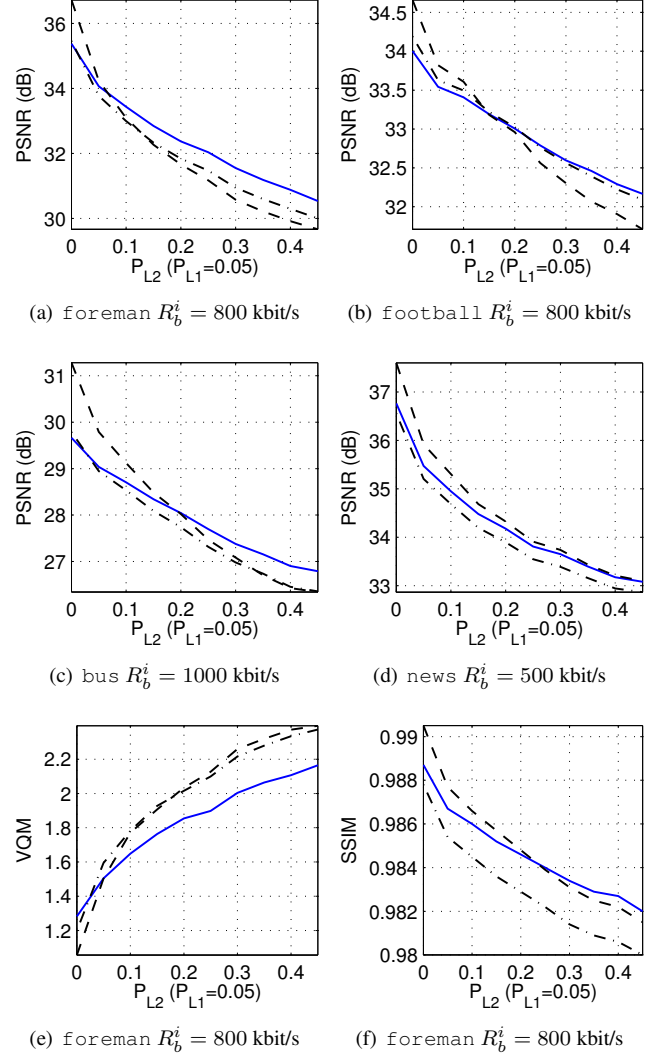
$$\begin{aligned} p_i^{\text{c}1} &\leftarrow \left[ p_i^{\text{MD}1,0}, p_i^{\text{MD}2,1} \right] \\ p_i^{\text{c}2} &\leftarrow \left[ p_i^{\text{MD}1,1}, p_i^{\text{MD}2,0} \right]. \end{aligned} \quad (4)$$

The packets  $p_i^{\text{cd}}$  are sent via two separate channels to the receiver, where they are decoded by 4 parallel decoders. In case some information has been lost, it is possible to exploit the spatial correlation existing between odd and even lines and between the different views to estimate the lost data. This estimation can be performed in different ways. In case the view  $v$  is missing either the packet  $p_i^{\text{MD}1,v}$  or the packet  $p_i^{\text{MD}2,v}$ , it is possible to estimate the missing lines interpolating the available ones within the same view. In case it is not possible to apply these concealment methods, the missing information is copied from the previous frame. The adoption of the MDDVC scheme proves to be highly effective, since the DVC decoder permits a correct decoding whenever the signal used for reference is sufficiently correlated with the signal to be decoded. The availability of different views increases the possibility of finding an adequate predictor for the blocks to be decoded and permits an accurate recovering of the transmitted information even in case of previous losses. Experimental results show that the adoption of the MDDVC scheme permits improving the quality of the reconstructed sequence at high loss rates both in terms of visual quality for each view and in terms of distortion in the disparity between each view.

## 6. EXPERIMENTAL RESULTS

The proposed DSC and hybrid DFD/DSC coders have been tested simulating the transmission of different video sequences in a scenario where packet streams are affected by losses. More precisely, each description is transmitted on an independent channel, which is simulated using a Gilbert model with burst length  $L_B = 4$ . The loss probability  $P_{L2}$  for description MD2 varies within the range  $[0.05, 0.45]$  with steps of 0.05, while the loss probability  $P_{L1}$  for description MD1 is fixed to 0.05. In this way, it is possible to grant a certain diversity between loss patterns that improves the performance of MDC schemes. In our tests we coded different CIF sequences at different bit rates  $R_b^i$ ,  $i = 1, 2$ , with GOP structure IPPP, slices of 22 macroblocks, and CABAC entropy coding. The adopted rate-distortion optimization strategy and the rate control algorithms are those defined within the JVT for the H.264/AVC coder.

The plots in Figure 5 report the values of the PSNR, VQM [18] and SSIM[19] metrics averaged over 10 channel realizations for a given  $P_{L2}$ . In each figure we compare the proposed MDDVC algorithm (referenced as DSC-based) with the traditional MDC scheme based on DFD (referenced



**Fig. 5.** Experimental results for different MDC algorithms with different residual coding units. The graphs report the average values of PSNR, VQM and SSIM metrics for the reconstructed sequences vs. the packet loss probability  $P_{L2}$  for description MD2 (the loss probability  $P_{L1}$  for description MD1 is set to 0.05). The solid lines denote the data for the DSC-based coder, the dashed line those for traditional DFD-based coder, and the dash-dotted line those for the PRISM-like coder.

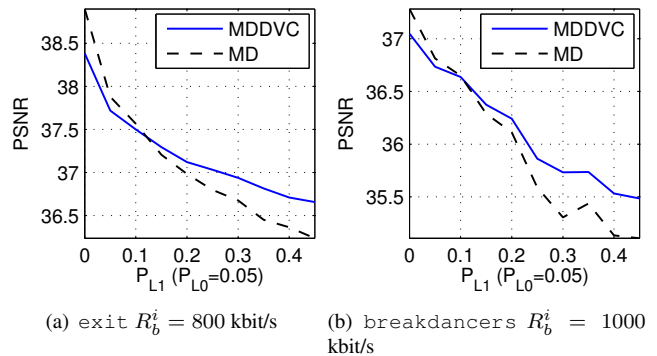
as DFD-based) and the MDDVC algorithm obtained replacing the syndrome generation unit in the scheme of Fig. 1(a) with the PRISM-like DSC coding scheme in [15] (referenced as PRISM-like). It is possible to notice that for most of the sequences the DSC-based scheme improves the quality of the reconstructed sequences whenever the loss probability becomes significant. As an example, Fig. 5(a) shows that for  $P_{L2} > 0.1$  the PSNR value for the DFD-based scheme

is always lower with respect to the proposed coder (the difference is approximately 1 dB for  $P_{L2} = 0.4$ ). The same improvement can be noticed using the other quality metrics VQM and SSIM (see Fig. 5(e) and 5(f)). Note that the performance of the PRISM-like coder is either the same or worse than the performance of the DFD-based approach. This fact is partly due to the adopted rate-distortion and rate-control algorithms, whose parameters are optimized and tuned for the H.264/AVC coder, and partly due to the difficulties in estimating and characterizing the correlation of transform coefficients (the DSC scheme we are considering do not have a feedback channel).

The same behavior can be noticed for other sequences too (see Fig. 5(b) and 5(c)), despite the crossing point between the DFD and the DSC plots depends on the characteristics of the video sequences. As a matter of fact, for the sequence *bus*, which presents a higher vertical correlation, the DSC-based approach becomes competitive for a higher  $P_{L2}$  value with respect to the sequence *foreman*. This fact is utterly evidenced by the results for the sequence *news* (see Fig. 5(d)), where the high vertical correlation and the low amount of motion allow the error concealment to perform quite well, and therefore, the PSNR values of the traditional DFD scheme are always higher or equal to those of the DSC scheme.

In the end, we tested the proposed MDDVC scheme with the transmission of stereoscopic sequences where the views taken by two cameras are independently coded and jointly decoded as described in Section 5. The captured video sequences were coded with varying target bit rate at 30 frame/s and GOP structure IPPP of 15 frames. We considered views 0 and 1 for multiview video sequences *breakdancers* and *ballet* from [20] and the sequence *exit* from [21]. The visual quality was evaluated using both traditional 2D metrics for coded video signals (like PSNR, SSIM, and VQM) and enhanced 3D visual metrics for stereoscopic videos. In this case, we adopted the SSIM Dd11 metric discussed in [22], which computes the SSIM metric on each view and combines its local value with a distortion measure evaluated on the disparity signal. In our approach, the disparity estimation was performed using a more traditional block matching algorithm in place of the belief propagation approach to reduce the overall computational complexity required to calculate the SSIM Dd11 value.

Figure 6 reports the average PSNR values for different sequences showing that for  $P_{L2} > 0.1$  the MDDVC permits obtaining a higher PSNR value with respect to the traditional MDC (up to 0.5 dB). In Figure 7 we report the average values of different metrics for the sequence *ballet*. It is possible to notice that the proposed approach improves the quality of the reconstructed sequence according to the other metrics as well. The value of SSIM metric for the proposed scheme is always higher or equal to the value for the traditional MDC (see Fig. 7(c)). Moreover, the results related to SSIM Dd11

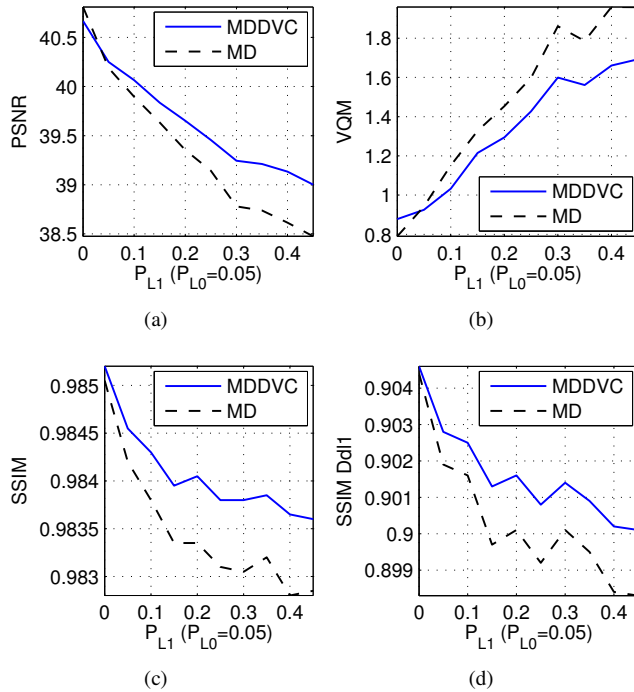


**Fig. 6.** Comparison of MDC and the proposed MDDVC scheme. The graphs report the PSNR values averaged over time and point-of-view of the reconstructed sequences vs. the packet loss probability  $P_{L2}$  for description MD2 (the loss probability  $P_{L1}$  for description MD1 is set to 0.05).

(see Fig. 7(d)) show that the MDDVC scheme is able to mitigate the distortion in the disparity information permitting a better 3D video experience to the end user. In this case, the SSIM Dd11 value related to the proposed MDDVC is always the highest.

## 7. CONCLUSION

The paper proposed an MDDVC scheme that generates two descriptions using a polyphase subsampling of odd and even rows and codes them using a DSC coder based on lossy syndromes. The proposed scheme is applied to the transmission of stereoscopic sequences over lossy channels exploiting the intrinsic robustness of the coding scheme. Experimental results show that for single view sequences the MDDVC approach outperforms its traditional MDC counterpart (up to 1 dB in PSNR) and other DSC solutions at high loss rates and whenever the vertical correlation of the input sequence is not too high. In the case of stereo video signals, the proposed solution preserves better the disparity information between the different views since its is able to mitigate the error drifting that characterizes predictive coding. Future work will be devoted to extend the designed approach to more effective MDC schemes and to a scalable coding of the input sequence in terms of temporal resolution and visual quality.



**Fig. 7.** Comparison of MDC and the proposed MDDVC scheme for the sequence *ballet*. The graphs report the average PSNR (a), VQM (b), SSIM (c), and SSIM Dd11 (d) values vs. the packet loss probability  $P_{L2}$  for description MD2 (the loss probability  $P_{L1}$  for description MD1 is set to 0.05).

## 8. REFERENCES

- [1] A. Norkin, A. Aksay, C. Bilen, G. B. Akar, A. Gotchev, and J. Astola, "Schemes for Multiple Description Coding of Stereoscopic 3D," *Lecture Notes in Computer Science*, vol. 4105/2006, pp. 730 – 737, Sept. 2006.
- [2] V. K. Goyal, "Multiple Description Coding: Compression Meets The Network," *IEEE Signal Processing Mag.*, vol. 8, no. 5, pp. 74–93, Sept. 2001.
- [3] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed Video Coding," *Proc. of the IEEE, Special Issue on Video Coding and Delivery*, vol. 93, no. 1, pp. 71–83, Jan. 2005, Invited Paper.
- [4] C. Yeo and K. Ramchandran, "Robust distributed multi-view video compression for wireless camera networks," in *Proc. of VCIP 2007*, vol. 6508, San Jose, CA, USA, Jan. 28 – Feb. 1, 2007, pp. 65 080P1 – 65 080P9.
- [5] J. Wang, X. Wu, S. Yu, and J. Sun, "Multiple Descriptions in the Wyner-Ziv Setting," in *Proc. of 2006 IEEE International Symposium on Information Theory (ISIT 2006)*, Seattle, WA, USA, June 9 – 14, 2006, pp. 1584 – 1588.
- [6] H. A. Karim, C. T. E. R. Hewage, S. Worrall, and A. M. Kondo, "Scalable Multiple Description Video Coding for Stereoscopic 3D," *IEEE Trans. Consumer Electron.*, vol. 54, no. 2, pp. 745 – 752, May 2008.
- [7] A. B. B. Adikari, W. A. C. Fernando, W. A. R. J. Weerakkody, A. Kondo, J. L. Martínez, and P. Cuenca, "DVC Based Stereoscopic Video Transmission in a Mobile Communication System," in *Proc. of the 2008 IEEE International Conference on Future Multimedia Networks (FMN 2008) (co-located with NGMAST2008)*, Cardiff, Wales, GB, Sept.17 – 18, 2008, pp. 439 – 443.
- [8] O. Crave, C. Guillemot, B. Pesquet-Popescu, and C. Tillier, "Robust Video Transmission Based on Distributed Multiple Description Coding," *EURASIP Journal on Wireless Communications and Networks: Special issue on Multimedia over Wireless Network*, no. 1, Jan. 2008.
- [9] M. Wu, A. Vetro, and C. W. Chen, "Multiple Description Image Coding with Distributed Source Coding and Side Information," in *Proceedings of SPIE: Multimedia Systems and Applications VII*, vol. 5600, Philadelphia, PA, USA, Oct. 25 – 27, 2004, pp. 120 – 127.
- [10] Y. Fan, J. Wang, J. Sun, P. Wang, and S. Yu, "A Novel Multiple Description Video Codec Based on Slepian-Wolf Coding," in *Proc. of the Data Compression Conference (DCC 2008)*, Snowbird, UT, USA, Mar. 25 – 27, 2003, p. 515.
- [11] A. Wang, Y. Zhao, and H. Bai, "Robust multiple description distributed video coding using optimized zero-padding," *Science in China Series F: Information Sciences*, vol. 52, no. 2, pp. 206 – 214, Feb. 2009.
- [12] C. Brites, J. Ascenso, and F. Pereira, "Modeling Correlation Noise Statistics at Decoder for Pixel Based Wyner-Ziv Video Coding," in *Proc. of 2006 Picture Coding Symposium (PCS 2006)*, Beijing, China, Apr. 24 – 26, 2006.
- [13] A. Aaron, R. Zhang, and B. Girod, "Wyner-ziv coding for motion video," in *Proceedings of Asilomar Conference on Signals, Systems and Computers 2002*, vol. 1, Pacific Grove, CA, USA, Nov. 2002, pp. 240 – 244.
- [14] R. Puri and K. Ramchandran, "Prism: A new robust video coding architecture based on distributed compression principles," in *Proc. of the 40<sup>th</sup> Allerton Conference on Communication, Control and Computing*, Allerton, IL, USA, Oct. 2002, pp. 402–408.
- [15] S. Milani, J. Wang, and K. Ramchandran, "Achieving H.264-like compression efficiency with distributed video coding," in *Proc. of VCIP 2007*, San Jose, CA, USA, Jan. 28 – Feb. 1, 2007, pp. 65 082Z1 – 65 082Z12.

- [16] T. Wiegand, "Version 3 of H.264/AVC," in *Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), 12<sup>th</sup> Meeting*, Redmond, WA, USA, July 17 – 23, 2004.
- [17] S. Milani, G. Calvagno, R. Bernardini, and P. Zontone, "Cross-Layer Joint Optimization of FEC Channel Codes and Multiple Description Coding for Video Delivery over IEEE 802.11e Links," in *Proc. of the 2008 IEEE International Conference on Future Multimedia Networks (FMN 2008) (co-located with NGMAST2008)*, Cardiff, Wales, GB, Sept.17 – 18 2008, pp. 472 – 478.
- [18] F. Xiao. (2000, Winter) DCT-based Video Quality Evaluation. MSU Graphics and Media Lab (Video Group). [Online]. Available: [http://compression.ru/video/quality\\_measure/vqm.pdf](http://compression.ru/video/quality_measure/vqm.pdf)
- [19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600 – 612, Apr. 2004.
- [20] 3DTV NoE Web site : Public material. [Online]. Available: <http://3dtv-research.org/publicSwLibrary.php>
- [21] MERL Multi view video sequences. [Online]. Available: <ftp://ftp.merl.com/pub/avetro/mvc-testseq>
- [22] A. Benoit, P. L. Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP Journal on Image and Video Processing, Special issue on "3D image and video processing"*, vol. 2008, Oct. 2008, ID 659024.