Multiple Description Distributed Video Coding Using Redundant Slices and Lossy Syndromes

Simone Milani and Giancarlo Calvagno

Abstract—During the last years, video coding designers have proposed robust coding approaches that combine Multiple Description Coding (MDC) schemes with Distributed Video Coding (DVC) principles. In this way, it is possible to obtain a better error resilience since the distortion drifting through the sequence is significantly mitigated by the DVC coding unit. The paper presents a Multiple Description Distributed Video Coder (MDDVC) that codes the input video signal generating a set of “lossy” syndromes for each pixel block and creates different descriptions multiplexing primary and redundant video packets. Experimental results show that at high loss probabilities the proposed solution improves the results of the original MDC approach.

Index Terms—Multiple Description Coding, redundant slices, Distributed Video Coding, H.264/AVC, robust video transmission.

I. INTRODUCTION

The provisioning of digital video services over wireless networks requires the implementation of appropriate strategies to deal with the presence of delays, packet losses and bandwidth limitations. Among the robust video coding architectures proposed in technical literature, Multiple Description Coding (MDC) schemes [1] have proved to be significantly more efficient with respect to traditional single description coding (SDC) architectures in contrasting the visual distortion caused by time-varying channel conditions. The possibility of characterizing a single video source via multiple correlated data streams (descriptions) permits reconstructing the transmitted video sequence corrupted by packet losses since the missing information is estimated exploiting the correlation existing between the different streams.

During the last years, video coding designers have proposed alternative solutions that rely on the information-theoretic concept of Distributed Source Coding (named Distributed Video Coding or DVC schemes) [2]. DVC approaches permit an independent coding of the data at the transmitter, and a robust predictive decoding at the receiver, which can be performed correctly provided that at least a sufficiently-correlated predictor is available at the decoder [3].

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 [4]–[7]. This paper proposes a novel MDDVC scheme that creates two descriptions including redundant slices in the video stream and adopts a hybrid DVC unit that processes the video signal both in the spatial and in the temporal domain. More precisely, the set of syndromes is generated in the pixel domain and compressed in the transform domain. In this way, the distributed source coding operations are separated from signal compression. Processing the signal in the pixel domain permits handling the temporal and spatial correlation more efficiently, while transform coding permits reducing the size of the coded bit stream. Experimental results will show that the proposed approach permits improving the PSNR value of the reconstructed sequence up to 1 dB with respect to the original MDC scheme.

In the following, Section II provides an overview of the implemented codec, while Section III describes the Distributed Source Coding (DSC) scheme that has been adopted to characterize the prediction residual signal. Experimental results in Section IV show how the proposed MDDVC approach improves the quality of the coded sequences at high loss rates. Conclusions are drawn in Section V.

II. THE PROPOSED MDC ARCHITECTURE

The MDDVC schemes proposed in literature have been derived replacing the traditional temporally-predictive video coder of traditional MDC approaches with a Wyner-Ziv video coding scheme (see [4] as an example). Similarly, the proposed MDDVC solution has been designed starting from the MDC scheme presented in [8], which has been developed using the concepts of primary and redundant slices defined within the H.264 standard [9]. Primary slices are used to code the primary picture and are associated to a normative decoding procedure. Redundant slices present an alternative representation of the image which is employed whenever the primary signal can not be correctly decoded. In our approach (like in [8]), redundant slices are characterized by the adoption of a higher Quantization Parameter (QP) with respect to the primary ones.

Figures 1 and 2 show the block diagrams of the employed encoder and decoder, which have been developed using the building blocks of the H.264/AVC coder. The adopted scheme is derived from the scheme of the codec in [8] made exception for the fact that the traditional residual coding unit is replaced by a syndrome coding unit (see Fig. 1), which is based on the principle of Wyner-Ziv coding. Note that in the proposed MDDVC approach, side information is estimated by the encoder via motion-search, and therefore, the computational
complexity is close to that of the original MDC scheme (see Section IV). Nevertheless, the adopted DVC paradigm permits mitigating the channel distortion produced by packet losses on the reconstructed video sequence.

The input video signal is partitioned into macroblocks (MBs) of 16 × 16 pixels, which are temporally predicted from the reference pictures stored in the frame buffer. Then, each 4 × 4 block within the input MB is coded using a Distributed Source Coding paradigm (described in Section III) permitting the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Fig. 1. Block diagram of the proposed encoder.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

The input video signal is partitioned into macroblocks (MBs) of 16 × 16 pixels, which are temporally predicted from the reference pictures stored in the frame buffer. Then, each 4 × 4 block within the input MB is coded using a Distributed Source Coding paradigm (described in Section III) permitting the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.

Section IV). Nevertheless, the adopted DVC paradigm permits the decoder to reconstruct the video sequence at the quality level related to a given quantization parameter $Q_{P_H}$. As a result, the MDDVC coder generates the primary RTP packet $p_{k,H}^H$, which codes the $k$th rows of macroblocks (one slice) within the $H$th frame. Then, the proposed approach generates a redundant packet $p_{k,H}^L$, which codes the input MB with a quantization parameter $Q_{P_L} > Q_{P_H}$ and permits reconstructing the transmitted sequences at a lower quality level. Note that the predictor block is the same since the $Q_{P}$ of the reconstructed video sequence.
provided that it has been correctly received. The syndrome

\[ s(i,j) = d(i,j) \chi(i,j) \]

Fig. 3. Example of syndrome generation and decoding. Note that using the
approaches based on Turbo/LDPC codes, which

\[ x_{n}(i,j) = \log p (x_{n}(i,j)) \]

Fig. 4. Probability of blocking the error drifting for the sequence foreman
at different quantization settings. The blocking probability is reported as a
function of \(QP_{H}\) and \(\delta QP_{L} = QP_{L} - QP_{H}\).

The syndrome generation and reconstruction processes for the pixel \(x(i,j)\). According to the distance \(|d(i,j)| = 7\) between \(x(i,j)\) and \(x_{p}(i,j)\), the syndrome \(s(i,j) = 0100\) (associated to \(Q_{4}\) is found from

\[ n = \lceil \log_{2} |d(i,j)| \rceil + 2 \]

respectively.

\[ P(|d'| < 2^{n-1}) = P(|d| + |e_{l}| < 2^{n-1}) \approx \text{erf} \left( \frac{2^{n-1}}{(\sigma_{d} + \sigma_{e})\sqrt{2}} \right) \]

The variance of \(e_{l}(i,j)\) can be computed in the transform

\[ \sigma_{e}^{2} = \frac{\Delta_{H}^{2}}{12} \left[ 1 - \left( \frac{\Delta_{H}}{\Delta_{L}} \right)^{2} \right] \]

IV. EXPERIMENTAL RESULTS

The proposed DSC coder has been tested simulating the transmission of different video sequences in a scenario where the two descriptions are transmitted over two independent lossy channels \(C_{1}\) (associated to MD1) and \(C_{2}\) (associated to MD2) simulated via a Gilbert two-state model with burst length \(L_{B} = 4\). The loss probabilities \(P_{L1}\) and \(P_{L2} = P_{L1}\) of \(C_{1}\) and \(C_{2}\) vary within the range \([0.05, 0.25]\) with steps of 0.05. In our tests we coded different CIF sequences at different

More precisely, the distortion \(e_{l}(i,j)\) in the transform domain can be

modelled by a uniform random variable as described in [8]. After the inverse

transformation, the quantization error in the pixel domain is shaped by the

coefficients of the inverse transform, and as a matter of fact, we approximate the

statistical model of the quantization error in the pixel domain with a

Gaussian variable.
CABAC entropy coding. The adopted rate-distortion optimization strategy is the one defined within the JVT for the H.264/AVC coder, and the relation between the quantization parameters is \( QP_L = QP_H + 6 \), which permits obtaining a rate increment around 30% with respect to a single description approach.

The plots in Figure 5 report the PSNR values averaged over 10 realizations for each channel \( C_d \) given its loss probability \( P_{Ld}, d = 1, 2 \). In each plot we compare the proposed MDDVC algorithm with the traditional MDC scheme (adopted in [8]). It is possible to notice that for most of the sequences the MDDVC scheme improves the quality of the reconstructed sequences whenever the loss probability becomes significant. As an example, Fig. 5(a) shows that for \( P_{Ld} > 0.1, d = 1, 2 \), the PSNR value for the MDC approach is always lower with respect to the proposed coder (the difference is approximately 1 dB for \( P_{Ld} = 0.2 \)). Note that at low loss percentages the performance of the MDDVC coder is either the same or worse than the performance of the MDC approach. This fact is partly due to the adopted optimization algorithms, whose parameters are optimized and tuned for the H.264/AVC coder, and partly due to the loss in compression gain which is proper to the DVC approach. However, MDC approaches proves to be competitive with other SDC technique at high loss rates, and therefore, the improvement offered by the MDDVC approach motivates its adoption.

The same behavior can be noticed for other sequences too (see Fig. 5(b) and 5(c)), even if the crossing point between the MDC and the MDDVC plots depends on the characteristics of the video sequences. As a matter of fact, for the sequence news (see Fig. 5(d)) the high correlation and the low amount of motion allow the error concealment procedure to perform quite well, and as a matter of fact, the PSNR values of the traditional MDC scheme are always quite close to those of the MDDVC scheme. Therefore, an adaptive algorithm that is able to switch from the MDC setting to the MDDVC setting depending on the characteristics of the video sequence and on the estimated channel loss probability proves to be an interesting topic for future investigations.

As for the complexity of encoding and decoding, Table I reports the relative increment of computational load with respect to the traditional MDC scheme in [8] (measured via the total encoding and decoding times). No significant increases can be noticed since the adoption of motion search at the encoder equalizes the amount of required calculation for the MDDVC scheme with respect to its MDC predecessor. Despite this, the proposed solution permits improving the robustness of the video transmission at high loss rates thanks to its effectiveness in blocking the distortion drifting.

**V. Conclusion**

The paper proposed an MDDVC scheme that employs a DVC coder based on lossy syndromes and generates two descriptions multiplexing primary and redundant packets defined in the H.264/AVC standard. The proposed scheme outperforms its traditional MDC counterpart (up to 1 dB in PSNR) at high loss rates thanks to the effectiveness of DVC in blocking the distortion drifting through the sequence. Future work will be devoted to extend the designed approach switching adaptively between the MDC and the MDDVC approaches according to the channels conditions and the characteristics of the sequence.

**References**


