Practical Distributed Video Coding in Packet Lossy Channels

Optical Engineering

SPIEDigitalLibrary.org/oe

# Practical distributed video coding in packet lossy channels

Linbo Qing
Enrico Masala
Xiaohai He

SPIE

# Practical distributed video coding in packet lossy channels

**Linbo Qing**
Sichuan University
College of Electronics and Information Engineering
No. 24 South Section 1, Yihuan Road
Chengdu 610065, China
E-mail: qing_lb@scu.edu.cn

**Enrico Masala**
Politecnico di Torino
Control and Computer Engineering Department
Corso Duca degli Abruzzi
24 - 10129 Torino, Italy

**Xiaohai He**
Sichuan University
College of Electronics and Information Engineering
No. 24 South Section 1, Yihuan Road
Chengdu 610065, China

**Abstract.** Improving error resilience of video communications over packet lossy channels is an important and tough task. We present a framework to optimize the quality of video communications based on distributed video coding (DVC) in practical packet lossy network scenarios. The peculiar characteristics of DVC indeed require a number of adaptations to take full advantage of its intrinsic robustness when dealing with data losses of typical real packet networks. This work proposes a new packetization scheme, an investigation of the best error-correcting codes to use in a noisy environment, a practical rate-allocation mechanism, which minimizes decoder feedback, and an improved side-information generation and reconstruction function. Performance comparisons are presented with respect to a conventional packet video communication using H.264/advanced video coding (AVC). Although currently the H.264/AVC rate-distortion performance in case of no loss is better than state-of-the-art DVC schemes, under practical packet lossy conditions, the proposed techniques provide better performance with respect to an H.264/AVC-based system, especially at high packet loss rates. Thus the error resilience of the proposed DVC scheme is superior to the one provided by H.264/AVC, especially in the case of transmission over packet lossy networks.  [DOI: 10.1117/1.OE.52.7.071506]

Subject terms: distributed video coding; Wyner-Ziv; packetization; packet lossy networks; side information.

Paper 121542SS received Oct. 23, 2012; revised manuscript received Feb. 8, 2013; accepted for publication Mar. 5, 2013; published online Mar. 25, 2013.

## 1 Introduction

The last decade has seen an exponential increase in the diffusion of video communication applications. Moreover, such applications currently can run on heterogeneous devices with very different capabilities, ranging from mobile devices with relatively small displays and limited power and computational resources to powerful personal computers with high-definition displays.[1] This advancement has been made possible by the intense research and development efforts in the field by academy and industry. A key contribution has been the increasing compression and resiliency performance offered by the new hybrid video coding standards such as H.264/ advanced video coding (AVC),[2] which allowed us to optimize the video presentation quality to end users while saving computational resources and bitrate. Recently, the emerging high-efficiency video coding (HEVC) standard[3] showed that it is possible to further improve coding efficiency, especially for the case of high-resolution video.[4,5]

Moreover, the usefulness of video applications in the mobile environment pushed researchers to investigate solutions to improve the error resiliency of video communications over wireless packet networks. This scenario has also been addressed in the design phase of new standards such as the H.264/AVC, which, differently from previous standards, incorporates error resiliency tools specifically designed for packet network technology.[6] Also, the introduction of a clear separation between the video coding layer and the network adaptation layer facilitates the design of adaptation schemes for peculiar scenarios such as the case of heterogeneous networks.[7]

### 1.1 *Distributed Video Coding: A New Paradigm*

While many research efforts have been devoted to the optimization of the traditional hybrid video coding scheme, for resource-limited environments, a number of researchers started to focus on a completely different paradigm, i.e., distributed video coding (DVC),[8] which promises to shift the computational complexity from encoder to decoder, with obvious benefits when the encoder is severely constrained in terms of energy and complexity. The key advantage of DVC is the possibility to perform encoding for each frame independently without negatively affecting quality performance. For a given bitrate, much lower encoding complexity is required with respect to the hybrid video coding scheme since the motion estimation step is not needed. However, complexity is shifted to the decoder, which needs to perform additional operations to reconstruct the original data by taking into account the correlation between several frames. Another outstanding advantage of DVC is the absence of a prediction loop in the encoder, which yields low encoding complexity and improved error resilience. Although the advantage of low encoding complexity of DVC is being eroded by new efficient video compression techniques for traditional video coding, the intrinsic error resilience of DVC turns out

0091-3286/2013/$25.00 © 2013 SPIE

to be a very attractive feature to implement robust video transmissions.

In practice, there are two main DVC architectures that have been proposed so far.[9,10] Unfortunately, DVC still provides scarce rate distortion (RD) coding performance compared with the theoretical upper bound; hence the majority of the works dealing with DVC focused on addressing this issue.[9] On the other hand, some researchers have also realized the potential of the intrinsic error resilience of DVC and introduced DVC into traditional video transmission for better error resilience. Unfortunately, few works focused on the performance of standalone DVC systems in case of transmission of DVC data over packet lossy channels, which is an important issue to address in order to move toward the deployment of effective DVC-based communication applications. Also when transmission issues are considered, existing works often analyze the transmission performance in theoretical scenarios, which can be far from actual network conditions. The detailed review of related work will be discussed in Sec. 2.

### 1.2 *Goals and Novelties of this Work*

Current DVC architectures typically rely on error correcting codes to implement the encoding algorithm. Therefore, those architectures are deemed inherently robust to transmission errors or losses. While in part this is true, however, it must be noted that error-correcting codes in DVC schemes are often optimized to recover the errors introduced by the "virtual channel," a fundamental element of the DVC block diagram, as it will be discussed in Sec. 3, and not by actual transmission using current packet network technologies. Indeed, the peculiar behavior of packet networks in terms of data loss patterns can strongly affect the performance of the DVC scheme, as it will be discussed in this work. Therefore, techniques aimed at improving the error resilience of DVC schemes in practical scenarios are strongly needed.

This paper aims at proposing new techniques to solve the practical issues encountered while transmitting DVC data on a real packet network. In more details, a number of new, practical enhancements to DVC schemes are proposed in this work to allow successful DVC communications to take place on packet networks.

The novelties and main contributions of this work with respect to the existing literature, reviewed in Sec. 2, can be summarized as follows: (1) An investigation of the suitability of different types of error correcting codes used in DVC for the case of transmission over a noisy channel. (2) A practical way to arrange the compressed DVC data into packets, which has many advantages, including the minimization of the impact of packet losses on the performance and the possibility to adapt the packet content using a practical rate-allocation algorithm which also considers channel conditions. (3) A practical rate-allocation algorithm whose mechanisms consider the condition of both the virtual channel and the real transmission; the key novelty is that the algorithm can compute the rate of the transmission over error-prone channels without relying on extensive feedback to the encoder. (4) Algorithms to generate and reconstruct the best possible side information (SI) in packet lossy channel conditions are given, so that the performance of the DVC scheme is maximized. (5) A reconstruction function to deal with the case of losses exceeding the error correction capabilities provided by the received parity bits and make full use of successfully decoded bit planes.

Simulation results are provided to show the effectiveness of the proposals when the previous techniques are incorporated into a common DVC architecture, i.e., the DISCOVER one introduced in Ref. 9. However, note that all the robustness techniques proposed in this work are suitable for any of the codec variants proposed in the literature, which are based on the original DISCOVER architecture. Thus, when the research community will develop new techniques to improve the coding efficiency of the DISCOVER architecture, our proposals can be readily incorporated in a straightforward way so that the global performance of communication can take advantage of the improvements in DVC research.

The paper is organized as follows. Section 2 provides an outline of the related work. Section 3 presents the theoretical background on DVC including the reference scheme adopted in this work. Section 4 briefly reviews the DVC coding architectures aiming at understanding, at a high level, the challenge and additional work that is needed for transmission of the coded data over practical real packet networks. Section 5 investigates the error resilience of different types of channel coding schemes for DVC transmission purposes. A proposal for an efficient packetization scheme is presented in Sec. 6, followed by Secs. 7 and 8, which explain how the proposed scheme deals with rate allocation, SI generation and reconstruction function in packet lossy networks, respectively. Section 9 presents simulation results, followed by conclusions in Sec. 10.

## 2 Related Work

Techniques for improving the reliability of video transmission over error-prone channels have always received a great deal of attention by the research community. A large number of them have been proposed in the literature, and summarizing all of them is out of the scope of this paper. However, it is interesting to note that many of them focused on introducing error resilient coding tools directly into existing video encoders, eventually leading to extensions of the standards themselves, as overviewed in Ref. 11. Moreover, the introduction of video coding standards such as H.264/AVC,[2] which has gained a tremendous popularity in recent years, has further increased the availability of such resilience tools, since H.264/AVC has been the first standard to specifically address the issue of transmission over packet networks starting from the design phase. Indeed, it is the first standard to isolate the codec core (the video coding layer) from the remaining parts.[12] This allows better efficiency in utilizing different types of channels, and in particular packet networks.[13]

Thus, when the first DVC architectures have been proposed, it has been immediately clear that one of the attractive advantages of DVC over traditional hybrid video codec schemes that needed specific tools for error resilience is its intrinsic ability to deal with transmission errors. This strength led many researchers to readily employ DVC schemes to exploit such a characteristic by, e.g., supplementing a principal, traditional video communication system with Wyner-Ziv (WZ) encoded video to enhance communication robustness or use it on its own. Both approaches are reviewed in following subsections.

### 2.1 *DVC-Aided Robust Video Transmission*

A number of researchers proposed to use WZ encoded stream to "protect" the original video stream, showing that significant gains are possible with respect to the traditional approach. For instance, Ref. 14 uses this scheme to mitigate the picture quality drop suffered by traditional forward error correction based systems, resorting, in case of necessity, to the WZ stream that, in turn, uses the corrupted video as SI. A similar technique can also be used to protect base layer information in a scalable video coding scheme, as done in Ref. 15. The same approach can also be applied to protect region-of-interests (ROI) within the frames of a video sequence, as done in Ref. 16. Others propose to add small information pieces for redundancy purposes, termed coset in Ref. 17, instead of a whole WZ video stream. However, Ref. 17 assumes a maximum coset erasure probability of $10^{-3}$, which strongly limits its applicability to practical scenario where the packet loss ratio (PLR) can be up to 20%. DVC principles are also introduced into multiple description image coding for better error resilience.[18]

### 2.2 *Standalone DVC Over Error-Prone Channels*

When used as a standalone system, DVC has shown great advantages for many appealing scenarios that require low-complexity encoding, such as wireless multimedia sensor networks, low-power surveillance, mobile camera phones, etc. For these types of applications, the traditional video coding paradigm can be too demanding, whereas DVC coding and transmission is best suited, since it can provide both coding efficiency and significant error resilience. However, most of the works dealing with DVC focus on improving the encoder rate-distortion performance,[8] since state-of-art systems are still far from their theoretical performance bound. To the best of our knowledge, very few works dealt with the issues caused by transmitting DVC compressed data over unreliable packet networks. In the following we attempt a review of the works dealing with this specific aspect, discussing advantages and limitations with particular reference to the techniques proposed in this paper.

The work in Ref. 19 presents a robust wireless video multicast solution based on the PRISM architecture,[10] focusing on the introduction of scalability in the DVC scheme. The work achieves better performance than scalable H.263+, considering a maximum of 10% packet loss rate. However, two main disadvantages reduce its applicability in practical scenarios: encoding complexity is increased by partial motion search operations and rate allocation mainly relies on offline training, which reduces the local efficiency of the algorithm. Moreover, the work does not consider the errors introduced by the packet losses (the typical impairment in real packet networks) that corrupt the SI, which can have a strong influence on the system error resilience performance.

A preliminary study of the error resilience of a transform domain WZ codec (TDWZ)[20] has been accomplished in Ref. 21, where several coding cases are studied and better error resilience than H.264/AVC, in "no motion" configuration [although the authors have not clearly described this point, the performance for H.264/AVC is similar to the one shown in Ref. 9, where H.264/AVC in "no motion" configuration (I-B-I scheme with no motion vectors) is studied for the same conditions and sequences], has been shown. However, most of their results (including the comparison with H.264/AVC) are based on unpractical assumptions: (1) in some cases, they assume the additional WZ data can always be requested through a feedback channel to implement a "decoding/transmission of more data/decoding" (DTD) loop, which would introduce high latency, as pointed out in Ref. 22, but most practical systems do not allow it; (2) each additional segment of bits is packetized into one packet, introducing a large overhead due to the very small packet payload; (3) in the cases where no feedback channel is used, they assume that the rate can be perfectly estimated at the encoder. They also do not analyze or address the issue of designing techniques to improve the error resilience, but simply perform experiments with the plain version of the codec and report the results which, due to the previous limitations, show significant performance degradation [i.e., about 5 dB peak signal-to-noise ratio (PSNR) performance loss at 350 kbps for the *foreman* sequence where only the key frames are corrupted].

A more comprehensive study of the error resilience characteristics of two of the main DVC architectures, namely the Stanford architecture (DISCOVER)[9] and PRISM architecture,[10] has been presented in Ref. 23. Their results have shown better error resilience of both DISCOVER and PRISM with respect to H.264/AVC (i.e., 4 dB PSNR loss for DISCOVER and 9 dB loss for H.264/AVC at 300 kbps), besides indicating that DISCOVER outperforms PRISM in RD terms. Similar to the work in Refs. 21, 23 also uses a feedback channel to implement the DTD loop to deal with loss of WZ data, which is unpractical due to latency, as previously explained. The authors also state that the comparison between DISCOVER and H.264/AVC is somehow not fair, since DISCOVER can ask more data through feedback channel in case of data loss while H.264/AVC cannot. In the latter case, performance should be better.

Summarizing the previous discussion, although the existing work provided some insight about the error resilience capability of DVC, many practical issues have been ignored. None of the works has taken into consideration the channel-coding scheme used in the DVC. In order to exploit, to the maximum extent and in a practical way, the intrinsic error resilience of DVC, this work proposes a number of new algorithms and novel solutions that address the gaps and limitations in the literature as more specifically discussed in detail in Sec. 4, in particular a comprehensive analysis of different channel codes for Slepian-Wolf coding (SWC) over error-prone channels, a practical way to arrange the compressed DVC data into packets showing how this aspect is fundamental to optimize transmission over packet networks, a rate-allocation mechanism that eliminates the DTD scheme, which is not suitable for practical scenarios, how to reconstruct the best possible SI in a scenario where key frames are corrupted, and a reconstruction function to deal with the case of decoding failure for some bit planes. These contributions, which significantly improve the existing literature, will be described in detail in Secs. 5, 6, 7, and 8.

## 3 Theoretical Background

### 3.1 *Compression of Source with Decoder Side Information (Slepian-Wolf Problem)*

Let $Xn$ be the source that must be transmitted to receiver. Source entropy is $H(X)$. Suppose that there exists a source

$Yn$, available at the receiver, which is correlated with $Xn$ (Fig. 1). According to the Slepian-Wolf theorem,[24] the source $Xn$ can be reliably transmitted if the following condition is satisfied:

$$R_x \geq H(X|Y), \tag{1}$$

where $R_x$ is the transmission rate of $Xn$.

### 3.2 *Capacity of Channels with Uncoded Side Information*

Consider a scenario similar to Fig. 1, but now there exist two independent parallel channels, and the inputs to both channel are encoded (Fig. 2). According to Shannon's coding theorem, the source $Xn$ can be reliably transmitted if the following condition is satisfied:

$$H(X) \leq C_1 + C_2, \tag{2}$$

where $C_1$ and $C_2$ are the channel capacity of channels 1 and 2, respectively.

Suppose $Xn$ is directly transmitted in channel 2 without coding by encoder 2, the source can be transmitted if, according to Ref. 25,

$$H(X|Z) \leq C_1, \tag{3}$$

where $Zn$ is the output of channel 2. If block channel coding is used, the block length of source $Xn$ is $k$, code word length is $n$, the source is transmissible if the following condition is satisfied asymptotically:

$$n/k \geq [H(X|Z)]/C_1. \tag{4}$$

## 4 Architectures for DVC

### 4.1 *DVC Coding Schemes*

The DISCOVER architecture described in Ref. 9 is one of the most popular and successful DVC schemes, and it is currently regarded as the state-of-art DVC coding solution. This work follows the DISCOVER architecture, which is depicted in Fig. 3. Video source is first split into interleaved key frames and WZ frames. Key frames are encoded independently, with a H.264/AVC intra encoder. WZ frames are discrete cosine transformed (DCT) and quantized. Coefficients are organized into bit planes and encoded by channel encoder. The channel code can be a turbo code (TC),[26] a low-density parity check code (LDPC),[27] or an LDPC accumulated code (LDPCA).[28] Then the parity bits (for TC and LDPC) or syndrome bits (for LDPCA) are stored in a buffer and transmitted to the decoder partially, under the control of the rate allocation (RA) module. The main task of the RA module is to estimate the minimum number of parity bits or syndrome bits to be sent for each bit plane and each band. RA can be accomplished with[29] or without[30] decoder's feedback, either in decoder or encoder.

At the decoder side, the key frames are first decoded independently with a H.264/AVC intra decoder. For each WZ frame, first its SI is generated from adjacent key frames. Then SI is DCT, quantized, and subdivided into bit planes. With the help of the virtual channel model between SI and WZ frame, the soft input information of each bit plane is calculated and used as input for the channel decoder module, where iterative decoding is executed. Once all the bit planes are decoded, all the coefficient bands are reconstructed, with the help of SI. Finally, the decoded WZ frames are generated by means of the inverse discrete cosine transform (IDCT).

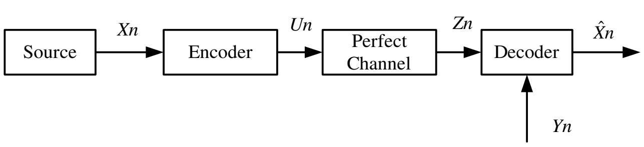



**Fig. 1** Compression with decoder side information.

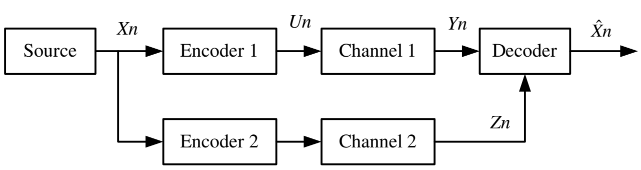



**Fig. 2** Compression with transmission of side information.

### 4.2 *Practical DVC Over Packet Lossy Channel*

Although the DISCOVER codec can achieve significant performance when forward and backward channels are noiseless, when considering practical transmission of a DISCOVER stream over packet lossy channel several points should be considered and optimized.

First, the parity/syndrome bits for all the bit planes must be divided into packets before transmission. In the DISCOVER codec,[9] the minimum number of syndrome/parity bits to be transmitted to decoder is first estimated on the basis of the minimal theoretical rate $R_{\min}$. If the WZ decoder fails, more syndrome/parity bits are requested from the encoder using a feedback channel, and decoding is repeated. In this paper, we define this process based on a feedback channel as the DTD process. For each DTD cycle, an additional small amount of parity/syndrome bits are transmitted. This is a very small payload to be sent in a single packet. So, in practical packet lossy channels, this DTD process needs to be avoided. An RA module should be introduced to decide the amount that is needed for WZ decoding at one time, then all the data are packetized using a reasonable packet length for transmission. In addition, the actual number of transmitted bits should be higher when packet losses are possible in the transmission channel, which means that the RA module should consider the PLR for rate allocation.

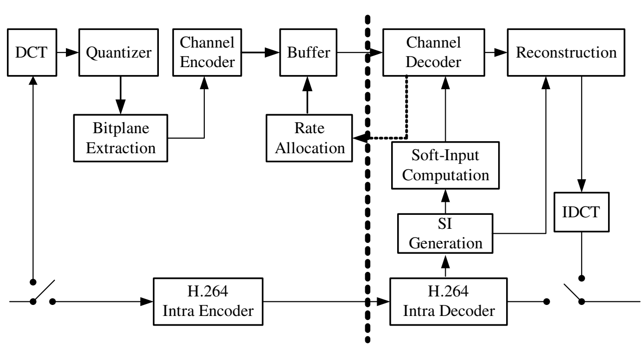



**Fig. 3** The simplified DISCOVER codec.

Second, the key frames are also affected by packet loss, and several blocks may be lost. The lost blocks can be concealed by traditional error concealment techniques based on the knowledge of the location of the missing blocks. Such information could be easily derived, for instance, from the sequence number of packets if the real-time transport protocol is used. Therefore, an error mask for key frame indicating the concealed areas can be easily constructed and should be considered for SI generation for better SI quality.

Finally, not all the bit planes can be decoded without errors depending on the amount of lost parity/syndrome bits. So the reconstruction function also needs to be modified.

The proposed practical distributed video coding scheme over packet lossy channel is shown in Fig. 4. This paper will mainly deal with the problems of rate allocation, packetization, channel decoder, SI generation, and reconstruction, whose blocks are shown in gray in Fig. 4.

## 5 Channel Coding in DVC

Roughly speaking, DVC is composed of quantization followed by Slepian-Wolf lossless source coding. Better SWC performance will lead to better performance of DVC in any circumstance. In this section we will focus on the SWC performance with transmission erasures. The source-coding problem of SWC is actually a channel coding one and near-capacity channel codes such as turbo and LDPC codes can be used to approach the Slepian-Wolf limits.[31] However, it has already been proven that LDPC can achieve better error-correction performance than TC,[32] which can result in better performance for practical DVC codecs. For LDPC-based DVC systems, there exists two schemes, i.e., the syndrome-based approach[9] and parity-based approach.[33] Since the correlation varies from frame to frame, band to band, bit plane to bit plane, the rate for perfect decoding should adaptively change. For parity-based SWC, the rate adaptability can be implemented by code puncturing.[33] For syndrome-based SWC, rate adaptability is achieved by syndrome accumulation.[9] Indeed, when considering transmission of DVC data over a perfect channel, the syndrome-based approach can outperform the parity-based approach. But when losses are considered, the situation changes. This issue is briefly discussed in the following subsections.

### 5.1 *Message Propagation Decoding Algorithms for LDPC*

An instance of an LDPC code can be specified by a parity-check matrix or a bipartite graph as shown in Fig. 5, where the circle nodes represent variable nodes and square nodes stand for the check nodes. The decoding algorithm for LDPC is based on the belief propagation algorithm, which can be solved by means of a message-passing algorithm (MPA) using bipartite graph.

The basic idea of MPA can be described as the extrinsic information being propagated between check nodes and variable nodes iteratively. Figure 6 shows the message propagation in the subbipartite of Fig. 5. The simplified MPA can be given as follows (more details can be found in Ref. 34):

1. Each variable node $v_i$ processes its input messages to calculate the output message $\vec{m}_{ij}$, which is then propagated to the neighboring check node $c_j$. $\vec{m}_{ij}$ concerns the temporal probabilities of the code bits $(z_i = b, b \in \{0, 1\})$ in current iteration, conditioned by input messages including the information of received channel sample $y_i$ (in Fig. 6, $i = 0$, $j = 2$).
2. Each check node $c_j$ processes its input messages to calculate the output message $\overleftarrow{m}_{ji}$, which is then

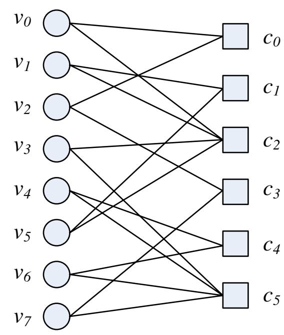



**Fig. 5** Bipartite of LDPC codes.

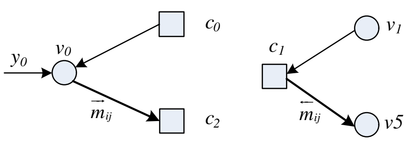



**Fig. 6** Message propagated between check nodes and variable nodes (left: message from $v_0$ to $c_2$; right: message from $c_1$ to $v_5$).

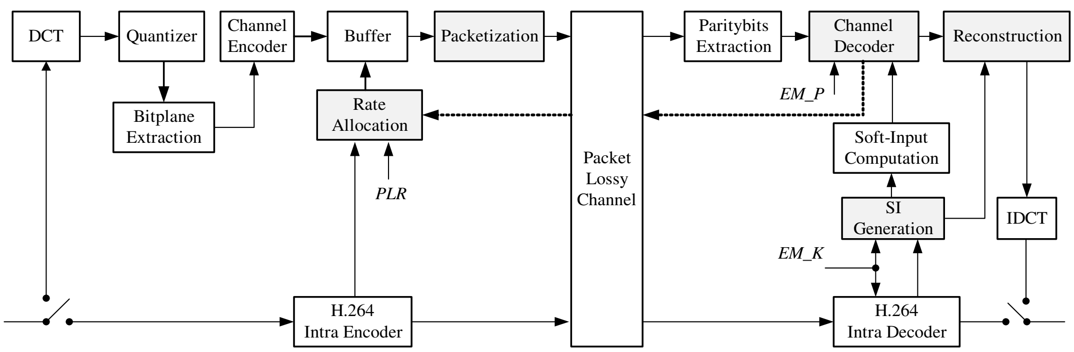



**Fig. 4** Proposed Wyner-Ziv coding and transmission scheme.

propagated to the neighboring variable node $v_i$. $\overleftarrow{m}_{ji}$ concerns the probabilities that check equation $c_j$ is satisfied (in Fig. 6, $j = 1$, $i = 5$).

3. Repeat steps 1 to 2 iteratively until stopping criterion has been met.

### 5.2 *Error Resilience of SWC Over Binary Erasure Channel*

Assuming the output bitstream of SWC is transmitted over a binary erasure channel (BEC), some bits in the code word will be erased or deleted due to poor channel condition. For the parity-based SWC, the lost bits cannot be distinguished from the ones that have been punctured, i.e., since the message propagation path remains the same, the only influence of this fact is that there is no information coming from the channel. However, as long as there are enough messages from neighboring check nodes, the lost bits can still be recovered.

For the syndrome-based SWC, the situation changes. The output message for each check node is the possibility of check equation $c_j$ to be satisfied. If the syndrome bits of current check node are lost, there is no way to calculate such possibility. In other words, all the message-propagation paths connected to this check node are deleted due to the loss of the syndrome bits. This will dramatically influence the performance of decoding. Therefore, parity-based SWC is more suitable for DVC over erasure channels than syndrome-based SWC.

### 5.3 *Experimental Results*

In order to validate the analysis in the previous subsection, experiments of SWC coding for a binary source $X$ are implemented. We assume that the correlation between $X$ and its side information $Y$ can be modeled by independent and identically distributed (i.i.d) binary symmetric channel statistics with crossover probability $p_c$. The output parity/syndrome bits are transmitted over a BEC with bit erasure rate $r_e$. The code word length is $n = 6336$, iteration number is 100, and the basic irregular parity check matrix are taken from Ref. 9. Two cases are considered for the experiments, i.e., fixed rate case and variable rate case.

1. Fixed rate setting: In this case, first the rate $R_0$ for perfect SWC decoding when there is no erasure is decided and fixed for the experiments. Then the source is SWC coded at rate $R_0$ and transmitted through BEC with different bit erasure ratio $r_e$. The decoding error ratio $r_d$ for different $r_e$ is given in Fig. 7.
2. Variable rate setting. In this case, the coding rate $R_d$ is first set as $R_0$ in case 1 and decoding is attempted. We assume the decoder can detect the decoding error rate. Similarly with original feedback-based SWC, the decoder will require more parity/syndrome bits for perfect decoding, potentially until all the parity/syndrome bits are transmitted (i.e., $R_d = 1$). But for each retransmission, the bits erasure ratio is still the same. In other words, there always exist bits erasures. Apparently, when the bit erasure rate $r_e$ exceeds a certain value, both parity-based SWC and syndrome-based SWC will fail to decode perfectly. For each crossover probability between $X$ and $Y$, the three-dimensional plots for final coding rate $R_d$, channel erasure ratio $r_e$, and decoding error ratio $r_d$ are given in Fig. 8.

As shown in Fig. 7, when there are bits erasures, both LDPC parity-based and LDPCA syndrome-based SWC

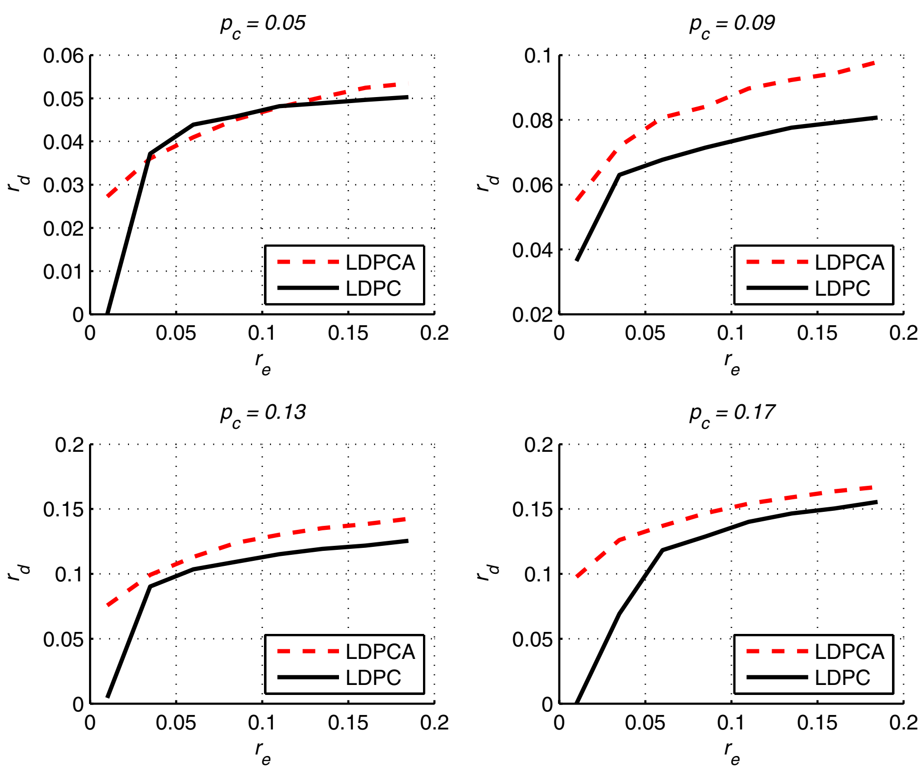



**Fig. 7** Decoding error ratio for different erasure ratios.

schemes cannot achieve perfect decoding, with the exception of parity-based SWC under some conditions (i.e., when $r_e >$ 0.01 at $p_c = 0.05$, 0.13, 0.17). But parity-based SWC has fewer errors than syndrome-based SWC. From Fig. 8 it can be seen that when retransmissions of parity bits are permitted, the syndrome-based SWC outperformed parity-based SWC in terms of coding rate $R_d$ when $r_e < 0.1$. However, when $r_e > 0.1$, the situation changes. The $R_d$ for syndrome-based SWC increases dramatically and exceeds parity-based SWC. Moreover, as $r_e$ increases, syndrome-based SWC cannot decode successfully even if all the syndromes are transmitted (i.e.,$R_d = 1$). On the other hand, the parity-based SWC can still accomplish perfect decoding, as long as enough parity bits are sent to decoder. Note that if $r_e$ exceeds a certain value, both the parity-based SWC and the syndrome-based SWC will not be able to decode the data, but the threshold value is, in general, much higher for the parity-based SWC than the syndrome-based SWC.

In conclusion, when $r_e$ is lower than certain value, the LDPCA syndrome-based SWC outperforms LDPC parity-based SWC. However, when $r_e$ exceeds a certain value, LDPCA syndrome-based SWC performs much lower than parity-based SWC, which means higher rate for the DVC system. In addition, LDPCA syndrome-based SWC cannot perfectly decode even if all the syndromes are transmitted as the $r_e$ increases, which would dramatically decrease the quality of DVC output frames. For these reasons, LDPC parity-based SWC is more suitable than LDPCA syndrome-based SWC for DVC over packet lossy channels where bit erasures are possible. Therefore, in the following sections, we only employ LDPC parity-based SWC for the DVC system.

## 6 Proposed Packetization Scheme

As analyzed in Sec. 5, LDPC parity-based SWC has better error resilience than LDPCA syndrome-based SWC. The systematic mode generation matrix $G$ is first generated and used to encode every bit plane of each coefficient. Then only the parity bits are buffered and sent on the actual channel while the original data are supposed to be sent on a virtual channel. The output of the virtual channel is the so-called side information, which is generated at the decoder side by means of the key frames. The channel decoder then uses the SI with the parity bits to recover the original data of the WZ frame. The next subsection will investigate the issue of how to packetize the required parity bits before sending them over a packet network.

### 6.1 *DVC Data Structure*

In order to present the proposed packetization scheme for DVC, the data structure should be first described. Suppose the frame size is $w \cdot h$, where $w$ and $h$ represents the width and height of the frame, respectively. Pixels are grouped into $4 \times 4$ blocks; then, as shown in Fig. 4, DCT is applied for each block and the resulting coefficients, one for each band $i$, are quantized according to a quantization parameter $Q_i$, $i \in [0, 15]$. For each band $i$ a number of bits for quantized symbol $X_i$, $i \in [0, 15]$ is decided. $Q_i$ decreases as the spatial frequency increases. For instance, Fig. 9 shows that the DC coefficient is represented using $Q_0 = 7$ bits while the highest frequency one has only $Q_{15} = 1$ bit. Then each coefficient $X_i$ is subdivided into bit planes, with the $j$'th most significant bit plane represented by $B_i^j$, $i \in [0, 15]$, $j \in [1, Q_i]$. Finally, the bits with the same band $i$ and bit plane number $j$ of all the DCT blocks are

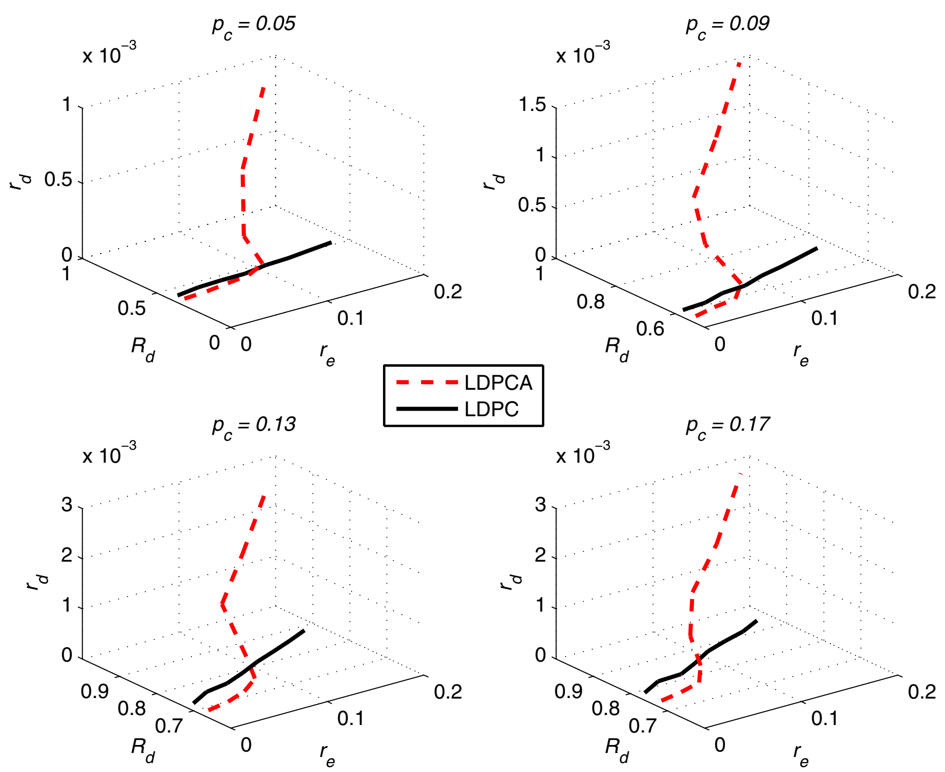



**Fig. 8** Decoding error ratio for different erasure ratios, coding rate, and cross-over probability.

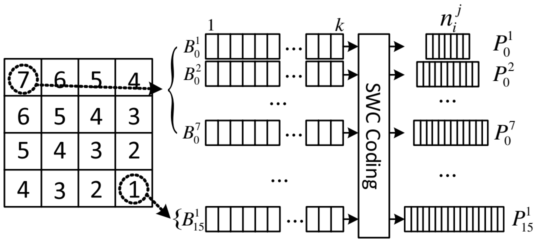



**Fig. 9** Data structure of WZC encoder.

grouped together. Each bit plane is independently SWC coded with the code word length of $k = w \cdot h/16$. The set of parity bits for $B_i^j$ for perfect decoding is represented by $P_i^j$ and the bit number is $n_i^j$.

### 6.2 *Rate Allocation Problem for DVC Over Binary Erasure Channel*

Before dealing with packetization issues, it is necessary to discuss the rate-allocation problem for DVC over binary erasure channels, since the variability of the amount of data to be sent has an impact on the packetization scheme, as explained in the next subsection. The coding and transmission system for each bit plane in Fig. 4 can be perfectly modeled by the "compression with uncoded side information" structure shown in Fig. 2, where channel 1 represents the real transmission channel and channel 2 is the virtual channel. For each bit plane, the correlation between source and side information varies, and the minimum rate for successful decoding varies consequently, so the required channel 2 capacity varies from bit plane to bit plane. Suppose channel 1 is a packet lossy channel, its status can be described by a number, the PLR $p$, yielding to a channel capacity $C1 = (1 - p)$.

Let $X_i^k$ represent the source sample vector of quantized symbols with source length $k$ for one specific coefficient band $i$, $i \in [0, 15]$, $B_i^j(X)$, $j = 1, 2, 3, \ldots$ denote the $j$'th most significant bit plane of $X_i^k$, $B_i^{\{j-1\}}(X)$ denote the symbol from earlier decoded bit planes. Then, according to Samuel's results in Ref. 35, the rate required for lossless compression of the $j$'th bit plane is given by

$$H(B_i^j(X)|B_i^{\{j-1\}}(X), Y_i^k), \tag{5}$$

where $Y_i^k$ is the SI of $X_i^k$.

As described earlier in this section, the information bits of each bit plane are coded using systematic LDPC coding, and only the parity bits are punctured and transmitted to the decoder. Given an $(n, k)$ LDPC code with the source length of $k$, code word length of $n$, and parity check matrix $H$ of $(n - k) \cdot n$, the compression ratio of each bit plane in DVC is

$$R = \mathrm{Frac}[(n - k)]/k, \tag{6}$$

where Frac$\{\bullet\}$ means a fraction of the object, which represents the puncturing ratio. According to the conditions given by Eq. (4), the compression ratio of $B_i^j(X)$ for lossless compression should satisfy

$$R_i^j \geq \{H[B_i^j(X)|B_i^{\{j-1\}}(X), Y_i^k]\}/(1 - p_i^j), \tag{7}$$

where $p_i^j$ is the erasure probability of transmission channel ($C1$ in Fig. 2) for the current bit plane, and $(1 - p_i^j)$ is the channel capacity. In other words, the lower bound of the transmission rate of each bit plane is given by Eq. (7). The number of parity bits $n_i^j$ allocated to $B_i^j(X)$ is

$$n_i^j = R_i^j \cdot k, \tag{8}$$

which is actually decided by the correlation of $Y_i^k$ and $X_i^k$ (i.e., the channel 2) and the bit erasure probability of current bit plane $p_i^j$ (i.e., the channel 1).

### 6.3 *Packetization Scheme*

Packet networks are characterized by the fact that all data are either completely lost or received correctly. In addition, data are first packetized into packets and then put into packet networks. So before transmission, the required parity bits should be first packetized into packets (see Fig. 4); the packet number $n_p$ can be decided as

$$n_p = \sum_{i,j} n_i^j / l_p, \tag{9}$$

where $l_p$ represents packet length in bit.

However, how to packetize the required parity bits (not necessary all the parity bits) for all the bit planes in Fig. 9 is still an open question. Two simple packetization schemes are proposed in the following.

1. Sequential packetization of parity bits, from low-frequency DCT coefficients to high-frequency coefficients, from MSB to LSB. In other words, all the required parity bits are assembled sequentially (i.e., $\{P_0^1, P_0^2, \ldots, P_0^7, \ldots P_{15}^1\}$ in Fig. 9) and then packetized into several packets.
2. Parallel packetization among different bit planes of parity bits. In this packetization scheme, each bit plane $P_i^j$ is divided into $n_p$ fixed length fragments, and an index (i.e., from 1 to $n_p$) is assigned to each one of them. Then the fragments with same index of each bit plane are packetized into one packet (see Fig. 10).

Suppose that the packet bit length is $l_p = 4000$, and the source frame size is common intermediate format $(352 \cdot 288)$ with $k = 6336$ DCT blocks. The packet length and block number (i.e., the frame length of LDPC coding) have the same order of magnitude.

For the sequential packetization, the erasure of one packet will cause the loss of a large portion of or even all of the

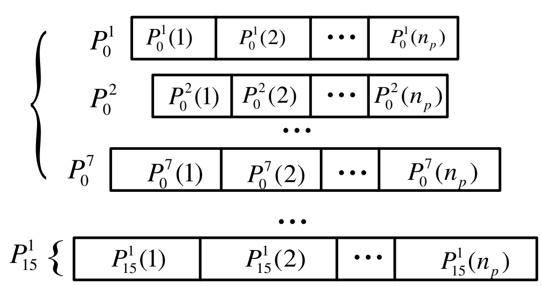


**Fig. 10** Parallel packetization.

required parity bits for one bit plane. Consequently, the decoder will fail to decode such bit plane. In addition, it is difficult to estimate the bit erasure probability $p_i^j$ of each bit plane, which is essential for bit allocation in Eq. (8).

On the contrary, parallel packetization has the great advantage that if a packet is lost, only a portion of data of each bit plane is lost, and the chances for the channel code to effectively recover the lost bits are high (see Fig. 7). At the same time, it guarantees that each bit plane of current frame has the same bit erasure probability, making it easier to estimate from channel conditions. Therefore, in this work, the parallel packetization scheme is employed for better error resilience.

## 7 Rate Allocation in Wyner-Ziv Video Coding

As discussed in previous sections, in order to make the DVC system more practical, the DVC encoder needs to know how many parity bits should be transmitted to the decoder for each bit plane. The correlation between source and SI for each bit plane varies, and the minimum rate for successful decoding varies consequently, so the required rate for decoding varies from bit plane to bit plane. At the same time, the packet loss probability of the channel should also be considered for rate allocation, when the bit stream of DVC is transmitted over packet lossy channels. In other words, the required number of parity bits for perfect decoding should be influenced by both the correlation and packet erasure ratio, since LDPC in Fig. 4 is used for "correcting" the bit errors in both the "virtual correlation channel" and the "real transmission channel." The theoretical rate can be decided by Eqs. (5) to (8).

There are three approaches for rate allocation in WZ codecs: (1) Ideal rate allocation. This scheme performs several decoding attempts, with increasing amounts of parity information, up to successful decoding or transmission of all the parity bits. The packet loss ratio is, of course, constant across the decoding attempts, since, in practical scenarios, the rate should be first allocated before transmission of parity information. The performance provided by this rate-allocation scheme will be regarded in this work as a lower bound indication. (2) Practical rate allocation in decoder, and communicated once to the encoder. In this scheme, the decoder performs the rate allocation and directly informs the encoder. (3) Practical rate allocation in encoder without feedback channel. In this scheme, the encoder performs the rate allocation. In this section, we will discuss the rate-allocation problem with particular reference to the practical transmission of DVC data over packet lossy channels.

### 7.1 *Ideal Rate Allocation: Lowest Rate Bound*

As discussed in Sec. 4.2, the basic DISCOVER codec scheme first requests partial parity bits and attempt decoding. If decoding fails, the decoder will request additional parity bits through the feedback channel and decode again. This DTD process (defined in Sec. 4.2) will be repeated iteratively until successful decoding of the current bit plane or the transmission of all the available parity bits. Typically, several DTD cycles are needed for decoding, introducing a high delay, which is often not suitable for practical video applications. However, this gradual decoding scheme guarantees to use the lowest rate for decoding, given a certain channel code. In this sense, it can be considered as the bound of the system performance. As discussed in Sec. 4.2, in our proposed system, the RA module should first allocate the amount that is needed for WZ decoding at one time before packetization and transmission. In this work we follow the DTD process to establish the system performance bound, so that it can be used as the reference to evaluate the error resilience of the DVC codec.

In the proposed system, the parallel packetization scheme shown in Fig. 10 is used. So the bit erasure ratio equals to $p_i^j$ under packet loss ratio $p$. Since the number of bits for each bit plane under a specific $p$ cannot be decided in advance, in our evaluation system we first compute the ideal number of bits $n_i^j$ for each bit plane when $p = 0$ and calculate the packet number $n_p$ for the current frame using Eq. (9). Then $n_p$ is fixed throughout all the evaluation for each frame. When $p > 0$, probably more parity bits will be required for decoding due to the possible decoding failure caused by the erasure of parity bits packets with a random erasure pattern. In order to decide the minimum rate needed for decoding given a specific $p$ value, the following rate-allocation mechanism is designed:

For each bit plane, suppose the overall parity bits output of LDPC encoder for $i$'th bit plane of $j$'th coefficient band bit plane is $P_i^j$, then the candidate parity bits block set is established as

$$\{P_i^j(s1), P_i^j(s2), \dots, P_i^j(si), \dots, P_i^j\},$$

where $P_i^j(si)$ is a rate-compatible punctured version of the whole output parity bits $P_i^j$, and each element of the set satisfies $P_i^j(si) \in P_i^j(sj)$, $i < j$ for rate compatibility. $P_i^j(s1)$ is the parity block that is needed for the error-free case (i.e., DISCOVER), and is first chosen and packetized into $n_p$ packets, which are subject to a certain packet erasure pattern before they are fed into the LDPC decoder for decoding. In case of decoding failure of $P_i^j(si)$, $P_i^j[s(i+1)]$ is chosen and packetized into $n_p$ packets again, and it is subject to the same packet erasure pattern before it is fed into the decoder to reattempt decoding. This process will repeat until successfully decoding. The data size of $P_i^j(si)$ for successful decoding can be used to decide the minimum rate. Note that the packet length will slightly increase for each reattempt, but this is not a problem if a reasonable packet size is considered in the initial phase.

As discussed in Sec. 6.2, the theoretical compression ratio of $R_i^j$ is influenced by both the correlation between source and SI and the erasure probability $p$. If $R_i^j > 1$, the theory ensures that the bit plane cannot be decoded in any case. In such a condition, there is no need to transmit the parity bits for the bit plane to the decoder, as long as the encoder is *informed* in advance of that decoding impossibility. In theory, this situation could be predicted by means of Eq. (7). In practice, in the proposed system, the rate is set to zero in the final result if decoding fails after all parity bits have been transmitted.

### 7.2 *Practical Rate Allocation*

In order to avoid the iterative operations in practical DVC schemes, the coding rate of each bit plane should first be estimated either at the encoder or decoder depending on the scheme. Then the encoder can determine the number of parity bits that should be transmitted to the decoder.

We also name all this process as bits allocation (BA). As discussed in Sec. 6.2, given the side information $Y_i$, the symbol from former decoded bit planes $B_i^{\{j-1\}}(X)$ and the bit erasure ratio $p_i^j$ for the specific bit plane $B_i^j(X)$, the compression ratio $R_i^j$ for decoding is given by Eq. (7) whose lower bound [i.e., the right-side expression in Eq. (7)] can be approximated as follows:[35]

$$H(B_i^j(X)|B_i^{\{j-1\}}(X), Y_i^k)/(1-p_i^j) \approx \frac{1}{N}\sum_{i=1}^{N} H(\beta_i)/(1-p_i^j), \tag{10}$$

where $H(\beta_i) = -\beta_i \log(\beta_i) - (1-\beta_i)\log(1-\beta_i)$ and $\beta_i$ is the probability of the bit in the current bit plane to be equal to 1, which is given by

$$\beta_i = p(B_i^j(X) = 1|B_i^{\{j-1\}}(X) = q_{Xi}^{\{j-1\}}, Y_i^k = y_i) = \frac{p(B_i^j(X) = 1, B_i^{\{j-1\}}(X) = q_{Xi}^{\{j-1\}}|Y_i^k = y_i)}{p(B_i^{\{j-1\}}(X) = q_{Xi}^{\{j-1\}}|Y_i^k = y_i)}, \tag{11}$$

where $q_{Xi}^{\{j-1\}}$ represents the quantized symbol decided by the previously decoded $(j-1)$ bit planes, $y_i$ represents the DCT coefficient of $i$'th band.

In order to calculate $\beta_i$ in Eq. (11) the correlation between $X$ and $Y$ should be estimated. Most of the previous works model it using a Laplacian distribution[36]

$$p(X|Y) = \frac{\alpha}{2}\exp(-\alpha|X-Y|), \tag{12}$$

where $\alpha$ is the Laplacian distribution parameter.

Now the BA problem has been transformed into the estimation of Laplacian parameter $\alpha$, which can be performed either in decoder[36,37] or encoder[38,39] for practical DVC systems. Both encoder- and decoder-based estimation have no access to some necessary information, namely, $X$ for decoder or $Y$ for encoder. But for decoder-based estimation, the motion field can be used to improve accuracy, since the decoder can perform computationally heavy operations. As for the encoder-based estimation, motion information is not available since in the proposed scheme the encoder computational complexity must be limited. Therefore, the decoder estimation is always better than the encoder one. On the other hand, if the estimation of $\alpha$ is implemented in the decoder, the rate estimation results still need to be transmitted to the encoder through a feedback channel, whereas this is not necessary for the encoder-based estimation.

In practical coding and transmission systems, assuming that quantization is performed according to the matrix shown in Fig. 9, there are 63 numbers, one for each bit plane, which should be transmitted to the encoder. The data size is therefore very small. In addition, the decoder RA only needs one transmission of RA results, which will not cause too much delay compared to a DTD scheme. On the other hand, in most of the packet networks, only the decoder can compute a reliable PLR estimate that can be used in Eq. (10). Therefore, the feedback channel is necessary even if the rate allocation is performed at the encoder. Considering the better performance of the decoder BA, the negligible influence in terms of delay, and the need of a feedback channel, we employ decoder BA in our system. If results in Eq. (10) exceed one, which yields $R_i^j > 1$, the theory ensures that this bit plane cannot be decoded, thus the rate is set to zero.

## 8 Side Information Generation and Reconstruction

Side information is one of the most important factors that influence the coding efficiency of DVC systems. As defined by Slepian-Wolf rate bound in Eq. (1), the rate for perfectly decoding is decided by the correlation between the source $X$ and side information $Y$. The higher the correlation, the lower the rate. In typical DVC frameworks (i.e., the DISCOVER architecture shown in Fig. 3), SI is generated by motion compensated interpolation (MCI)[9] using the adjacent key frames. In the considered practical coding and transmission problems of DVC shown in Fig. 4, the streams of both the key frames and WZ frames are corrupted by packet erasures. The condition of key frames should be taken into consideration for better SI generation. Moreover, when packet losses exceed a certain value, decoding failure of whole bit planes may occur; hence the *reconstruction* module in Fig. 4 should be modified to maximize the decoding performance by taking full advantage of all the successfully decoded bit planes. This section thoroughly investigates these two issues.

### 8.1 *Key Frames Under Packet Loss*

Key frames are the best information to use in order to compute the SI for DVC. Therefore, it is important to investigate the performance of video streams using key frames only under packet loss conditions for reference purposes. Key frames are typically intra-encoded. A sequence composed of all intra-encoded frames is resilient to packet losses when compared to differentially encoded video streams, since there is no temporal error propagation, which is one of the main causes of video quality degradation.

In order to be as much error resilient as possible, we first divide key frames into slices with the basic element of $4 \times 4$ block. Then the video frames are encoded so that all slices have approximately the same size, and each slice is put into one packet. This scheme maximizes the quality of the decoded video since the data contained in every packet can be decoded independently of the others. Thus all received data contribute to improve the quality of the reconstructed video. In addition, concealment techniques can be applied to reconstruct the missing data. Concealment in traditional video coding is often implemented by simply copying the data from the same position in the reference frames. In our system, we employed a simple temporal concealment technique, which substitutes the missing data with the one in the same position in the previous key frame.

### 8.2 *Side Information Generation with Corrupted Key Frames*

Since there is a strong correlation between adjacent frames, the simplest way to generate SI is directly copying adjacent frames or averaging the adjacent forward and backward key frames.[26] Motion field should be considered for more sophisticated SI generation. There are two major traditional algorithms to generate SI at the decoder, the motion compensated interpolation/extrapolation and hash-based motion estimation. The former one has more consistent performance.[9]

The algorithm used in this work is MCI as done in the DISCOVER framework.

For the case of packet losses, the algorithm needs to be modified in order to improve its performance. When key frames are sent on a packet lossy network, it is possible that some frames need concealment due to missing information. The traditional video decoder can easily determine which portion of the video (usually one or more slices) is missing and act consequently. For the purpose of generating the SI in DVC, the algorithm should consider the areas that have been erased and concealed.

The motion vectors of real-world sequences have a centrally biased distribution.[40] In other words, most of the blocks are stationary with little or even no motion with respect to adjacent frames. As shown in Fig. 11, $Y_t$ represents the SI for current WZ frame. For each block $B_t(x,y)$ located at pixel location $(x,y)$ in $Y_t$, if the block at the same position in adjacent key frames [i.e., $B_{t-1}(x,y)$ or $B_{t+1}(x,y)$] is lost, the chances of a poor performance of the MCI is high. In such circumstances, the SI generation algorithm should be modified based on the loss conditions of the various parts of the forward and backward key frames. In order to improve the SI, but without significantly increasing the decoder complexity, a simple algorithm has been proposed as follows:

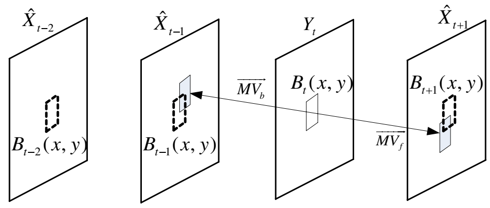



**Fig. 11** Modified SI generation with concealed key frame.

1. Each block of initial SI of current WZ frame is first generated by MCI without considering block erasures.

$$B_t(x,y)=\frac{1}{2}\hat{X}_{t-1}(x+mv_{xb},y+mv_{yb})+\frac{1}{2}\hat{X}_{t+1}(x+mv_{xf},y+mv_{yf}),\tag{13}$$

where $(mv_{xb},mv_{yb})$ and $(mv_{xf},mv_{yf})$ are the backward and forward motion vector represented by $\vec{MV}_b$ and $\vec{MV}_f$ in Fig. 11, computed using the algorithm in Ref. 9.

2. Let $E_{t+1}$ and $E_{t-1}$ be the block loss pattern of forward key frame and backward key frame, where $E(x,y)=1$ means the pixel located at $(x,y)$ is lost and vice versa for 0. Each block of SI is refined by

$$\hat{B}_t(x,y)=\begin{cases}B_t(x,y) & E_{t-1}(x,y)=0 \text{ and } E_{t+1}(x,y)=0\\ B_{t-1}(x,y) & E_{t-1}(x,y)=1 \text{ and } E_{t+1}(x,y)=0\\ B_{t-2}(x,y)/3+B_{t+1}(x,y)\cdot 2/3 & E_{t-1}(x,y)=0 \text{ and } E_{t+1}(x,y)=1\\ B_{t-2}(x,y) & E_{t-1}(x,y)=1 \text{ and } E_{t+1}(x,y)=1\end{cases}.\tag{14}$$

Table 1 shows the improvements of the proposed algorithm over MCI, for different $Q_p\_I$ of key frames and packet loss ratios. The proposed algorithm has very good performance in case of packet loss in key frames, even if sequences present different motion characteristics.

### 8.3 *Reconstruction Function Under Bit Plane Decoding Failure*

In packet lossy channels, it might still happen that, due to a high number of packet losses, the received data is not enough to reconstruct all the bit planes, hence all the bits of each DCT coefficient. Therefore, in this case it is necessary to make full use of successfully decoded bit planes for all the DCT coefficients for the best results. In this work, we propose to use the following strategy. Since quantization has been employed to compute the coefficient, this operation implicitly defines intervals within the valid range of values, and each interval can be specified by the quantization value as its index. Adding more bits is equivalent to reducing the size of those intervals, which can be perfectly defined by the traditional reconstruction function given as follows:[26]

$$\hat{X}_i=\begin{cases}Y_i, & Y_i\in[\hat{p}_i^l,\hat{p}_i^u]\\ \hat{p}_i^l, & Y_i<\hat{p}_i^l\\ \hat{p}_i^u, & Y_i>\hat{p}_i^u\end{cases},\tag{15}$$

where $Y_i$, $i\in[0,15]$ represents the $i$'th DCT coefficients of the SI frame, $\hat{p}_i^l$ is the lower bound value of the quantization interval corresponding to the quantized symbol $q_{Xi}$, and $\hat{p}_i^u$ is the upper bound value of the quantization interval.

For a specific band $i$ and its quantized symbol $q_{Xi}$, $i\in[0,15]$, suppose there are $l_b$ lost bit planes, and the set of lost bit planes (i.e., the unsuccessfully decoded bit planes) is represented by $\mathsf{L}=\{L_i^b,b\in[0,15]\}$, then there are $2^{l_b-1}$ possible quantized symbols. Let us define the set of possible quantized symbol as $\mathbb{Q}=\{q_{Xi}(y),y\in[0,2^{l_b-1}-1]\}$, the corresponding interval setting as $\Omega=\{[\hat{p}_{Xi}^l(y),\hat{p}_{Xi}^u(y)],y\in[0,2^{l_b-1}-1]\}$, and the boundary setting as $\mathsf{B}=\{\hat{p}_{Xi}^z(y),y\in[0,2^{l_b-1}-1],z\in\{l,u\}\}$, then the proposed modified reconstruction function is given by

$$\hat{X}_i=\begin{cases}Y_i, & Y_i\in[\hat{p}_{Xi}^l(y),\hat{p}_{Xi}^h(y)]\\ \hat{p}_{Xi}^{\min}, & \text{otherwise}\end{cases},\tag{16}$$

where

$$\hat{p}_{Xi}^{\min}=\arg\min_{\substack{z\in[l,n]\\ y\in[0,2^{l_b-1}-1]}}\{|\hat{p}_{Xi}^z(y)-Y_i|\}.\tag{17}$$

In other words, we first find the possible quantization bin $\Omega$ compatible with the quantization parameter and available bit planes that are perfectly decoded. If the value of the corresponding coefficient in the SI fits into the interval setting determined by $\Omega$, the value of the SI is used, otherwise

**Table 1** PSNR improvement of proposed SI generation over the MCI algorithm (dB).

| | $Q_p\_I$ | | | |
|---|---|---|---|---|
| Plr | 22 | 24 | 26 | 28 |
| | | Coastguard | | |
| 0.01 | 0.02 | 0.03 | 0.02 | 0.02 |
| 0.05 | 0.14 | 0.16 | 0.12 | 0.11 |
| 0.10 | 0.28 | 0.29 | 0.25 | 0.23 |
| 0.15 | 0.40 | 0.40 | 0.36 | 0.35 |
| 0.20 | 0.40 | 0.50 | 0.47 | 0.45 |
| | | Foreman | | |
| 0.01 | 0.04 | 0.05 | 0.04 | 0.04 |
| 0.05 | 0.25 | 0.20 | 0.20 | 0.17 |
| 0.10 | 0.45 | 0.41 | 0.36 | 0.35 |
| 0.15 | 0.61 | 0.55 | 0.51 | 0.51 |
| 0.20 | 0.71 | 0.66 | 0.64 | 0.63 |
| | | Flower garden | | |
| 0.01 | 0 | 0 | 0 | 0 |
| 0.05 | 0.04 | 0.05 | 0.06 | 0.06 |
| 0.10 | 0.13 | 0.12 | 0.14 | 0.14 |
| 0.15 | 0.20 | 0.19 | 0.20 | 0.21 |
| 0.20 | 0.27 | 0.25 | 0.27 | 0.27 |
| | | Soccer | | |
| 0.01 | 0 | 0 | 0 | 0 |
| 0.05 | 0.01 | 0 | 0 | 0.01 |
| 0.10 | 0.01 | 0.01 | 0 | 0.02 |
| 0.15 | 0.01 | 0.02 | 0.03 | 0.04 |
| 0.20 | 0.02 | 0.04 | 0.04 | 0.05 |

either the upper or lower boundary of the interval, which is nearest to the value of the SI, is used as the decoding output. In this way, a value can be assigned to each DCT coefficient also when not all bit planes could be successfully decoded.

## 9 Simulation Results

In order to validate the error-resilience performance of the proposed schemes and provide fair comparisons with the traditional H.264/AVC scheme, we carried out several experiments. First, we present the video coding parameters and the three types of experimental setups used in the experiments. Second, the experimental results for each setup are reported and commented.

### 9.1 *Coding Conditions and Experimental Setup*

In all the experiments, the luminance of four standard video test sequences with widely different characteristics (first 200 frames of foreman, coastguard, flower garden, and all frames of soccer) have been encoded at 30 fps, quarter common intermediate format resolution with a group of picture (GOP) length of 2, using both the proposed DVC system and H.264/AVC, to provide a comprehensive view about the performance of the proposed schemes with respect to H.264/AVC. The GOP pattern "I-WZ-I" is employed for DVC and the GOP pattern "I-B-I" with no P-type frames is employed for H.264/AVC. I-type frames for DVC are the frames referred to as the key ones in the previous discussion and are also H.264/AVC-encoded. For the WZ frames, five quantization matrixes following Ref. 41 are used in DVC for different RD points. The quantization parameters (QP) for I-type frames are set so that the average quality (measured according to PSNR) of the WZ frames and key frames is similar. Apart from the modules whose algorithms are proposed in this work, the other modules of the DVC codec mainly follow the DISCOVER system described in Ref. 9. For H.264/AVC test, various QP values [i.e., (16, 18, 20, 22, 24, 26, 28)] have been employed in order to test several H.264/AVC video bitrates. The test model software JM v.11.0[42] main profile has been employed to perform H.264/AVC and DVC key frame encoding.

In the experiments, the RD performance bound for the case of no losses is established first. Then, we consider packet lossy networks with uniform packet loss probability. The packet length is set to be about 500 bytes for both I-type frames and B/WZ frames. The transmission performance over packet loss channels with different packet loss ratio [i.e., $\mathrm{PLR} = (1\%, 5\%, 10\%, 15\%, 20\%)$] is evaluated. For each condition 20 experiments with different randomly generated packet loss patterns are performed, and their results have been averaged. Note that since we kept the maximum packet length fixed, the number of packets varies for I-type frames at different QP. For example, when $\mathrm{QP} = 28$ the data size of the fifth frame of the foreman sequence is 2896 bytes, and the number of packets needed is six. This implies that the probability of the whole frame loss is $\mathrm{PLR}^6$, which is much lower than PLR. In other words, this packetization scheme will provide much better SI quality than putting the I-type frame into a single packet. As for WZ/B frames, since the amount of data is generally lower, the number of packets is lower, and the chances of whole frame loss are higher than I-type frames. In this case, our proposed algorithm can handle it by directly using SI as the WZ frames are decoded. For the B frames case a standard, temporal error concealment strategy is used, where missing information is replaced with the one in the same position in the previous frame.

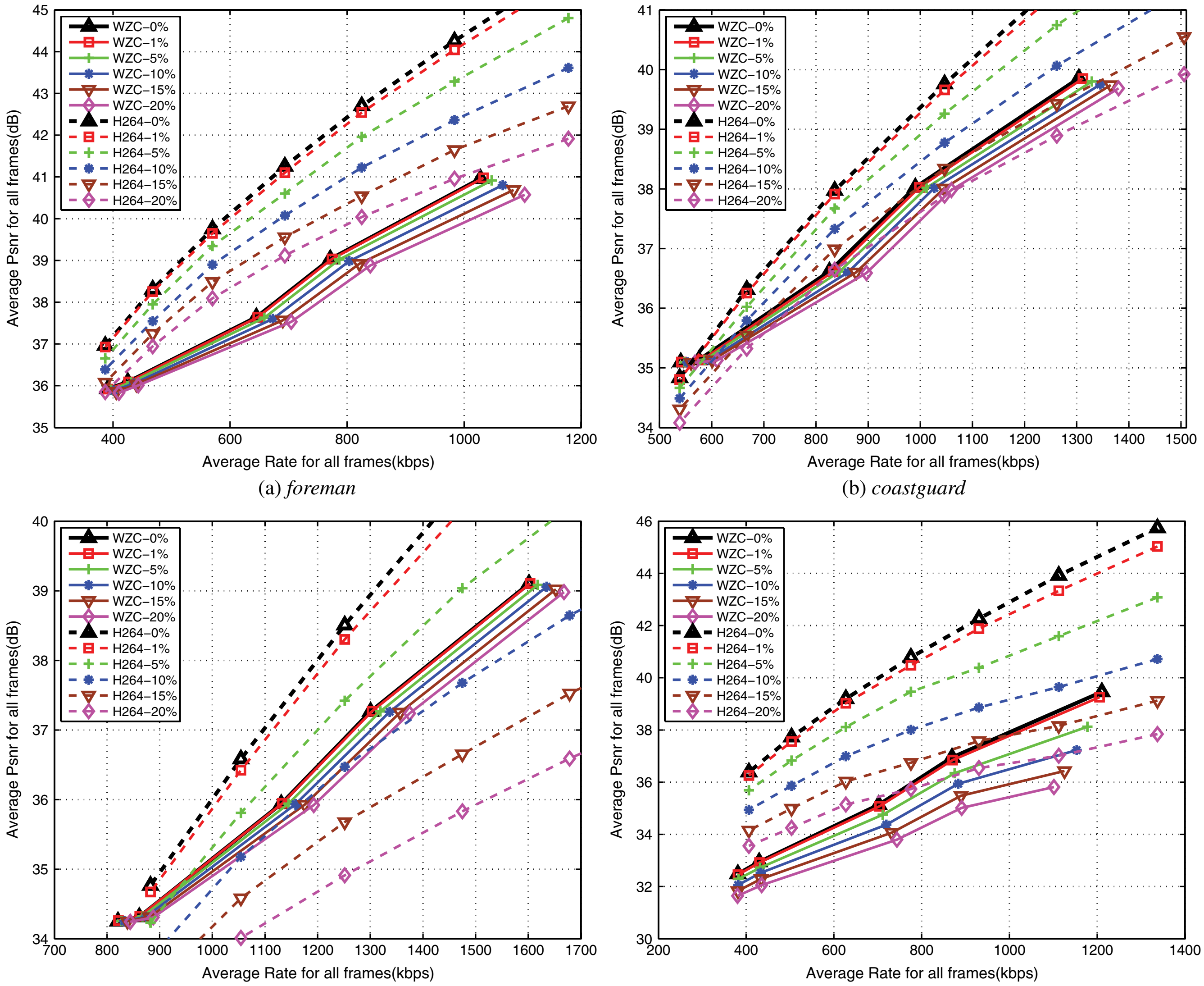
All RD performance graphs show the PSNR value computed as the average over all frames. Also, note that the results showing WZC performance refer to our proposed parallel packetization scheme, since the performance using sequential packetization would be very poor, besides being arduous to implement due to the difficulties in estimating the bit erasure probability for each bit plane as explained in Sec. 6.3. Three types of setups are defined as follows to evaluate the contribution of each part of the proposed scheme:

1. Packet losses only on B/WZ frames with ideal rate allocation: In this setup, the packet losses only affect the B and WZ frames (i.e., all key frames are loss-free). The main purpose of this experiment is to evaluate the error resilience of the WZ frames themselves.
2. Packet losses on all frames with ideal rate allocation: This setup considers a more realistic situation where losses affect all transmitted packets. Both I as well as B and WZ frames are subject to losses. The concealment algorithm for I frames and SI generation algorithm are the ones described in Sec. 6. The ideal rate-allocation algorithm presented in Sec. 8.1 is used in WZ coding (i.e., the decoder requests more parity bits from the encoder through a feedback channel until successful decoding or the transmission of all the parity bits).
3. Packet losses on all frames with practical decoder rate allocation: In these experiments, rate allocation is performed in the decoder using the algorithm in Sec. 8.2. The other conditions are the same as setup No 2. Since the rate for each frame is decided and transmitted to the encoder only once, this setup is much more practical than the aforementioned setup No 2.

### 9.2 *Experimental Results and Analysis*

#### 9.2.1 *Packet losses only on B/WZ frame with ideal rate allocation*

Figure 12 shows the RD performance considering packet losses ranging from 0% to 20% but affecting only B and WZ frames. In this scenario, video sequences coded using H.264/AVC (H264) often show better performance




(a) *foreman* (b) *coastguard*

(c) *flower_garden* (d) *soccer*

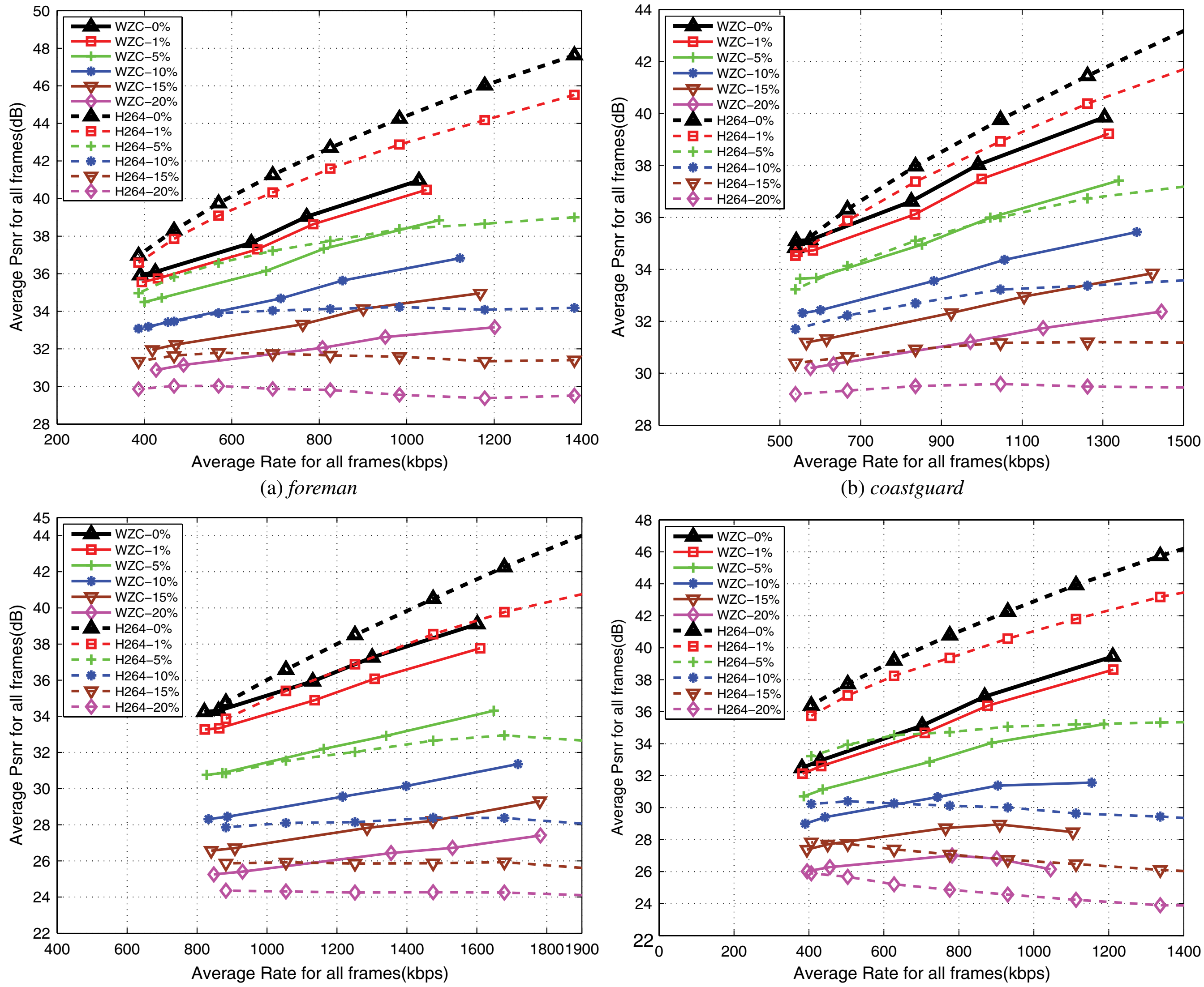
**Fig. 12** RD performance of H.264/AVC (dashed line) and DVC schemes (solid line) in presence of losses on B and WZ frames only.

compared to the DVC counterparts (WZC). In all sequences, the performance gap can be explained considering that B frames are more efficient than WZ ones in RD terms. These results are expected, since the state-of-art DVC frameworks are still far from traditional H.264/AVC in terms of compression performance.

However, it should also be noted that the performance of the DVC scheme, for all sequences, is only marginally affected by the increased PLR, while in case of the H.264/AVC sequences it degrades significantly. So the proposed WZ coding and transmission scheme has much better error resilience than H.264/AVC. This is, in general, an advantage for DVC as new higher performance DVC schemes will be proposed. For the case of flower garden and coastguard, for high bitrates and PLR, WZ coded sequences even slightly outperform the H.264/AVC ones, up to 0.3 dB for coastguard and 2 dB for flower garden in PSNR. The main reason is that these two sequences have more uniform motion due to camera panning, which allows the MCI algorithm to generate better SI, which resulted in better overall system performance.

### 9.2.2 *Packet losses on all frames with ideal rate allocation*

For a more realistic situation where losses affect all packets, Fig. 13 shows the RD performance of both the H.264/AVC (H264) and the proposed DVC framework with ideal rate allocation in decoder (WZC). While at low PLR, the H.264/AVC sequences present a better quality compared with DVC ones, for all sequences the situation reverses as PLR and bitrate increase. For instance, the crossover point for 5% PLR is at about 1000 kbit/s for foreman, coastguard, and flower garden, while it is about 1200 kbit/s for soccer. For higher PLR values, the crossover point moves toward lower bitrates: for 10% PLR it is about 500 kbit/s (600 kbit/s for soccer), while for higher PLR the DVC scheme with the proposed SI generation algorithm for packet




(a) *foreman* (b) *coastguard*

(c) *flower_garden* (d) *soccer*

**Fig. 13** RD performance of H.264/AVC (dashed line) and DVC schemes (solid line) when losses affect all frames, where ideal rate allocation is assumed.

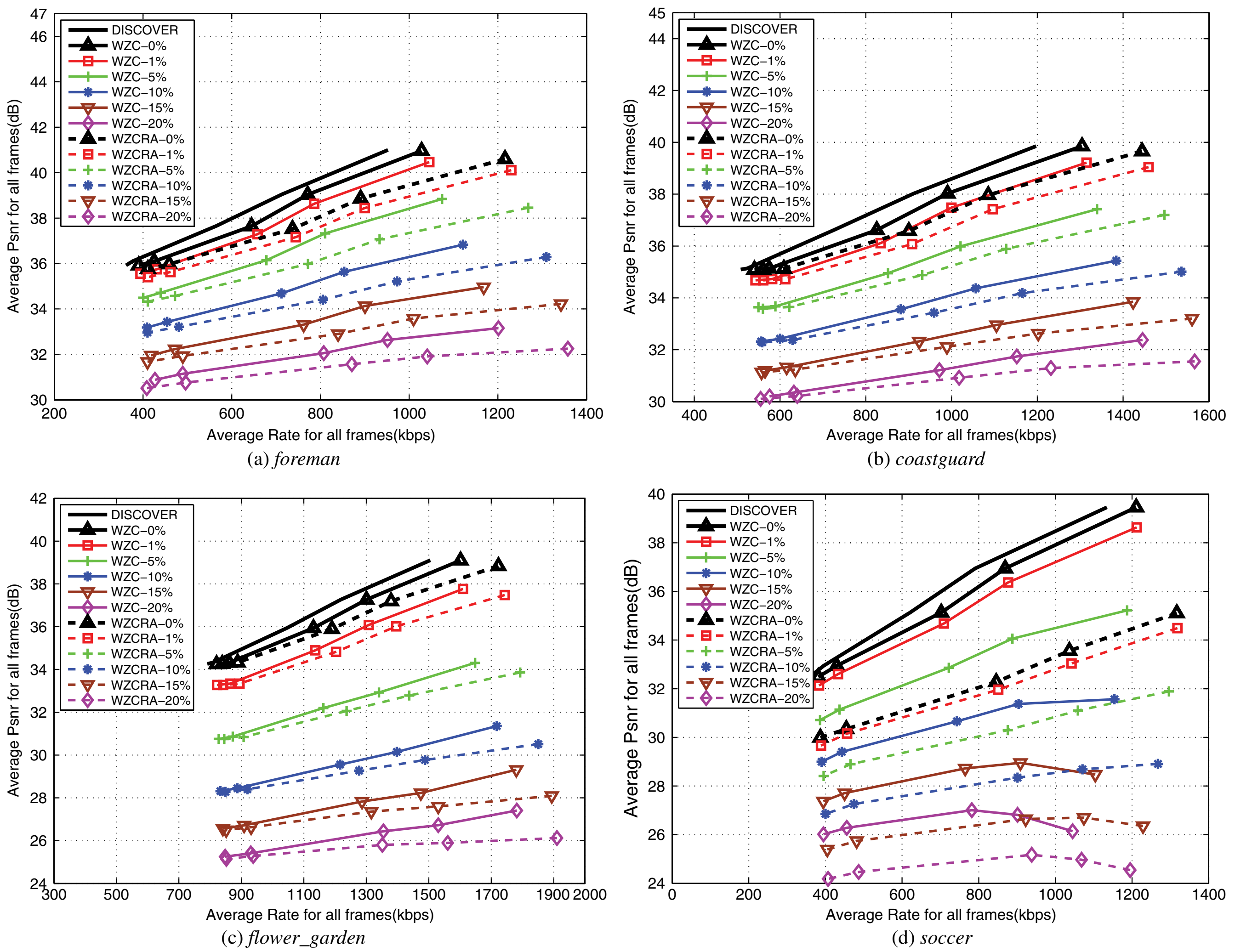
loss scenarios outperforms H.264/AVC at almost all tested bitrates.

### 9.2.3 *Packet losses on all frames with decoder rate allocation*

For the proposed practical DVC system with decoder rate allocation that does not rely on extensive feedback information from the receiver as in widespread DVC schemes, Fig. 14 shows the RD performance comparing ideal rate allocation (WZC) and practical decoder rate allocation (WZCRA). The performance of the DISCOVER codec based on LDPCA and the DTD loop in case of no losses are also given in Fig. 14 for reference. Note that in Sec. 4.2 we have shown that the DISCOVER codec cannot directly be employed for transmission in packet lossy channel due to the use of a DTD loop that would severely have an impact on the system latency. In addition we have also shown in Sec. 5.3 that the LDPCA-based SWC performance is much worse than the LDPC-based SWC when the erasure rate is greater than 10%. Due to the lack of a publicly available DISCOVER implementation that can handle data losses, we could not show the DISCOVER performance at PLR > 0. However, the previous considerations suggest that its performance would significantly drop as the PLR increases, besides the fact that the DTD loop does not allow to use it in practical scenarios.

As expected, the DISCOVER based on LDPCA performance is slightly better than our LDPC-based system at PLR = 0 (i.e., 0.3 to 0.8 dB). The performance slightly decreases for the case of WZCRA with respect to WZC due to the exchange of less information in WZCRA. The reduction ranges from 0.2 to 1 dB for most sequences (i.e., foreman, coastguard, flower garden). Note that the gap for the soccer sequence is more pronounced, ranging from 3 to 5 dB PSNR, suggesting that sequences with high motion particularly suffer from the absence of extensive receiver feedback. This is reasonable since high motion results in poor SI generated by MCI. As a consequence the overall performance decreases.

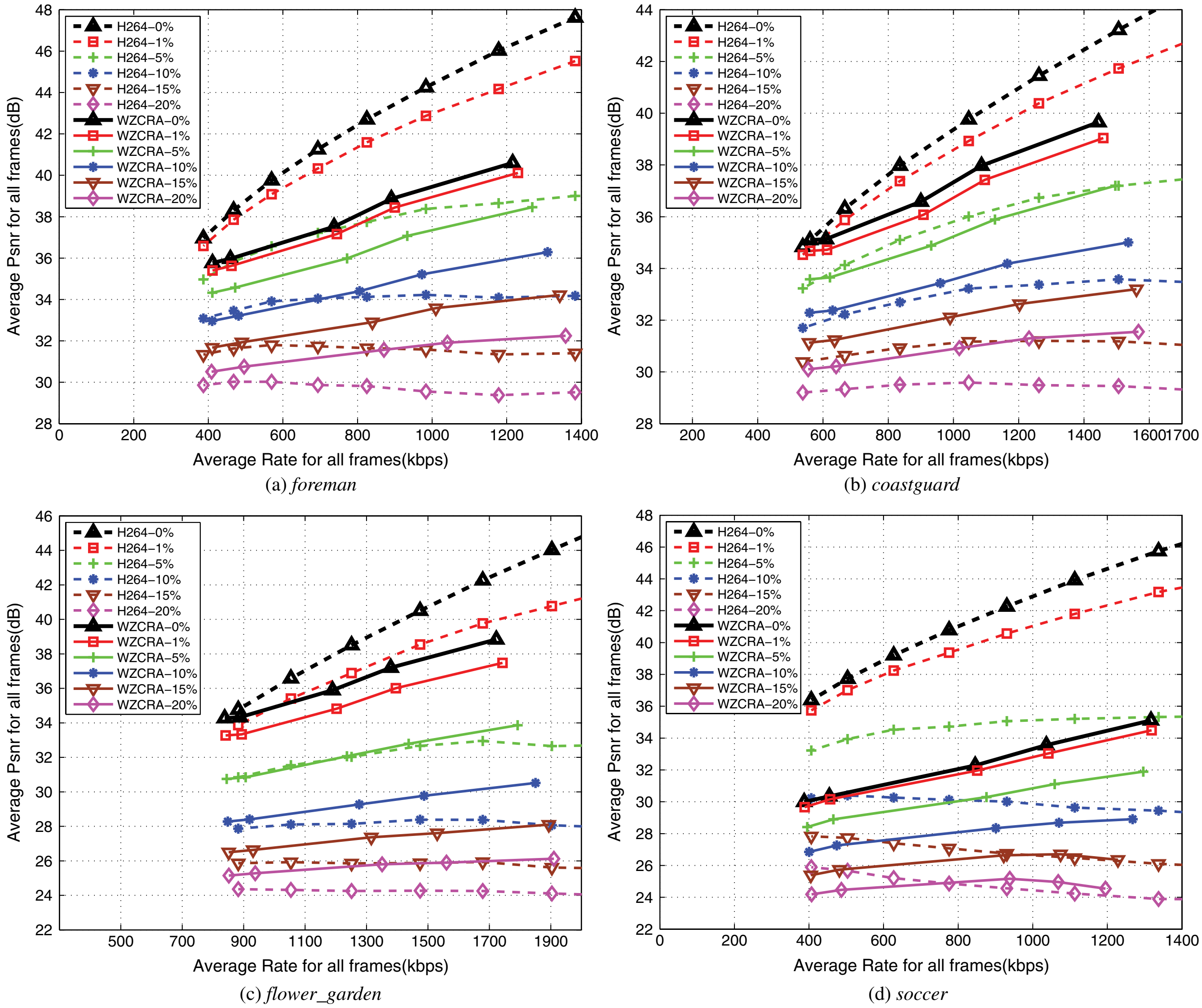
In order to clearly show the error-resilience performance, the RD performance of the proposed practical DVC scheme with decoder rate allocation (WZCRA) are also directly compared with H.264/AVC (H264) in Fig. 15. The results make our proposal for a practical DVC scheme even more interesting, although the performance, as expected, slightly decreases due to the limitation on feedback information. The behavior shown in the previous set of results in Eq. (2) still holds, i.e., DVC is more convenient as the




(a) *foreman* (b) *coastguard*

(c) *flower_garden* (d) *soccer*

**Fig. 14** RD performance of DVC with ideal RA (solid line) and proposed RA (dashed line).




(a) *foreman* (b) *coastguard*

(c) *flower_garden* (d) *soccer*

**Fig. 15** RD performance of H.264/AVC (dashed line) and the proposed practical DVC scheme (solid line) with the proposed rate allocation algorithm.

PLR increases. For most sequences, the proposed DVC scheme outperforms H.264/AVC when PLR > 10%. Even for the soccer sequence with very high motion (which causes very poor DVC coding performance), the proposed practical DVC scheme outperforms H.264/AVC when the rate exceeds 700 kbps at 20% PLR and 1 Mbps at 15% PLR.

### 9.3 *Discussion*

As shown by the experimental results, the proposed practical DVC system is effective and practical for transmission over packet lossy channels. The key strengths are that the DTD loop has been discarded, the data size is allocated before packetization, and the data are packetized properly into packets with reasonable length. Figures 12 and 13 clearly show that if the rate is perfectly allocated for both the virtual channel and the real transmission channel, the proposed practical DVC system can achieve better performance than H.264/AVC at high PLR. Note the proposed DVC framework is based on the start-of-art DISCOVER architecture, which still presents a gap in terms of coding performance with respect to H.264/AVC for the case of no transmission error. Therefore, we can conclude that the proposed practical system presents much better error resilience than H.264/AVC. Specifically, the careful choice of the channel coding scheme based on a comprehensive analysis of the most suitable channel codes and the proposed packetization scheme that takes into account the error resilience and the rate allocation algorithm now ready for use in a practical scenario have significantly improved the error resilience of the DVC system. The SI generation based on the loss pattern of the key frame and the reconstruction function, which makes full use of successfully decoded bit planes, significantly improved the decoding quality of video frames. Although the loss pattern is needed for SI, this information can be easily extracted by inspecting, e.g., packet fields in the network layer protocols such as the sequence number of the real-time transport protocol.

Figures 14 and 15 have also shown that although the proposed practical rate-allocation algorithm introduced some performance losses for low packet loss rates, the final system is still better than H.264/AVC when the network condition degrades. This shows that our proposed rate-allocation algorithm is effective.

While our system achieves good performance, the modifications and additions with respect to the original DISCOVER architecture are limited in terms of complexity. No additions are needed for the encoder except for the packetization, which is, however, necessary for any transmission system. For the decoder case, the only additional complexities are as follows: (1) Estimating the rate for each bit plane depending on the quality of the received SI and on the PLR which is typically provided by transmission protocol such as the real-time transport protocol. (2) Reconstructing the information on the basis of all the already successfully decoded bit planes. Since the most complex tasks are confined within the decoder, the previous modifications perfectly integrates into the original DISCOVER architecture.

## 10 Conclusion

This paper presented a framework to optimize the quality of video communications based on distributed video coding in packet lossy network scenarios. The peculiarities of the distributed video-coding paradigm have been reviewed in order to investigate the issues in transmitting such type of compressed data over packet lossy networks. A number of adaptations have been proposed to take full advantage of the intrinsic robustness of distributed video coding to data loss. A new packetization scheme is proposed, modifications to rate-allocation mechanism have been investigated in detail in order to make them more practical, new procedures for the SI generation, and reconstruction functions are also proposed for better performance. Results show that all these new proposals significantly contribute to the performance of DVC in a practical scenario. Moreover, results also include extensive performance comparisons of the improved DVC schemes with respect to a conventional packet video communication using the H.264/AVC coding standard. The proposed techniques provide comparable or better performance than a traditional H.264/AVC-based system especially when PLR increases. The proposed techniques are an effective first step in the direction of a practical system for a high-performance error-resilient DVC transmission over traditional packet lossy networks. Future work will be devoted to perform further comparison with the emerging HEVC standard.

### Acknowledgments

This work has been supported in part by the EU FP7-PEOPLE-IRSES-2009 S2EuNet Project under Grant Agreement No. 247083, National Natural Science Foundations of China (61201388), and Doctor Station Foundation of National Ministry of Education of China (20110181120009).

### References

1. H. K. Arachchi and W. A. C. Fernando, "H.264/AVC in mobile environment: a review," *Opt. Eng.* **44**(9), 097006 (2005).
2. ITU-T & ISO/IEC, "Advanced video coding for generic audiovisual services," doc. number H.264 (ITU-T) and 14496-10 (ISO/IEC), ITU, Geneva, Switzerland, p. 680, (January 2012), http://www.itu.int/rec/T-REC-H.264-201201-I/en.
3. ITU-T & ISO/IE, "High efficiency video coding" doc. number H.265 (ITU-T) and 23008-2 (ISO/IEC), text has been consented ("approved") (25 January 2013), http://www.iso.org/iso/catalogue_detail.htm?csnumber=35424.
4. B. Li, G. Sullivan, and J. Xu, "Comparison of compression performance of HEVC working draft 4 with AVC high profile," JCTVC-G399, http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=3656, (2011).
5. P. Hanhart et al., "Subjective quality evaluation of the upcoming HEVC video compression standard," *Proc. SPIE* **8499**, 84990V (2012).
6. K. E. Psannis and Y. Ishibashi, "Efficient error resilient algorithm for H.264/AVC: mobility management in wireless video streaming," *Telecommun. Syst.* **41**(2), 65–76 (2009).
7. , "Quality scalability in H.264/AVC video coding," *Proc. SPIE* **5960**, 59601O (2006).
8. B. Girod et al., "Distributed video coding," *Proc. IEEE* **93**(1), 71–83 (2005).
9. X. Artigas, J. Ascenso, and M. Dalai, "The DISCOVER codec: architecture, techniques and evaluation," in *Proc. of Picture Coding Symposium*, EURASIP, Lisbon, Portugal (2007).
10. "PRISM: a new 'reversed' multimedia coding paradigm," in *Proc. Int. Conf. on Image Process.*, Vol. 1, pp. 617–620, IEEE, Piscataway, New Jersey (2003).
11. Y. Wang et al., "Error resilient video coding techniques," *IEEE Signal Process. Mag.* **17**(4), 61–82 (2000).
12. T. Wiegand et al., "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circ. Syst. Video Technol.* **13**(7), 560–576 (2003).
13. S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circ. Syst. Video Technol.* **13**(7), 645–656 (2003).
14. S. Rane, P. Baccichet, and B. Girod, "Systematic lossy error protection of video signals," *IEEE Trans. Circ. Syst. Video Techol.* **18**(10), 1347–1360 (2008).
15. M. Ramon, F.-X. Coudoux, and M.-G. Gazalet, "Efficient priority encoding transmission scheme based Wyner-Ziv coding and spatial scalability," *Proc. SPIE* **5960**, 59605K (2006).
16. Z. Xue et al., "Error-resilient scheme for wavelet video codec using automatic ROI detection and Wyner-Ziv coding over packet erasure channel," *IEEE Trans. Broadcast.* **56**(4), 481–493 (2010).
17. A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv coding of video: an error-resilient compression framework," *IEEE Trans. Multimedia* **6**(2), 249–258 (2004).
18. W. Wang, Y. Zhao, and J. S. Pan, "Multiple description image coding based on DSC and pixel interleaving," in *Proc. Int. Conf. Intell. Information Hiding and Multimedia Signal Process.*, pp. 691–694, IEEE, Piscataway, New Jersey (2008).
19. M. Tagliasacchi et al., "Robust wireless video multicast based on a distributed source coding approach," *Signal Process.* **86**(11), 3196–3211 (2006).
20. C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain Wyner-Ziv video coding performance," in *IEEE Int. Conf. on Acoustics, Speech and Signal Process.*, Vol. 2, p. II, IEEE, Piscataway, New Jersey (2006).
21. J. Q. Pedro et al., "Studying error resilience performance for a feedback channel based transform domain Wyner-Ziv video codec," in *Proc. Picture Coding Symposium*, p. 4, EURASIP, Bruxelles, Belgium (2007).
22. C. Brites, J. Ascenso, and F. Pereira, "Feedback channel in pixel domain Wyner-Ziv video coding: myths and realities," in *Proc. 14th European Signal Process. Conf.*, p. 5, EURASIP, Bruxelles, Belgium (2006).
23. C. Tonoli, P. Migliorati, and R. Leonardi, "Error resilience in current distributed video coding architectures," *J. Image Video Process.* **2009**, 1–17 (2009).
24. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theor.* **19**(4), 471–480 (1973).
25. S. Shamai and S. Verdù, "Capacity of channels with uncoded side information," *Euro. Trans. Telecommun.* **6**(5), 587–600 (1995).
26. A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," in *Proc. of the 36th Asilomar Conf. on Signals, Systems and Computers*, Vol. 1, pp. 240–244, IEEE, Piscataway, New Jersey (2002).
27. D. He, A. Jagmohan, and L. Lu, "Wyner-Ziv video compression using rateless LDPC codes," *Proc. SPIE* **6822**, 68221U (2008).
28. D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *Signal Process.* **86**(11), 3123–3130 (2006).
29. D. Kubasov, K. Lajnef, and C. Guillemot, "A hybrid encoder/decoder rate control for Wyner-Ziv video coding with a feedback channel," in *9th IEEE Workshop on Multimedia Signal Process.*, pp. 251–254, IEEE, Piscataway, New Jersey (2007).
30. M. Morbee, J. Prades-Nebot, and A. Pizurica, "Rate allocation algorithm for pixel-domain distributed video coding without feedback channel," in *Proc. Acoustics, Speech and Signal Process.*, Vol. 1, pp. 521–524, IEEE, Piscataway, New Jersey (2007).
31. Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Proc. Mag.* **21**(5), 80–94 (2004).
32. P. Jagatheeswari and M. Rajaram, "Performance comparison of LDPC codes and turbo codes," *Eur. J. Sci. Res.* **54**(3), 465–472 (2011).
33. L. Liu and E. J. Delp, "Wyner-Ziv video coding using LDPC codes," in *Proc. Signal Process. Symp.*, pp. 258–261, IEEE, Piscataway, New Jersey (2006).
34. W. E. Ryani, "An introduction to LDPC codes," in *CRC Handbook for Coding and Signal Processing for Recording Systems*, B. Vasic and E. Kurtas, Eds., pp. 1–23, CRC Press, Boca Raton, Florida (2004).

35. S. Cheng and Z. Xiong, "Successive refinement for the Wyner-Ziv problem and layered code design," *IEEE Trans. Signal Process.* **53**(8), 3269–3281 (2005).
36. C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. Circ. Syst. Video Technol.* **18**(9), 1177–1190 (2008).
37. L. Qing, X. He, and R. Lv, "Modeling non-stationary correlation noise statistics for Wyner-Ziv video coding," in *Proc. Int. Conf. Wavelet Analysis and Pattern Recognit.*, pp. 316–320, IEEE, Piscataway, New Jersey (2007).
38. T. Sheng, X. Zhu, and G. Hua, "Feedback-free rate-allocation scheme for transform domain Wyner-Ziv video coding," *Multimedia Syst.* **16**(2), 127–137 (2010).
39. C. Brites and F. Pereira, "Encoder rate control for transform domain Wyner-Ziv video coding," in *Proc. of IEEE Intl. Conf. on Image Process.*, Vol. 2, pp. 5–8, IEEE, Piscataway, New Jersey (2007).
40. J. Y. Tham, S. Ranganath, and M. Ranganath, "A novel unrestricted center-biased diamond search algorithm for block motion estimation," *IEEE Trans. Circ. Syst. Video Technol.* **8**(4), 369–377 (1998).
41. A. Aaron et al., "Transform-domain Wyner-Ziv codec for video," *Proc. SPIE*, **5308**, 520–528 (2004).
42. Joint Video Team of ISO/IEC MPEG & ITU-T VCEG, "Joint Model Number 11.0 (JM-11.0)," (29 January 2013), http://iphome.hhi.de/suehring/tml/download.

**Linbo Qing** received a PhD in communication and information systems from the Sichuan University, Chengdu, China, in 2008. From February to August 2011, he was a visiting researcher at the Agder Mobility Laboratory, University of Agder, Norway, where he worked on data aggregation and security in mobile communications. From September 2011 to February 2012, he was a visiting researcher at the Internet Media Group, Politecnico di Torino, Turin, Italy, where he worked on distributed video coding and transmission. He is currently an assistant professor in the College of Electronics and Information Engineering at Sichuan University. He coauthored over 30 scientific publications in image processing related topics. His main research interests include image processing, video coding and transmission, and information theory.

**Enrico Masala** received an MSc (summa cum laude) and a PhD in computer engineering from the Politecnico di Torino, Turin, Italy, in 1999 and 2004, respectively. In 2003, he was a visiting researcher at the Signal Compression Laboratory, University of California, Santa Barbara, where he worked on joint source channel coding algorithms for video transmission. He is currently an assistant professor in the Control and Computer Engineering Department at the Politecnico di Torino. He coauthored over 40 scientific publications in multimedia communications related topics. His main research interests include processing, coding, and robust transmission of multimedia signals (especially video) over packet networks.

**Xiaohai He** received his PhD in biomedical engineering from the Sichuan University, Chengdu, China. He is currently a professor in College of Electronics and Information Engineering at Sichuan University. He is an editor of the *Journal of Information and Electronic Engineering* and the *Journal of Data Acquisition & Processing*. His research interests include image processing, pattern recognition, and image communication. He is a senior member of the Chinese Institute of Electronics.