# Multiple description video coding and iterative decoding of LDPCA codes with side information


Olivier Crave [1,2], Christine Guillemot [1], and Béatrice Pesquet-Popescu [2]

[1] IRISA / INRIA, Rennes, FRANCE

[2] TELECOM ParisTech, Signal and Image Proc. Dept., Paris, FRANCE


## Abstract


Multiple description coding (MDC) with side information (SI) at the receiver is particularly relevant for robust transmission in sensor networks where correlated data is being transmitted to a common receiver, as well as for robust video compression. The rate-distortion region for this problem has been established in [**?**]. Here, we focus on the design of a practical MDC scheme with SI at the receiver. It builds upon both MDC principles and Slepian-Wolf (SW) coding principles. The input source is first quantized with a multiple description scalar quantizer (MDSQ) which introduces redundancy or correlation in the transmitted streams in order to take advantage of the path diversity. The resulting sequences of indexes are SW encoded, that is separately encoded and jointly decoded. While the first step (MDSQ) plays the role of a channel code, the second one (SW coding) plays the role of a source code, compressing the sequences of quantized indexes. In a second step, the cross-decoding of the two descriptions is proposed. This allows us to account for both the correlation with the SI as well as the correlation between the two descriptions.

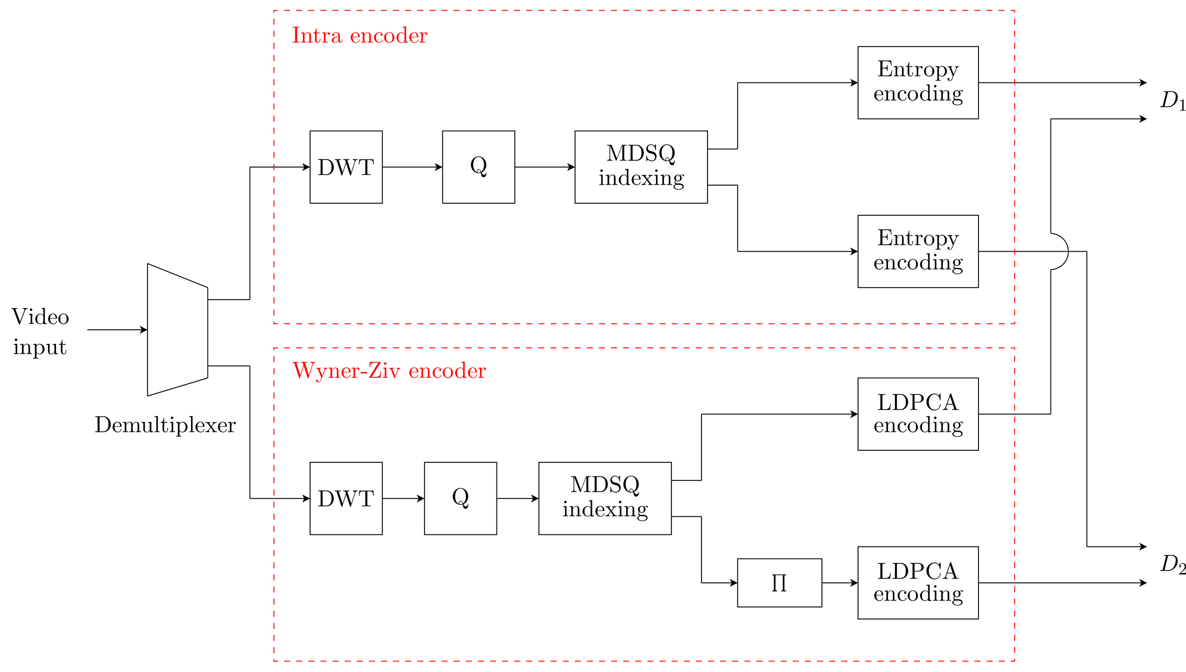

## Introduction

- robust distributed video coding
- multiple description coding
- increase the robustness of distributed video coding while keeping good rate-distortion performance
- video sequence structured into key frames and Wyner-Ziv frames
- for each type of frame, two descriptions are generated by a multiple description scalar quantizer and sent on a loss-prone channel
- when both Wyner-Ziv descriptions are received, they are jointly decoded along with the side information
- investigate the influence of the amount of redundancy and of the iterative decoding of the descriptions on the performance
- goals:
  - improve the robustness of distributed video coding while at the same time keeping good RD performance
  - a low-complexity at the encoder
  - preventing the introduction of quality fluctuations in the decoded video sequences along time
- solution: generating two descriptions for each type of frame (KF and WZF) by a multiple description scalar quantization (MDSQ)
- two balanced descriptions that contain partial information about WZF and KF
- to exploit the redundancy between the descriptions, jointly iteratively decoding the descriptions at the central decoder

## Multiple description video coding with side information

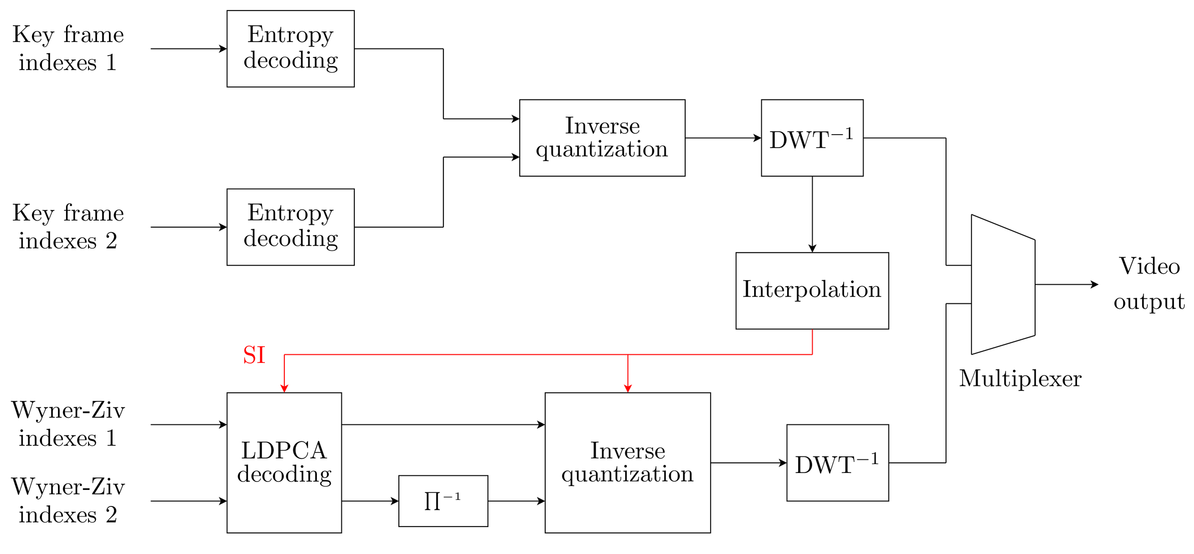
### Encoding

- split the sequence:
  - key frames coding: intra encoder (JPEG2000)
    1. spatial discrete wavelet transform (DWT)
    2. multiple description scalar quantization (MDSQ)
    3. entropy coded using variable-length codes
  - transform-domain Wyner-Ziv coding:
    1. DWT
    2. MDSQ (same index assignment)
    3. bitplane per bitplane encoding by low-density parity-check accumulate (LDPCA) codes
    4. only the produced accumulated syndromes are buffered into two descriptions
- two balanced descriptions that contain partial information about every frame in the video input




### Central decoding

1. entropy decoding of the key frame indexes
2. transform coefficients reconstructed and inverse transformed
3. side information interpolated from the key frames
4. decode the Wyner-Ziv indexes with the accumulated syndrome bits transmitted in small amounts upon the decoder request via the feedback channel
5. indexes decoded separately or jointly (cross-decoding)
6. matrix lookup to get the quantized coefficients
7. reconstruction with the help of the side information
8. inverse transform to retrieve the Wyner-Ziv frames which are combined with the previously decoded key frames

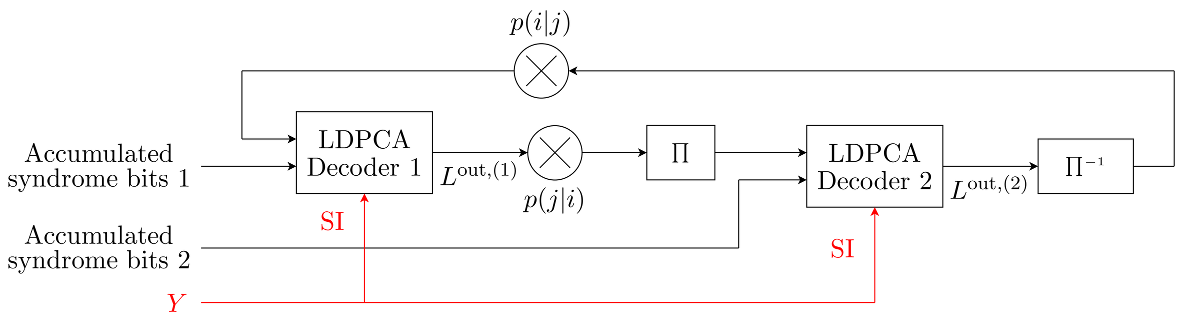



### Side decoding

- same procedure with only one description

## Iterative decoding of LDPCA codes

- joint decoding of descriptions with side information
- goal: exploit the correlation between descriptions
- principle: exchange extrinsic informations between the two decoders




LDPCA cross-decoding with side information

$\{X_n, n=1,2,...,N\}$ are the samples of a memoryless i.i.d source It is encoded at an average rate of $r$ bits per sample (bps) per channel two correlated bitstreams, $\boldsymbol{u}^{(s)}=\{u_1^{(s)},\dots,u_{rN}^{(s)}\}, s=1,2$ Each bitstream is separately encoded using an LDPCA encoder At the receivers, a bitstream of information bits is obtained from the SI, $\boldsymbol{y}=\{y_1,\dots,y_N\}$. transfer of information from the 1st decoder to the 2nd decoder When the estimates of the source bits of the first decoder converge, the log-likelihood ratio (LLR) $L^{\text{out},(1)}_{(k-1)r+t}$ is sent from the check nodes to the variable node $v^{(1)}_{(k-1)r+t}$ where $k=1,\dots,N$, $t=1,\dots,r$ send to the other decoder as an *a priori* information:

$$P(u^{(2)}_{(k-1)r+t}=1)=P(u^{(2)}_{(k-1)r+t}=1|u^{(1)}_{(k-1)r+t}=1)\times P(u^{(1)}_{(k-1)r+t}=1) + P(u^{(2)}_{(k-1)r+t}=1|u^{(1)}_{(k-1)r+t}=0)\times P(u^{(1)}_{(k-1)r+t}=0)$$

$\forall k\in\{1,\dots,N\}$ (i.i.d samples),

$$P(u^{(2)}_{(k-1)r+t}=1|u^{(1)}_{(k-1)r+t}=1)=\sum_{\substack{l:b_t(l)=1\\ m:b_t(m)=1}} P(j=m|i=l)$$

$$P(u^{(2)}_{(k-1)r+t}=1|u^{(1)}_{(k-1)r+t}=0)=\sum_{\substack{l:b_t(l)=0\\ m:b_t(m)=1}} P(j=m|i=l)$$

where $l\in\{1,\dots,M\}$, $m\in\{1,\dots,M\}$, $M$ is the size of the alphabet of the indexes, and $\{b_t(l), t=1,\dots,r\}$ is the binary representation for the quantizer index $l$. $i$ and $j$ are the row and column indexes of the index assignment matrix.

The conditional probabilities are obtained from the index assignment matrix by using the distribution model of the quantized wavelet coefficients in the subbands.

- distribution model:
  - approximation subband: Gaussian
  - detail subbands: Laplacian
- parameters estimated from the SI

The LLRs for the second description are obtained from the bit probabilities

$$P(u^{(2)}_{(k-1)r+t}=1)=\sum_{\substack{l:b_t(l)=1\\ m:b_t(m)=1}} P(j=m|i=l)\times P(u^{(1)}_{(k-1)r+t}=1) + \sum_{\substack{l:b_t(l)=0\\ m:b_t(m)=1}} P(j=m|i=l)\times P(u^{(1)}_{(k-1)r+t}=0),$$

$$P(u^{(2)}_{(k-1)r+t}=0)=1-P(u^{(2)}_{(k-1)r+t}=1),$$

$$L^{\text{in},(2)}_{(k-1)r+t}=\log\frac{P(u^{(2)}_{(k-1)r+t}=1)}{P(u^{(2)}_{(k-1)r+t}=0)}$$

These LLRs are used by the second decoder as *a priori* information. After convergence, the second decoder generates *extrinsic* log-likelihoods for the first decoder. An interleaver before the encoding of one of the descriptions is necessary to make sure that the information contained in one description is not redundant with the information contained in the other description for a given bitrate. The transfer of information back to the first decoder is carried out in a similar fashion. For a given bitrate for the accumulated syndrome bits, this cross-decoding is performed until the two descriptions are decoded or until the number of iterations reaches a certain threshold (the results shown in Section **??** were obtained for a threshold set to 18), in which case more accumulated syndrome bits are requested by the decoder.

## Reconstruction

Optimal inverse quantization where the noise $U$ has a Laplacian distribution with zero mean and variance $2/\alpha^2$ :

$$\hat{x}_{opt}=E[x|x\in\bigcup_{k=1}^{K}[z_i^d,z_{i+1}^k),y]=\left(\sum_{k=1}^{K}g(k)\right)\Big/\left(\sum_{k=1}^{K}h(k)\right)$$

where

$$g(k)=\begin{cases}\left(\frac{1}{\alpha}+z_i^k\right)e^{\alpha(y-z_i^k)}-\left(\frac{1}{\alpha}+z_{i+1}^k\right)e^{\alpha(y-z_{i+1}^k)} & \text{if } y<z_i^k,\\ \left(\frac{1}{\alpha}-z_i^k\right)e^{-\alpha(y-z_i^k)}-\left(\frac{1}{\alpha}+z_{i+1}^k\right)e^{-\alpha(z_{i+1}^k-y)}+2y & \text{if } y\in[z_i^k,z_{i+1}^k),\\ \left(\frac{1}{\alpha}-z_i^k\right)e^{\alpha(z_i^k-y)}-\left(\frac{1}{\alpha}-z_{i+1}^k\right)e^{\alpha(z_{i+1}^k-y)} & \text{if } y\geq z_{i+1}^k\end{cases}$$

and

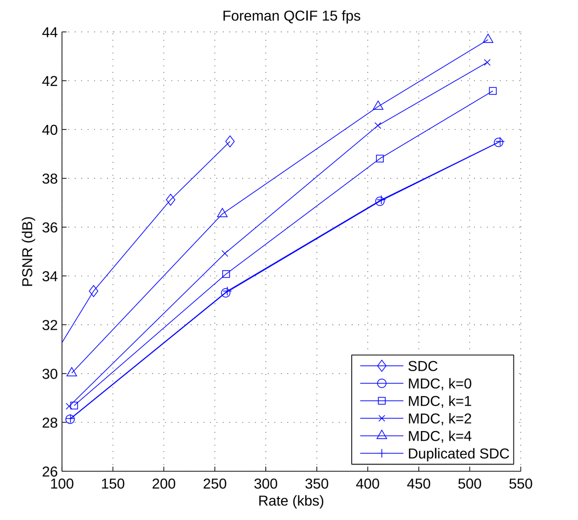
$$h(k)=\begin{cases}e^{\alpha(y-z_i^k)}-e^{\alpha(y-z_{i+1}^k)} & \text{if } y<z_i^k,\\ 2-e^{-\alpha(y-z_i^k)}-e^{-\alpha(z_{i+1}^k-y)} & \text{if } y\in[z_i^k,z_{i+1}^k),\\ e^{\alpha(z_{i+1}^k-y)}-e^{\alpha(z_i^k-y)} & \text{if } y\geq z_{i+1}^k.\end{cases}$$

The number $K$ of intervals of quantization for a given $x$ depends on the number of received descriptions and the number of diagonals in the index assignment matrix.

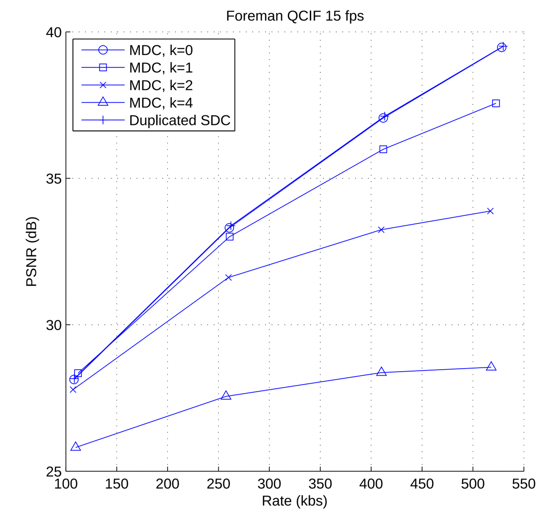
## Experimental results

- FOREMAN (QCIF format at 15 Hz)
- GOP size = 2
- bitrates for the key frames: 114, 266, 418 and 532 kbs
- intra codec based on JPEG2000 (VM 5.2)

### Key frames

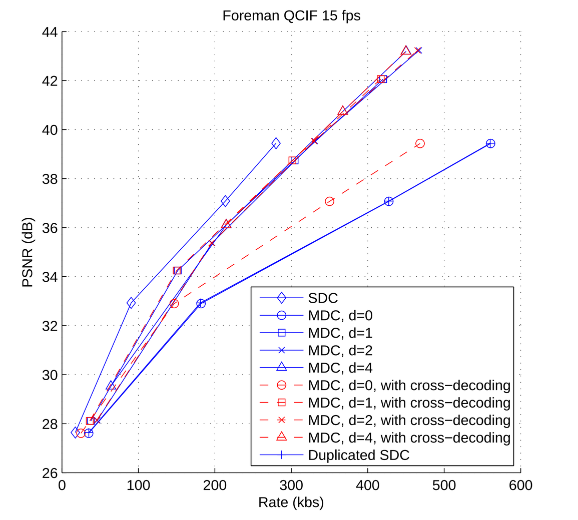



Central decoder

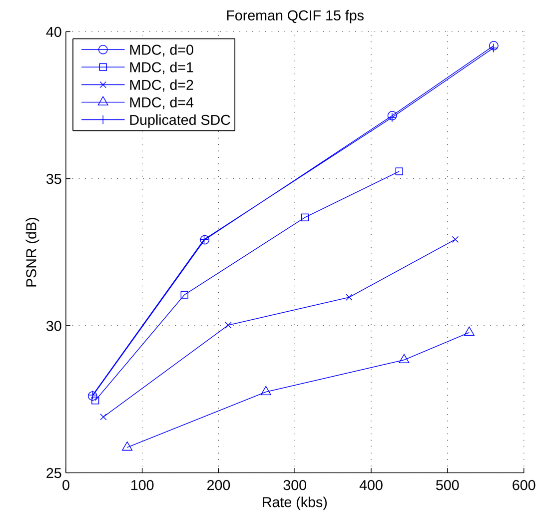



Side decoders

### Wyner-Ziv frames

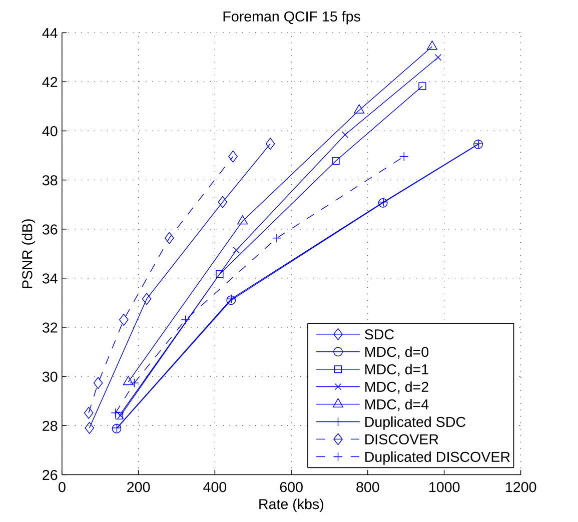



Central decoder

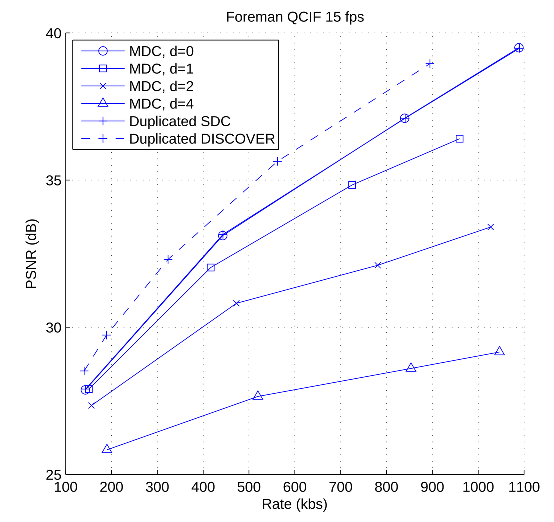



Side decoders

### All the frames




Central decoder




Side decoders

## Conclusion

- robust balanced multiple description video coding scheme with SI
- less redundancy leads to performance closed to the SDC scheme
- cross-decoding leads to a rate reduction when the correlation between the descriptions is maximized




Mail: <crave,pesquet>@telecom-paristech.fr, cguillem@irisa.fr