Multiple description coding of visual data

Gabriella Olmo

Politecnico di TORINO
Department of Electronics
Corso Duca degli Abruzzi 24
10129 Torino
Gabriella.olmo@polito.it
www.telematica.polito.it/sas-ipl/

Pre- and post-processing MDC for still images

(how to merge MDC and multimedia standards)

Introduction

- Previously described methods cannot be used jointly with co-decoding standard tools
- In fact, descriptions cannot be compliant JPEG or JPEG2000 streams
- This is a strong limitation to the use of MDC for encoding audio visual data
- In the following we will focus on still images, video will be discussed afterwards

Pre- and post-processing MDC for images

- Basic idea: to split the image into pixel subsets, or sub-images that:
  - Share some fundamental information about the original image (a certain degree of overlapping)
  - Share statistical properties with the original image, in order that ordinary compression tools can be used on the sub-images without (too much) efficiency loss
- The sub-images are encoded with the preferred encoder to generate the descriptions

Pre- and post-processing MDC

- At the decoder, each received description can be separately decoded with the standard tool
- In case of two description reception, the two decoded sub-images are post-processed in order to recover the full quality signal
- If a description gets lost, the decoded one is used to estimate the missing information

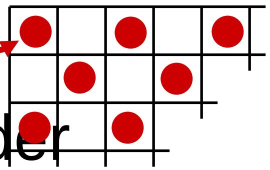
Tuning the redundancy

- In order to introduce an amount of controlled redundancy between the descriptions, so achieving the desired tradeoff between the two quality levels, the original image must be properly processed prior to splitting

1. Over sampling the original image prior to the description generation is a means to improve the correlation characteristics
2. Performing an overlapped image splitting so as to minimize the prediction error of the lost data is another approach

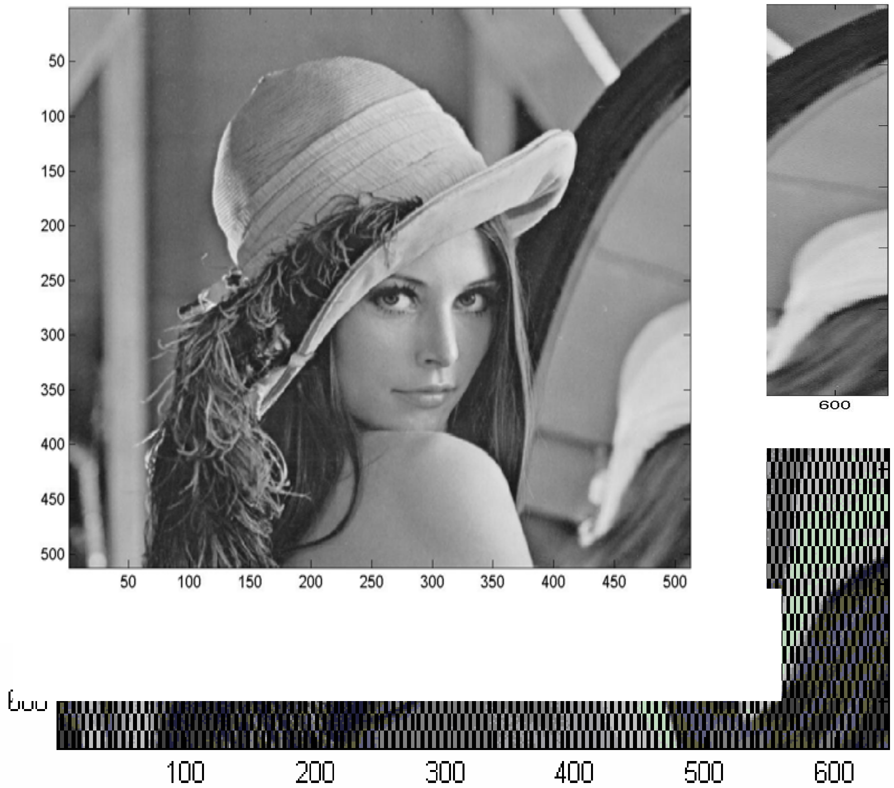
Spatial oversampling

- Basic idea: controlling the redundancy by over sampling the image prior to splitting
  - zero padding techniques in the frequency domain
  - quincunx or “directional” sub-sampling to create sub-images



  - Sub-images are encoded by a standard encoder
  - In case of one description reception the lost sub-image is estimated by linear first order estimator from the neighbors pixels

- **How does it work?**



"Multiple description image coding using Pre- and Post-processing" S.Shirani, M.Gallant, and F. Kossentini,

2D over sampling

- When the post-processing stage is not available, the decoder will not crash and is still able to output a valid image, which in turn will exhibit distortion
- This method simply introduces redundancy in a uniform fashion, leading to:
  - Suboptimal performance
  - Large computational complexity due to 2D operations

1D over sampling

- In order to introduce the desired correlation, a directional zero-padding scheme can be used
  - Direction of expansion is that which exhibits the best correlation properties. In other words, the direction along which the image shows a smoother behavior.
- In order to minimize deformation, sub-images are created by taking the vectors of data in a direction perpendicular to that of expansion

1D over sampling

Original image

DCT transform

Zero padding

IDCT transform

Down sampling

Encoding subimage

Decoded subimage

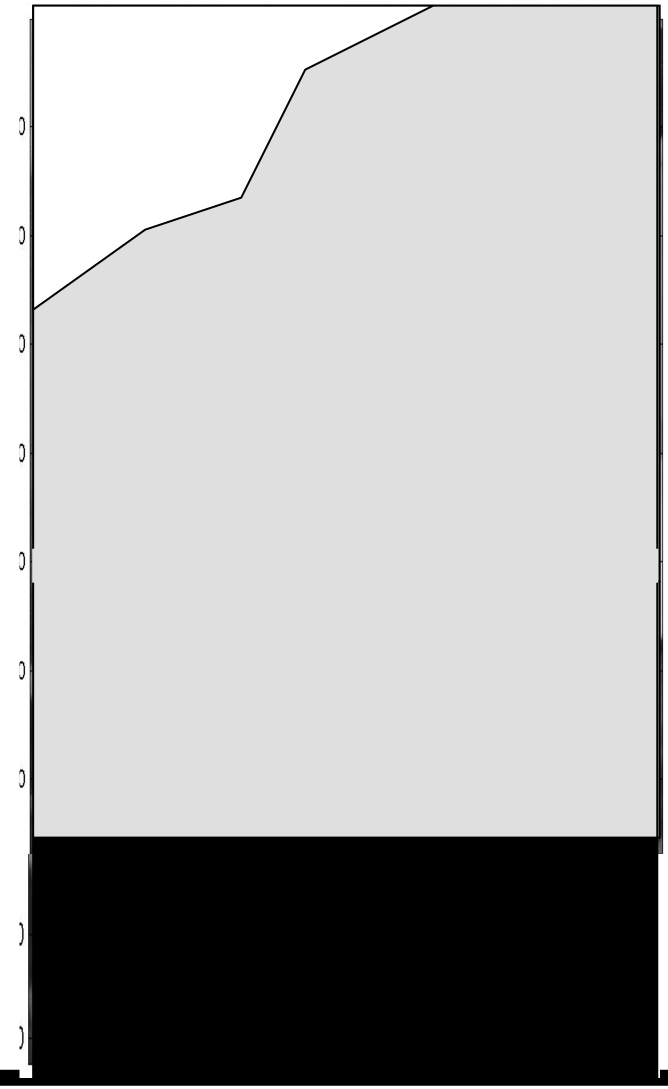
Up sampling

Estimation

DCT transform

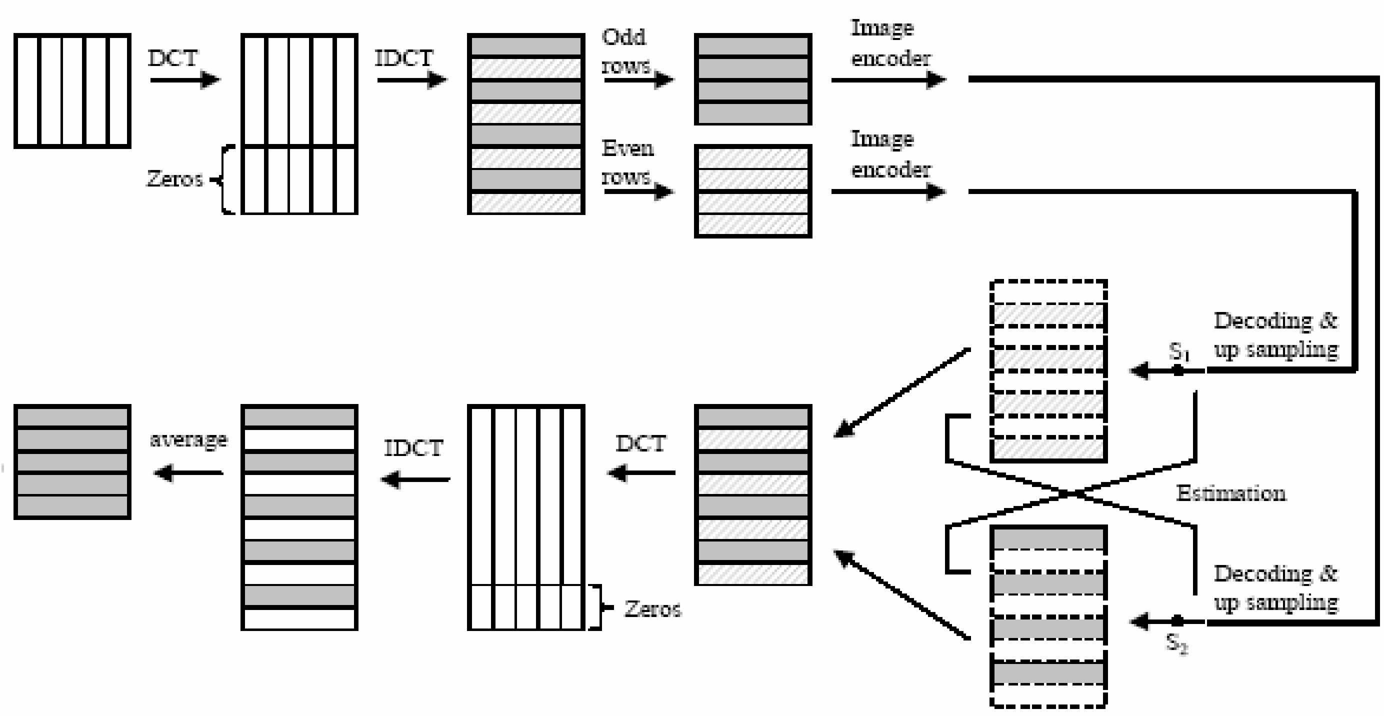
Zero padding

IDCT transform



PPP MDC – 1D over sampling

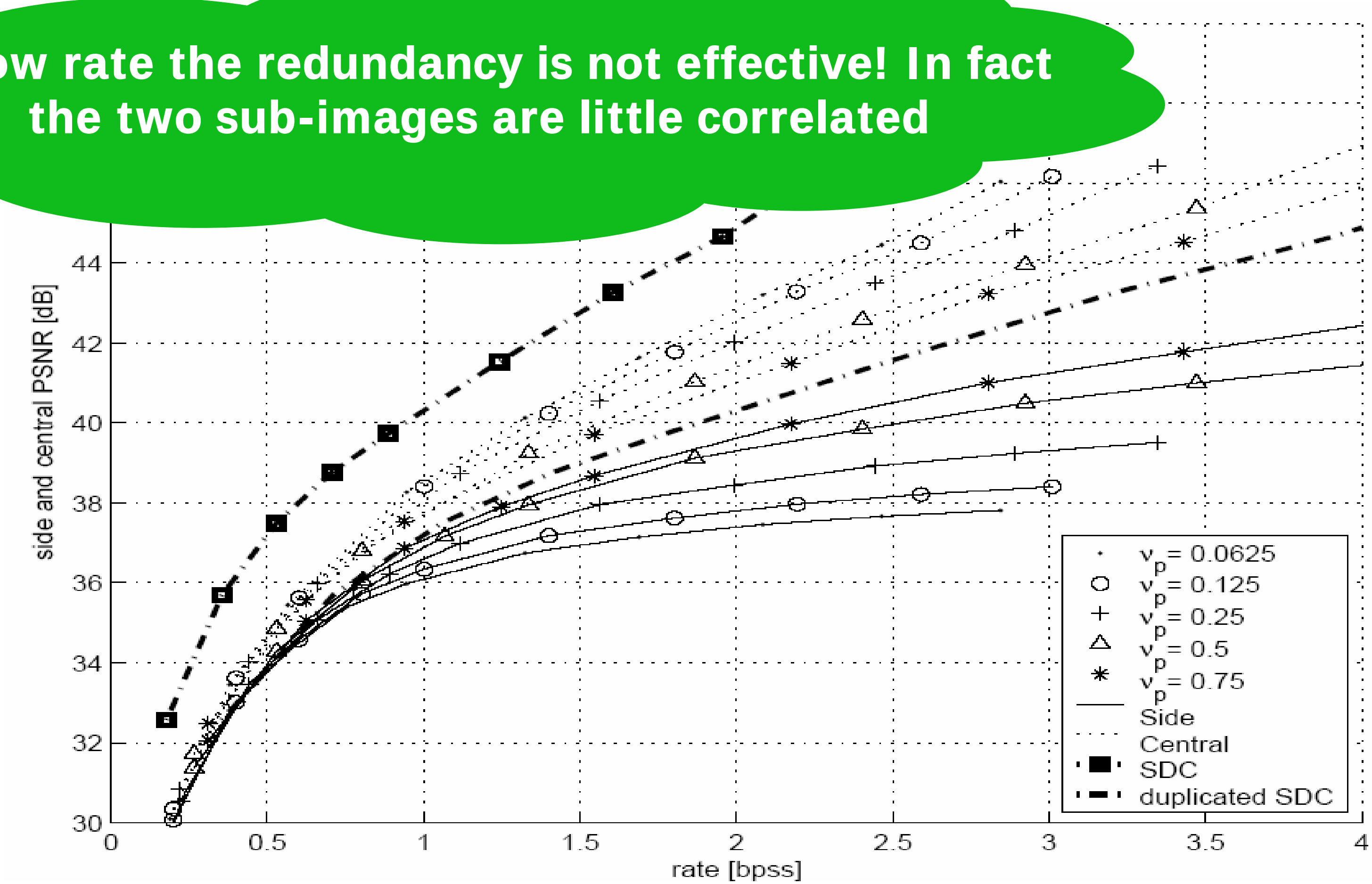
- **How does it work?**



"Directional multiple description scheme for still images":G.Olmo; T.Tillo

1D over sampling

At low rate the redundancy is not effective! In fact the two sub-images are little correlated



Performance of directional PPP MDC for image *Lenna*

PPP MDC based on pixels predictability

- Basic idea: to split the image so as to minimize the prediction error of lost description
- Encoder selects the more predicable direction
- Two sub images generated:
  - A tunable amount of pixels are present on both sub-images
  - In order to determine which pixels need to be replicated, the encoder emulates the estimation stage in the side decoder → pixels that minimize distortion are duplicated

- **How does it work?**

Original image

LPV selection

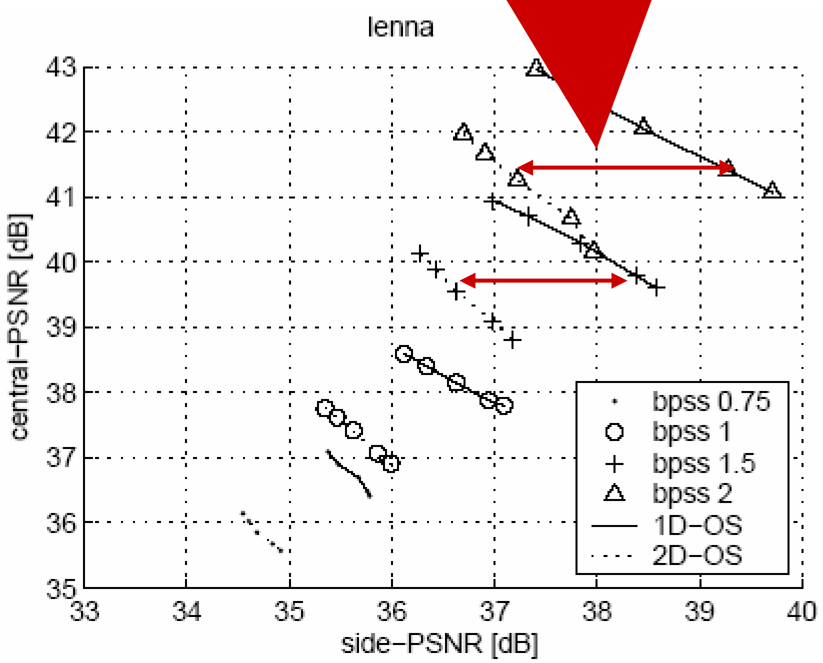
Image to be encoded

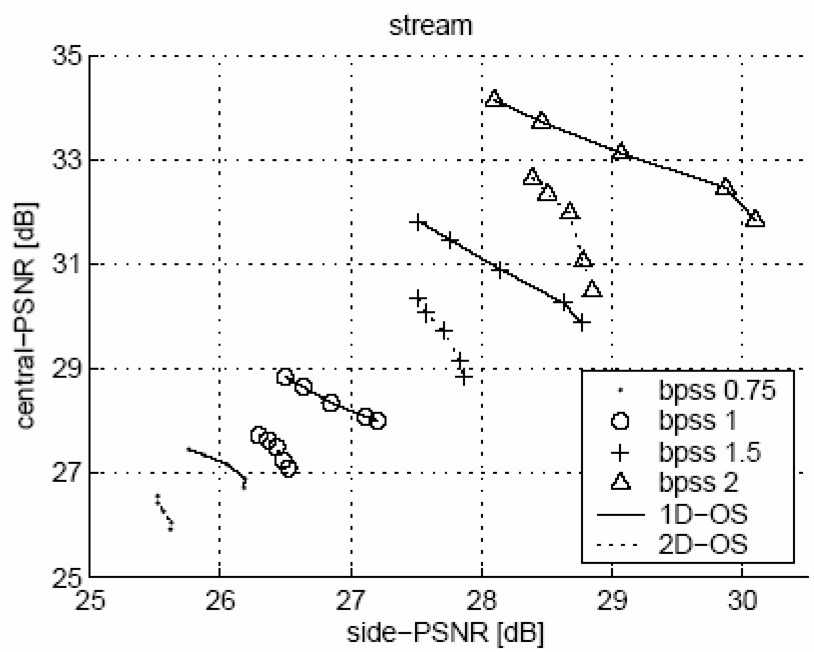
The position of the inserted extra rows/columns is communicated to the decoder by means of binary string

- **Decoder**
  - Single description reception case:
    - The received description is decoded.
    - Sub-image is upsampled according to the under-sampling pattern.
    - Lost pixels estimated from neighbors.
  - Two description reception case:
    - Both received description are decoded.
    - Sub-images are merged according to their under-sampling patterns
    - Duplicated pixels are averaged

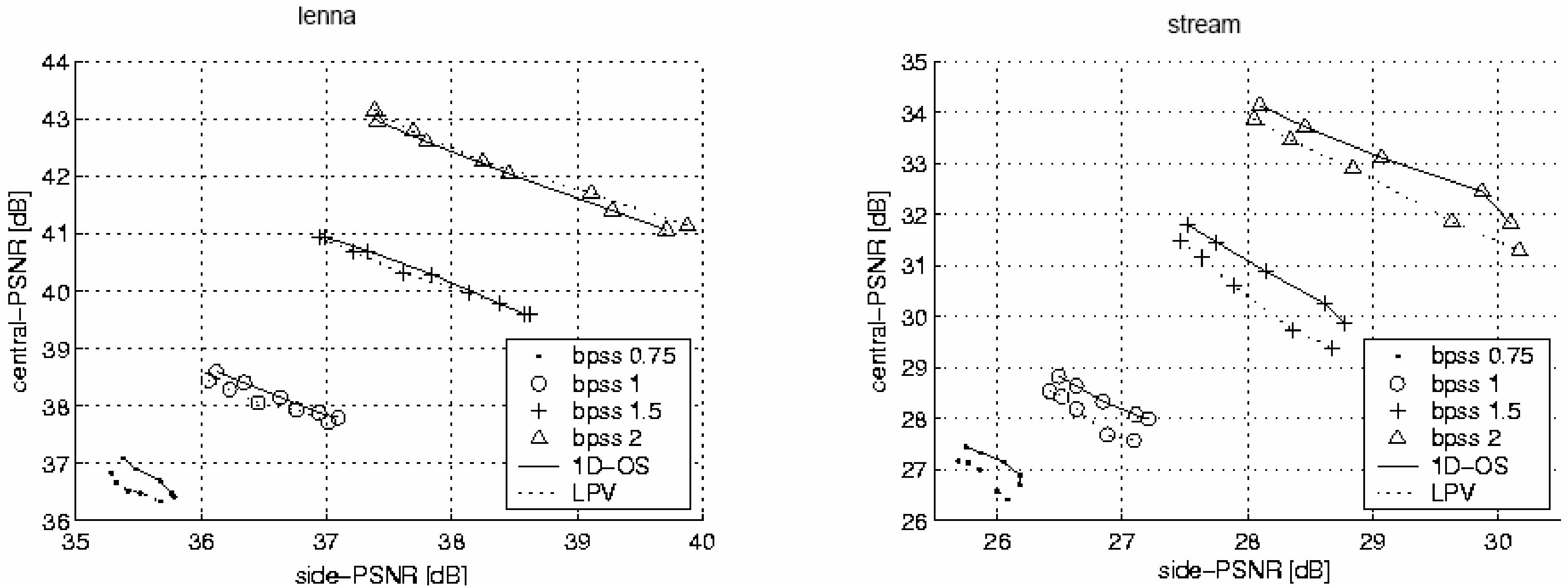
PPP MDC performance

1D exhibits better performance with less complexity





PPP MDC performance



All the spatial based PPP MDC methods turn out to be ineffective in the low bit rate region

MD for post-compression based rate allocation

(exploiting what some standards make available)

JPEG 2000 rate allocation

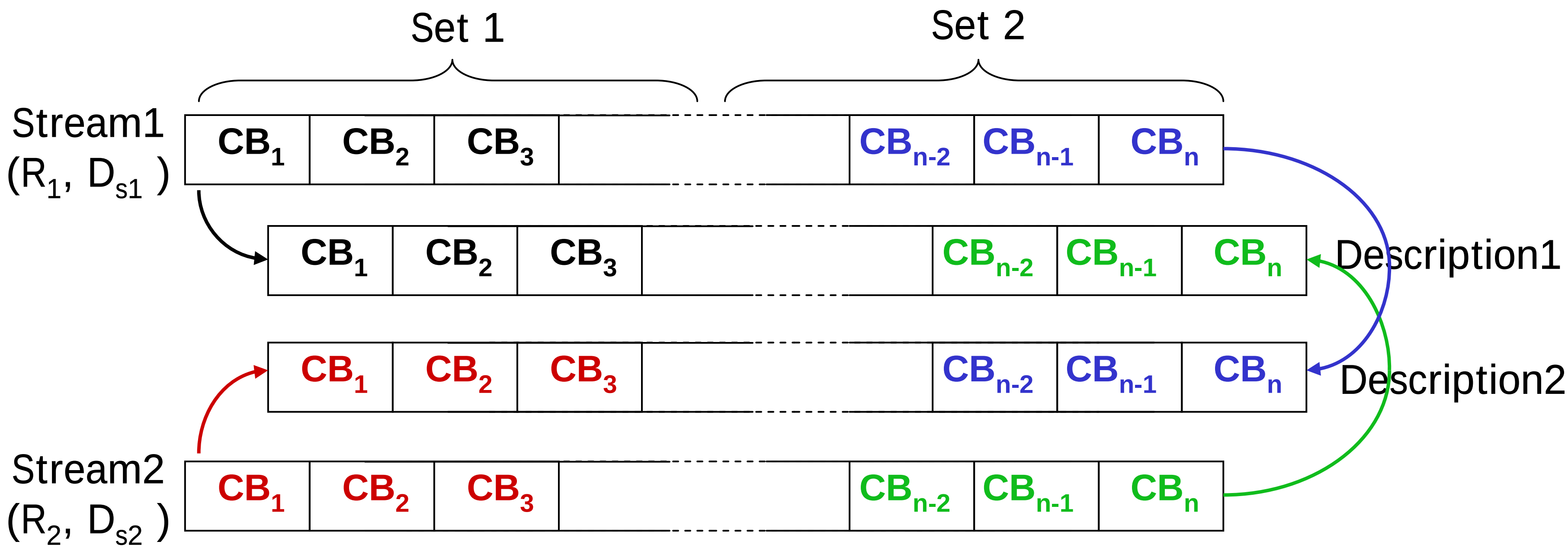
- Coefficients are grouped in *code-blocks (CBs)*
- A distortion measure D_i and a rate R_i are associated to each CB B_i
- An additive distortion measure is assumed
- All CBs are independently compressed and for each of them the rate-distortion curve is computed
- A post-compression stage determines the rate R_i to which each CB should be compressed so as to minimize D for a given rate R

MD for post-compression based rate allocation

- **Basic idea**
  - Generate two streams at different rates
  - Mix the CBs of the two streams in order to get balanced descriptions
- **Encoder:**
  - Generates two streams at rate $R_2 < R_1$
  - The CBs are grouped into two sets having *similar* RD characteristics
  - The first description is obtained by properly combining the CBs of set 1 taken from the first stream, with the CBs of set 2 taken from the second one

MD for post-compression based rate allocation

$$\forall\ CB_i \in Set1 \Rightarrow \exists\ CB_k \in Set2 : D_i(r) \approx D_k(r)$$



MD for post-compression based rate allocation

- **Central decoder:**
  - Both descriptions are pre-processed and merged in a single stream, where, for each CB, the best representation is selected
  - The resulting stream is then decoded.
    - Central distortion is $\mathbf{D_{s1}}$
- **Side decoders:**
  - The received description is simply decoded by the used tool
    - The average side distortion is $\mathbf{(D_{s1} + D_{s2})/2}$
  - This modality can also be adopted in case the receiver is not equipped with an enhanced decoder.

Redundancy tuning?

- **Redundancy = R_2!!**

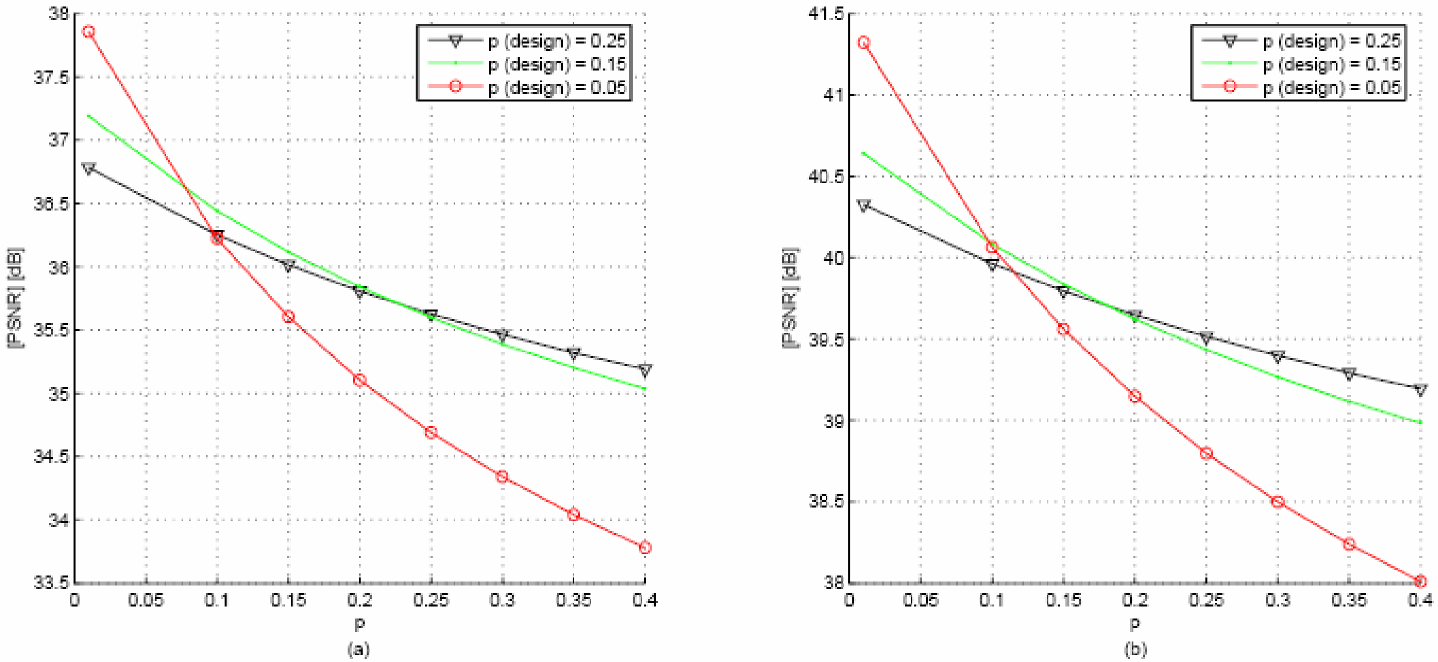
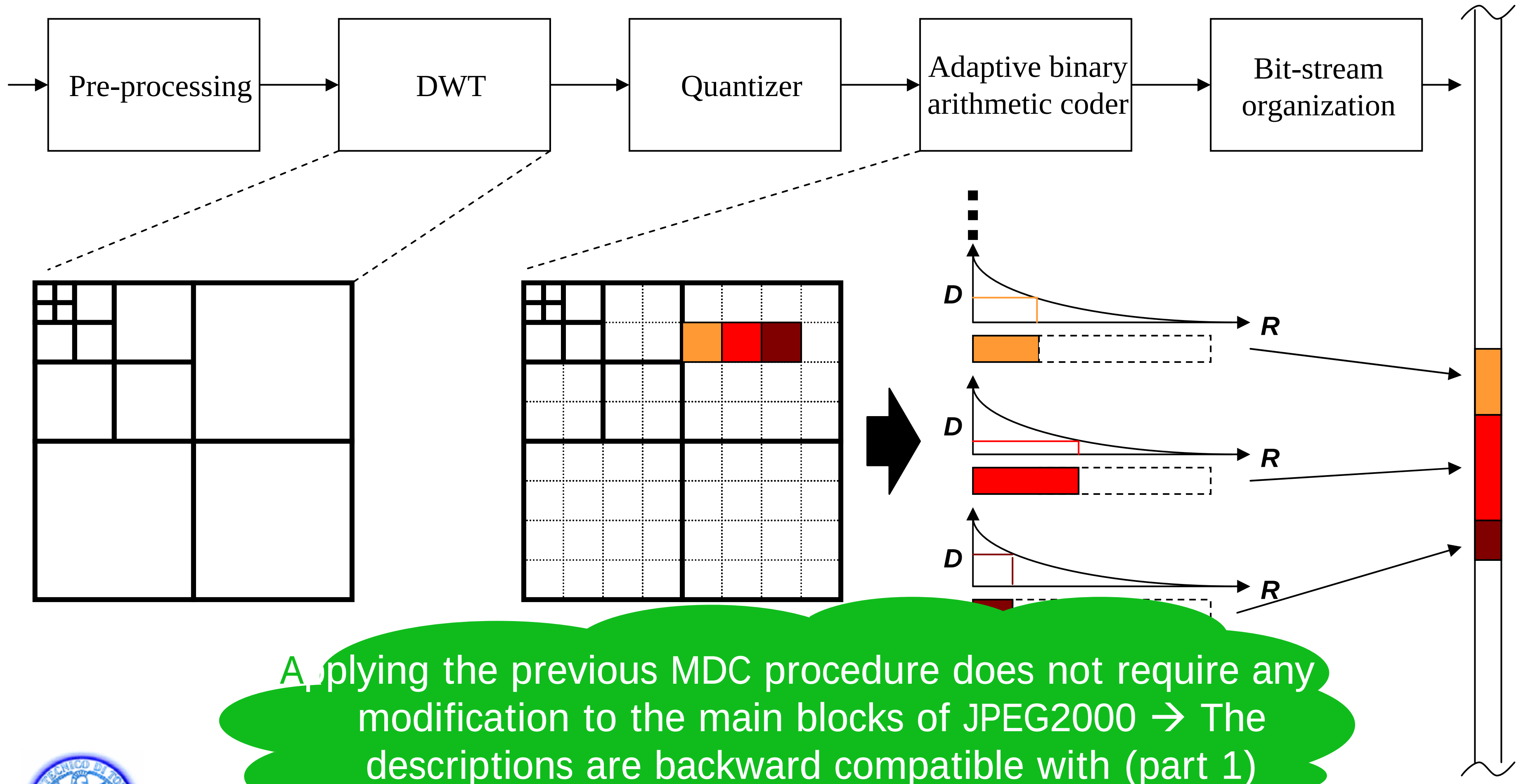


Fig. 6. The expected end-to-end PSNR versus the probability of description loss for image: a) Goldhill; b) Lenna; with design parameter $p = 0.05, 0.15$ and 0.25.

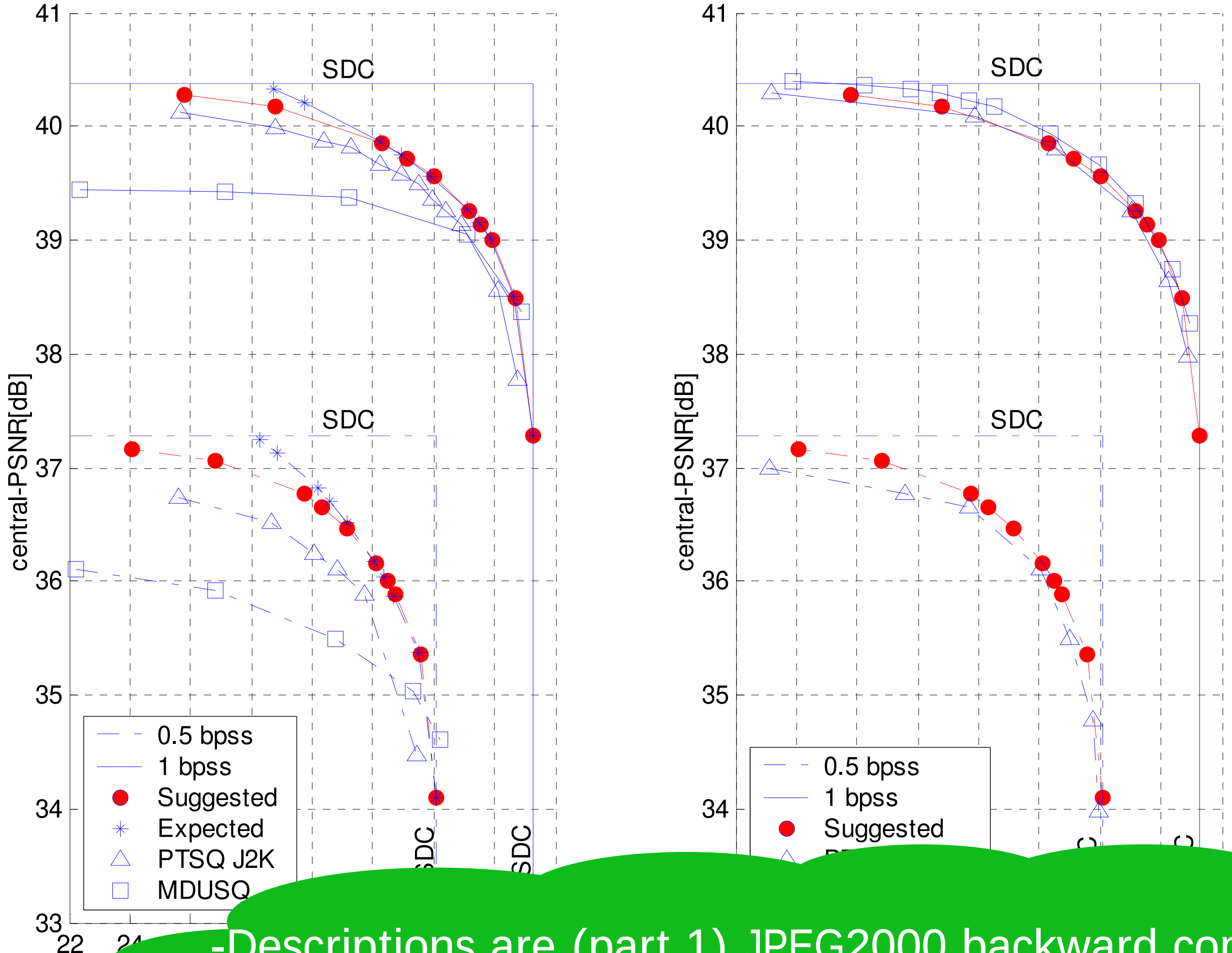
MD generation exploiting the JPEG2000 core



Applying the previous MDC procedure does not require any modification to the main blocks of JPEG2000 → The descriptions are backward compatible with (part 1) JPEG2000 standard

MD generation exploiting the JPEG2000 core

"A novel multiple description coding scheme compatible with the JPEG200 decoder":T.Tillo; G.Olmo



-Descriptions are (part 1) JPEG2000 backward compatible.

-Redundancy can be easily controlled → match the network

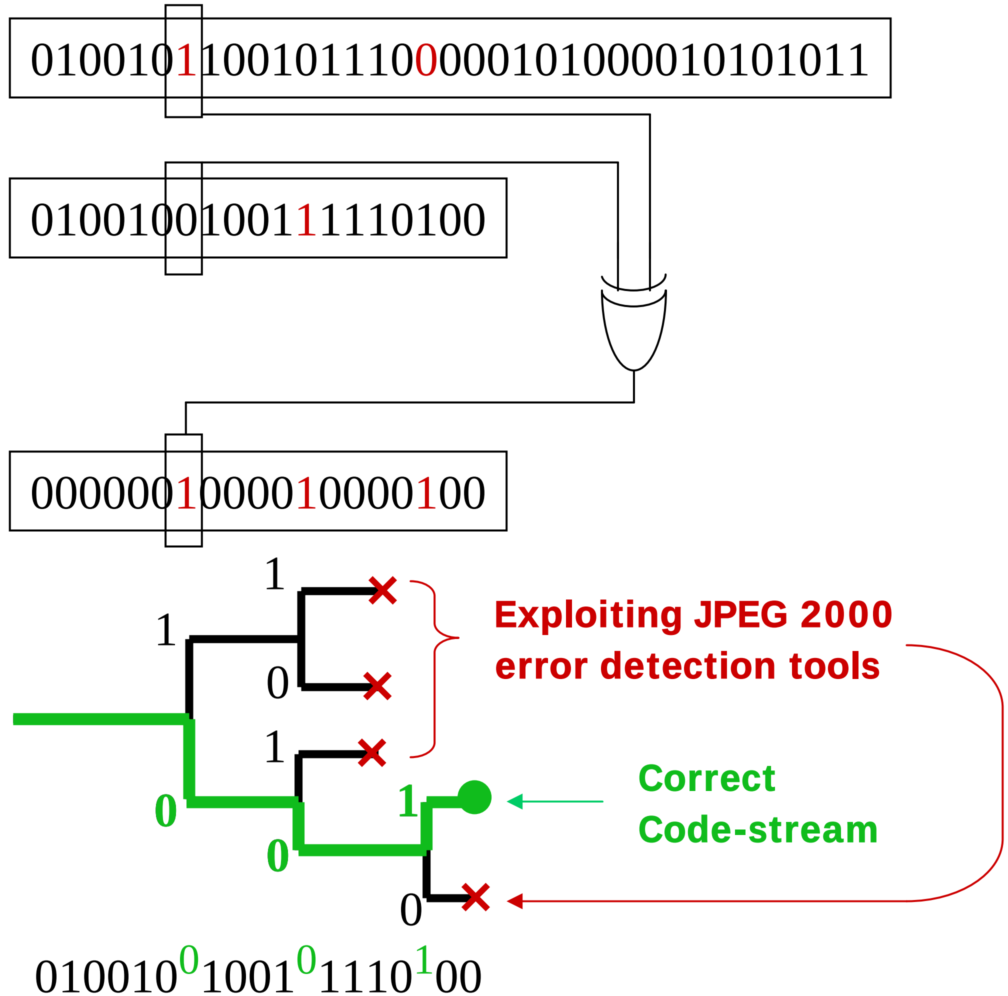
- **Preliminary**
  - In JPEG2000, the CB stream encoded at low bit rate is completely embedded in a higher rate one.
  - As a consequence, the bit-stream of an image encoded at rate $R_2 < R_1$ is entirely embedded in that encoded at rate R_1
  - This property can be exploited in order to identify the locations of bit errors in the shortest stream representing the code-block.
  - The decoder can exploit standard error resilience tools offered by JPEG 2000 in order to recognize the correct error pattern

stream of finer
representation of code-block

stream of coarser
representation of code-block

result of ex-oring

arithmetic decoder branching



Resilient MDC for JPEG2000

"flexible error resilient scheme for jpeg 2000":T.Tillo; M.Grangetto; G.Olmo

Results on image Woman (2560*2048 pixel) for a binary symmetric channel.

BER	MDC algorithm		UEP RS on 5 layers *
	2 descr. received	1 descr. received	
Error free	36.04	29.61	36.06
10^{-4}	35.17	26.64	35.05
10^{-3}	31.10	19.61	26.91

+4dB

This approach can guarantee a base quality, delivered by the bitstream encoded at the lower rate R_2 , in case one description is received error free or both streams are received affected by bit errors.

* "Unequal protection of JPEG2000 code-stream in wireless channels": A. Natu; D. Taubman,

Resilient MDC for Motion JPEG2000

"Multiple Description Coding with error correction capabilities: an application to Motion JPEG 2000":T.Tillo; M.Grangetto; G.Olmo

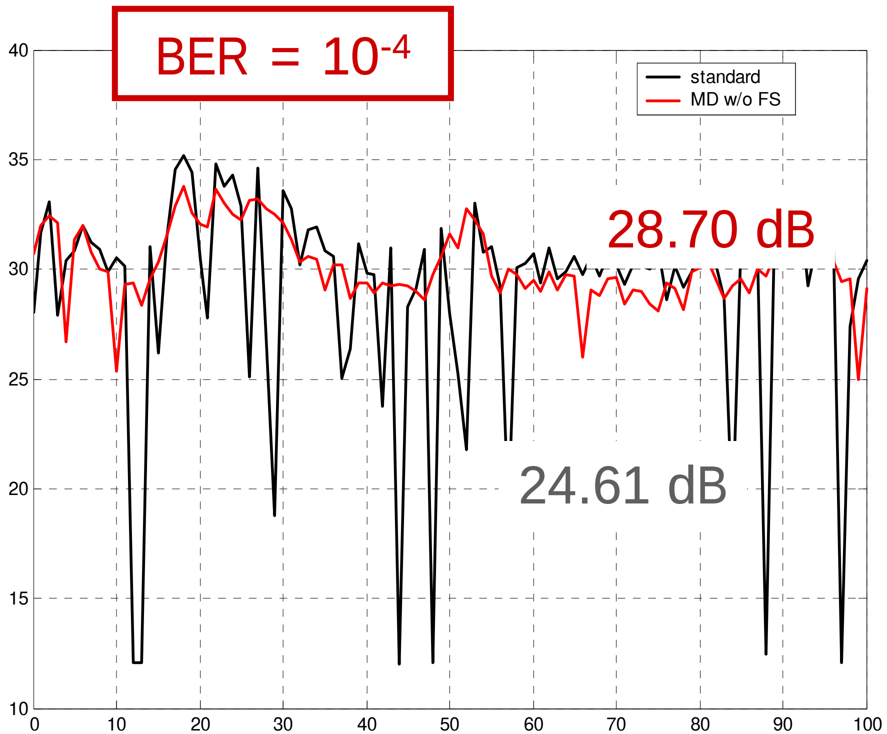
- CIF-YUV 4:2:0 format, 15 fps.
- Overall bit rate 700 kbps (0.5 bpp per frame)
- 4 resolution levels, 32X32 code blocks, ρ=0.2
- Headers assumed to be protected by a FEC, $R_c = 0.5$
- BSC, BER = 10^{-4}, a maximum of 64 branches allowed → maximum 6 errors per coding-pass can be concealed

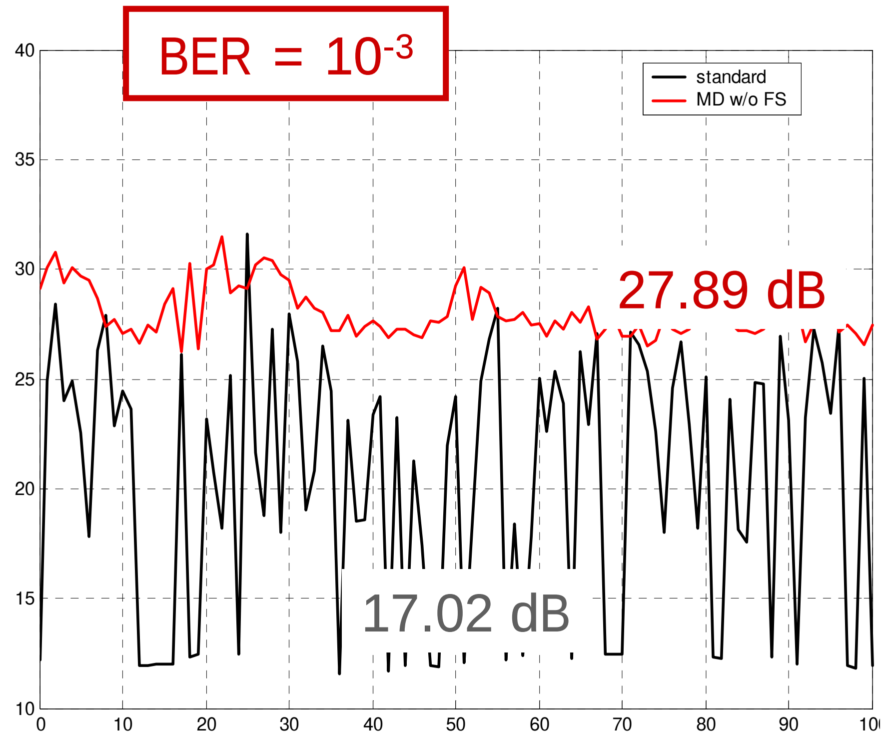
Standard decoder

MD decoder

Resilient MDC for Motion JPEG200

- PSNR of luminance component
- The average PSNR values are computed as the linear average of MSE over 1000 frames.
- $\rho = 0.4$

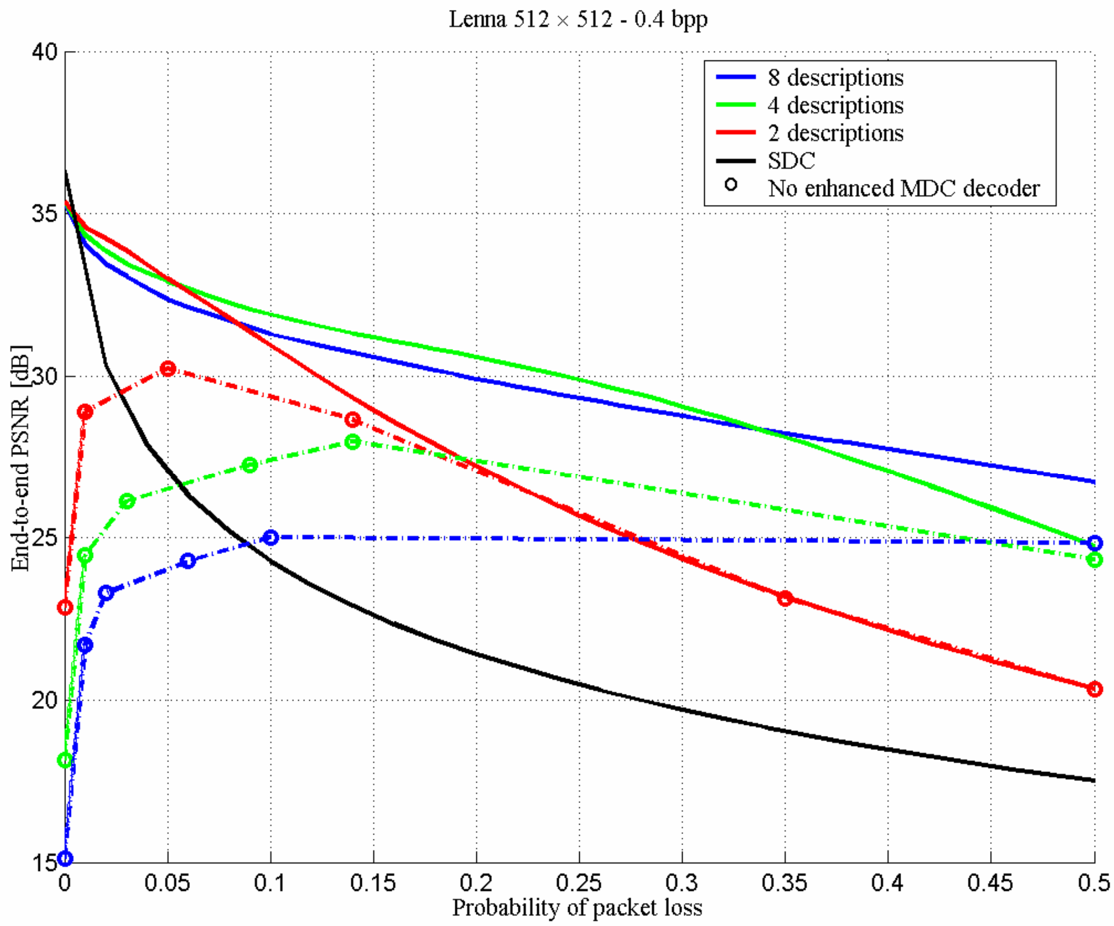




Multi-rate allocation for N balanced MDC

- The R-D based scheme can be easily extended to the case of N > 2 descriptions
- Starting from 2 or more streams encoded at different bit rates, strategies can be conceived in order to get N balanced descriptions

Multi-rate allocation for n balanced MDC



Multi-rate allocation for n balanced MDC

(a) (b) (c) (d)

Fig. 3. The obtained image by decoding: a) one; b) two; c) three and d) four descriptions ($N = 4$).

References

1. S.Shirani, M.Gallant, and F. Kossentini, “Multiple description image coding using Pre- and Post-processing,” *IEEE Intl. Conf. On Information Technology: Coding and Computing*, 2-4 April 2001
2. T. Tillo, E. Baccaglini, G. Olmo, “A flexible multi-rate allocation scheme for balanced multiple description coding applications,” IEEE Intl. Workshop on Multimedia Signal Processing, Shanghai, Nov. 2005
3. E. Baccaglini, T. Tillo, G. Olmo, “Network adaptive Multiple Description Coding for JPEG2000,” ICIP 2005 - IEEE ICIP, Genova, Sept. 2005
4. G. Olmo, T. Tillo,”A novel multiple description coding scheme compatible with the JPEG 2000 decoder,” *IEEE Signal Processing Letters,* Vol. 11, No. 11, pp. 908-911, November 2004
5. T. Tillo, M. Grangetto, G. Olmo,”Multiple Description Coding with error correction capabilities: an application to Motion JPEG 2000,” ICIP 2004, Singapore, October 2004
6. T. Tillo, M. Grangetto, G. Olmo, ”A flexible error resilient scheme for JPEG 2000,” 2004 IEEE International Workshop on Multimedia Signal Processing, September 2004
7. T. Tillo, G. Olmo, ”A low complexity pre-post processing multiple description coding for video streaming,” IEEE ICTTA 04, Damascus, April 19-23, 2004
8. G. Olmo, T. Tillo, “Directional multiple description scheme for still images,” IEEE 10th Int. Conference on Electronics, Circuits and Systems, Sharjah (UAE), December 14-17, 2003
9. G. Olmo, T. Tillo,”ROI-preserving multiple description coding based on pre- and post-processing to standard encoders,” Sixth Baiona Workshop on Signal Processing in Communications Baiona (Spain), September 8-10, 2003

MDC for motion compensated video applications

(where things get really challenging)

MDC for video delivery

- Appealing for real time interactive applications for which retransmission is often not feasible
- It simplifies the network design (no feedback or retransmission necessary; all packets treated the same)
- But MDC for motion compensated video is more than just applying an MD image coder to each frame!!!

MDC for motion compensated video

- In MC prediction, whenever the encoder uses a signal for prediction that is unavailable at the decoder side, a mismatch occurs, which propagates until next I frame
- The most efficient prediction would be formed from data from both channels
- If separate prediction is used for each channel, no mismatch occurs but prediction is less efficient
- A primary challenge is to balance mismatch control and prediction efficiency: avoiding mismatch is not always convenient

Classification of MD predictors

Predictor class	Definition
A	**No mismatch.** **The prediction uses the same signals known at the decoder, independently of its state.** **Example: Two individual predictions (two-state encoder)**
B	**Use same prediction used by reference SDC scheme. Prediction error minimized; no redundancy added in the prediction process. Mismatch will occur. Optional: to send a compressed version of the mismatch**
C	**It has parameters that can be selected to trade the loss in prediction efficiency and the amount of mismatch. Optional: to send a compressed version of the mismatch**

Types of redundancy

Type	Cause of redundancy	Symbol
(a)	**Coding the prediction error signal(s) using MD techniques (e.g. correlating transform, MDSQ)**	ρ_a
(b)	**Using a predictor that is less efficient than SD**	ρ_b
(c)	**Sending explicit signal for mismatch reduction**	ρ_c
(d)	**Sending side information (e.g. duplicated motion vectors)**	ρ_d

$$R = R^* + \rho_a + \rho_b + \rho_c + \rho_d$$

Video redundancy coding (Wenger)

- The sequence is temporally downsampled into two subsets (every other frame)
- Each subset is encoded using a SD coder
- No mismatch occurs (Class A predictor)
- Side decoder: missing frames are interpolated form received ones
- Decreased correlation between adjacent frames → redundancy governed by the source, not controllable

MD transform coding – MDTC (Reibman et al.)

- **Version with no mismatch control**: A single prediction is formed using a SD predictor with information available at the central decoder (Class B)
- The prediction error is DCT transformed and pairs of DCT coefficients have correlation introduced by means of correlating transform.
- ρ_a is controlled by adjusting the transforms applied to each coefficient pair. Larger ρ_a decreases the side distortion.
- Motion vectors are duplicated so that ρ_d is non zero. I frames are periodically inserted

MD transform coding – MDTC

- **Version with mismatch control**: Three prediction loops are formed, to mimic the possible states of the decoder
- The central prediction error is encoded as before (Class B)
- The difference between central and side prediction is encoded in the descriptions.
- The quantization parameter of such a side information (which is essentially ρ_c redundancy) must be carefully selected

Multiple description motion compensation (Wang-Lin et al.)

- Three prediction loops
- Central loop estimates frame n from frames n-1 and n-2, with motion vectors MV1 (between n and n-1) and MV2 (between n and n-2)
- Side loops estimate frame n from frames n-i, i=1, 2 respectively
- The mismatch is added to each description along with MVi
- Class C predictor as mismatch and efficiency can be traded off by means of the predictor details

Independent flow MD video coding IF-MDVC (Franchi et al.)

- Polyphase downsampling of each video frame and separate predictive encoding of each one
- Each frame is subject to oversampling by means of 2D DCT/zero padding in order to augment the data correlation prior to splitting.
- The MD generation is external to the prediction loop: no mismatch occurs (Class A).
- In case of N>2 descriptions, the advantage of having no mismatch control signals is larger than the efficiency loss of having different prediction loops.

Using MD video coding on a lossy packet network

- Information is partitioned in multiple packets each of which has probability p of being lost
- If one description fits one packet, the usual assumption of ideal MD network is still valid (description either perfectly received or lost)
- If the descriptions are packetized in L packets, a description may be partially lost, and losses may appear in both descriptions
- $2^L - 1$ possible states at the receiver; mismatch control is challenging
- Mismatch coding infeasible; strategies include insertion of I-frames, sync-frames, SP-frames

Conclusions

- MDC for MC video is challenging due to the mismatch problem
- In packet networks, descriptions may be partially received, and this represents a novelty with respect to classical MD paradigm
- Many proposals, but no performance comparison possible due to different test conditions!

A LOT OF WORK TO BE DONE!!!

References

1. Y. Wang , A. R. Reibman, S. Lin, “Multiple description coding for video delivery,” Proceedings of the IEEE, vol. 93, no. 1, pp. 57-68, Jan. 2005
2. S. Wenger, “Video redundancy coding in H.263+,”Audio-Visual services over packet networks conf., Aberdeen, U.K., 1997
3. A. Reibman et al., “Multiple description coding of of video using motion compensated prediction,” IEEE Trans. Circuits Systems Video Technol., vol. 12, no.3, pp. 193-204, March 2002
4. Y. Wang, S. Lin, “Error resilient video coding using multiple description motion compensation,” IEEE Trans. Circuits Systems Video Technol., vol. 12, no. 6, pp. 438-452, June 2002
5. N. Franchi, M. Fumagalli, R. Lancini, S. Tubaro, “Multiple description video coding for scalable and robust transmission over IP,” IEEE Trans. Circuits Systems Video Technol., vol. 15, no. 3, pp. 321-334, March 2005

THANK YOU

Contact information:

- Gabriella Olmo – Department of Electronics, Politecnico di Torino – Ph. +39 011 5644094 – email gabriella.olmo@polito.it
- Tammam Tillo - Department of Electronics, Politecnico di Torino/ Istituto Superiore "Mario Boella" – Ph. +39 011 2271210 – email tammam.tillo@polito.it