TELECOM ParisTech
Institut Mines-Télécom

Coding for the Network: Scalable and Multiple description coding

Marco Cagnazzo

Overview

- Examples and motivations
- Scalable coding for network transmission
- Techniques for multiple description coding

Coding for the Network

- **Traditional compression techniques:**

Minimum size for a given quality

- **Assumptions:**
  - Known bandwidth
  - Known decoding parameters (resolution, frame-rate)
  - No errors
- **Network video delivery**
  - Different links and users
  - Unreliable links

Coding for the Network

- **Basic idea: from a monolithic representation to a structured representation**

- **Layered (i.e. hierarchical) sub-streams: scalable video coding (SVC)**

- **Equal importance sub-streams: multiple description coding (MDC)**

Image transmission

1. **Sequential coding**
2. **Progressive coding (scalable)**
3. **Multiple description (MDC)**

- **5 packets**
- **2 cases:**
  - no error
  - packet 2/5 not received

Examples: transmission without errors

Packet 1

Non-scalable

Scalable

Multiple descriptions

Examples: transmission without errors

Packet 2



Non-scalable

Scalable

Multiple descriptions

Examples: transmission without errors

Packet 3

Non-scalable

Scalable

Multiple descriptions

Examples: transmission without errors

Packet 4

Non-scalable

Scalable

Multiple descriptions

Examples: transmission without errors

Packet 5

Non-scalable

Scalable

Multiple descriptions

Examples: transmission with errors

Packet 1

Non-scalable

Scalable

Multiple descriptions

Examples: transmission with errors

Packet 2: **LOSS**

Non-scalable

Scalable

Multiple descriptions

Examples: transmission with errors

Packet 3

Non-scalable

Scalable

Multiple descriptions

Examples: transmission with errors

Packet 4

Non-scalable

Scalable

Multiple descriptions

Examples: transmission with errors

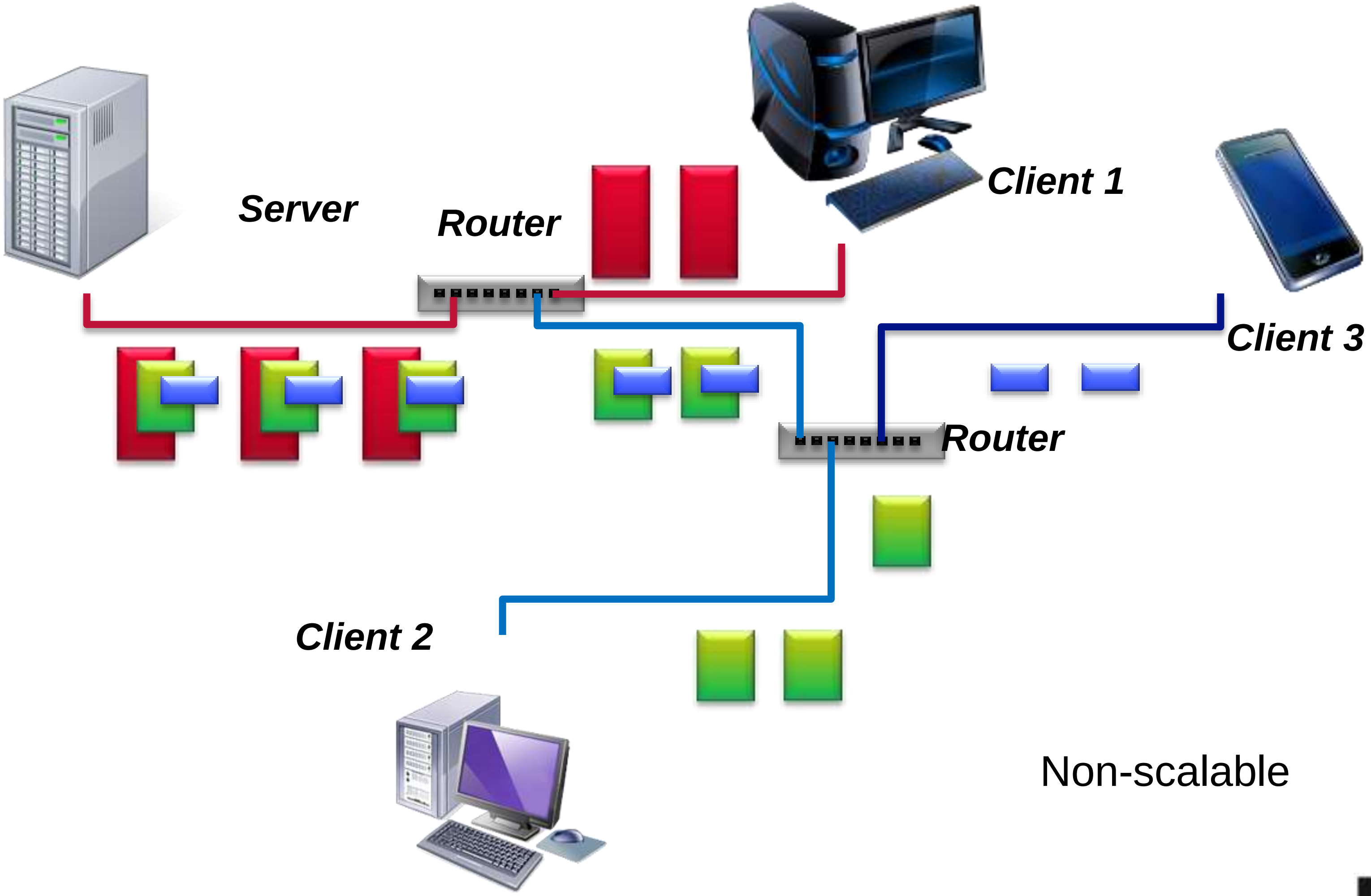
Packet 5

Non-scalable

Scalable

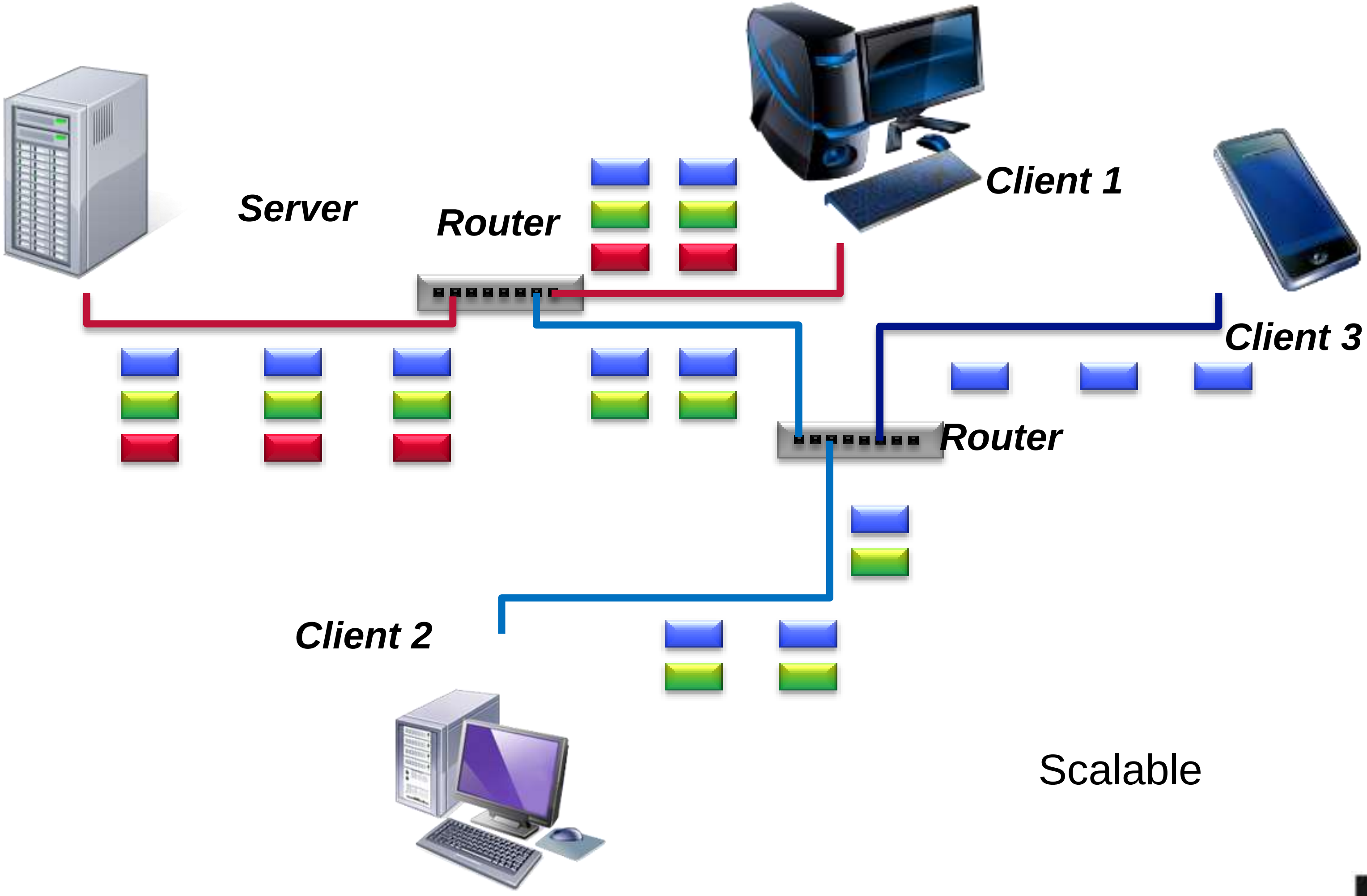
Multiple descriptions

Examples: heterogeneous users and networks



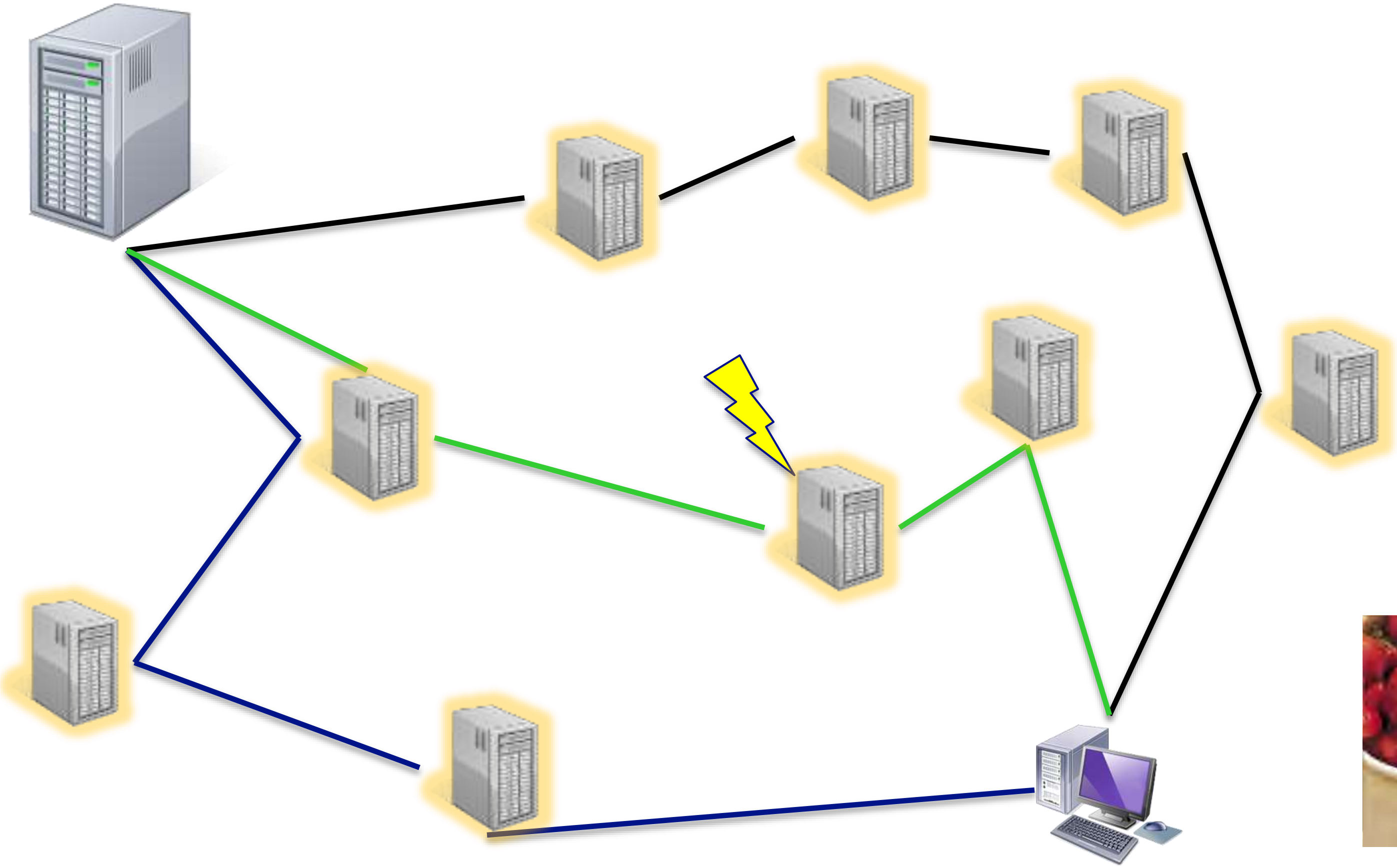
Non-scalable

Examples: efficient network use

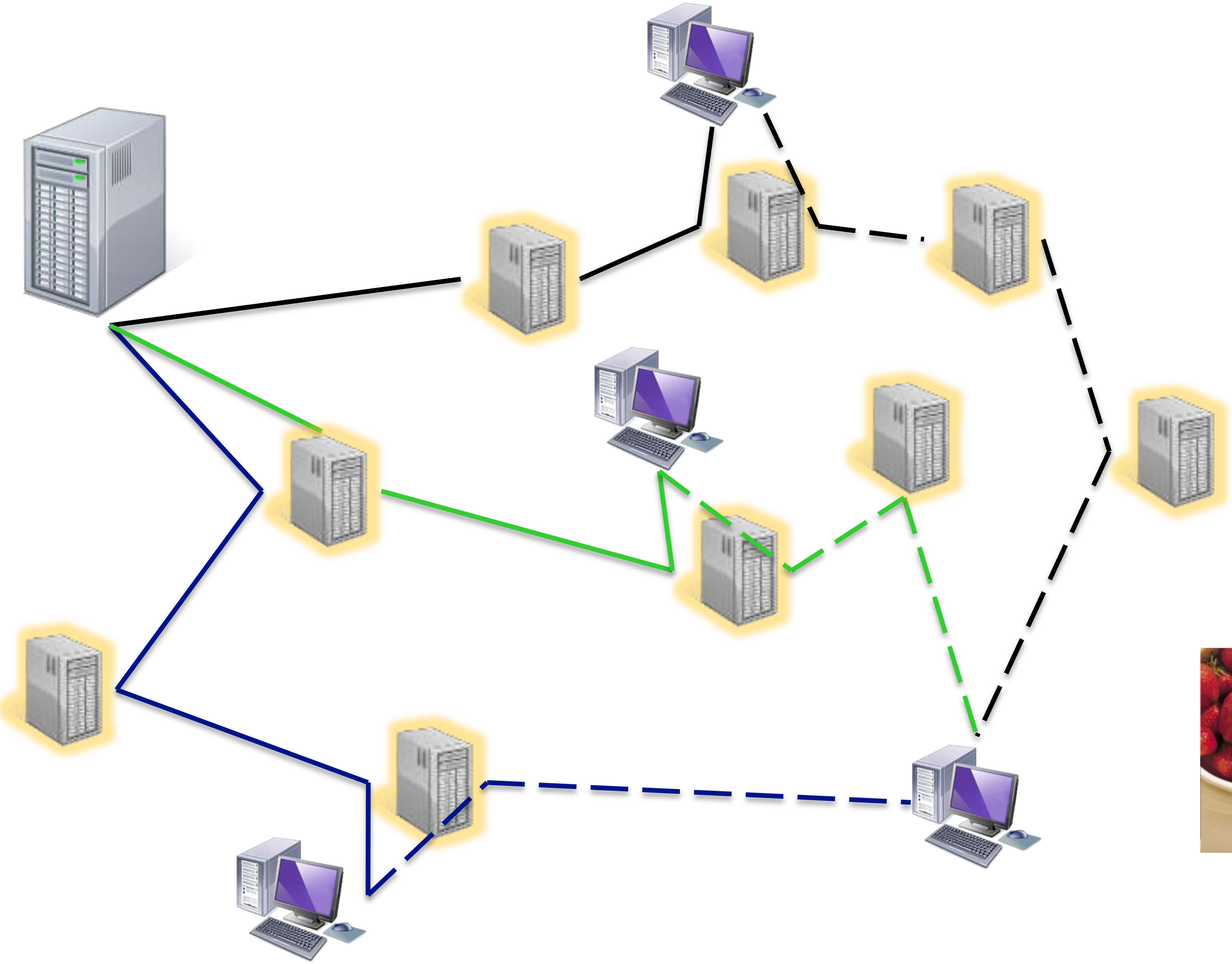


Scalable

Multiple Description Coding



Multiple Description Coding + P2P



Overview

- **Examples and motivations**
- **Scalable coding for network transmission**
  - Principles and applications
- **Techniques for multiple description coding**

SVC approach

- **Encode once**
  - Different versions obtained by combining sub-streams
- **Intelligent network (multiple multicast groups)**
  - Packet replication only where needed
  - Efficient network use

Scalability in video coding

Definition: A scalable video stream is a compressed representation of a video such that:

- The representation is made up of *layers*
- Layers provide *incremental* refinement of the sequence
- The representation is *efficient* in term of reconstructed image quality for a given rate

Kinds of scalability

- **SNR scalability**
  - aka quality, bit-rate scalability
- **Space scalability**
  - aka resolution scalability
- **Time scalability**
  - aka frame-rate scalability
- **Other (object scalability, …)**

Scalability: pros…

- Encode once, decode many!
- Optimal use of the network
- Space saving on the server
- Each user is provided exactly with the demanded information
- Error resilience (but what if base layer is lost?)
- Dynamic decoding rate

Scalability: cons

- **Complexity increases**
- **Rate-distortion performances are degraded**
  - This goes worse and worse when more layers are added
  - Is it better to degrade base layer or details?
- **Fully useful only if the infrastructure is changed**
  - Multicast routers are needed

Scalability: issues

- **How can we obtain quality, space and time scalability?**
- **How can we achieve scalability with low impact on rate-distortion performance?**
- **How can we achieve scalability with low complexity increase?**
- **How can we trade-off between base layer and details degradation?**

Overview

- **Examples and motivations**
- **Scalable coding for network transmission**
- **Techniques for multiple description coding**
  - Principles
  - MDC techniques

MDC vs. SVC

- **Scalable coding: base + enhancement**
  - If base is lost, enhancements are useless
  - TCP-based transmission: all packets arrive…
  - … but retransmission delay is usually much higher than packet interarrival

- **In conclusion, SVC relies too much on transport layer, which could fail for several reasons**

- **A different transmission model could help when transmission is not reliable and delay is a constraint**

Multiple description coding: principles

- **The encoder creates 2 description of the same importance**
  - Trivial example: odd and even frames
- **Descriptions sent over different (logical) channels**
- **One description is enough for acceptable quality**
- **Receiving 2 descriptions provides the best quality**
- **Generalization: *N* descriptions**

MDC: pros and cons

Pros

- **Good quality without retransmission**
  - Real-time and interactive applications
  - Simplification of network design: no feedback channel needed
- **Traffic dispersion**
  - Better than scalable coding since all streams have the same importance
  - Application to P2P distribution using multiple paths

Cons

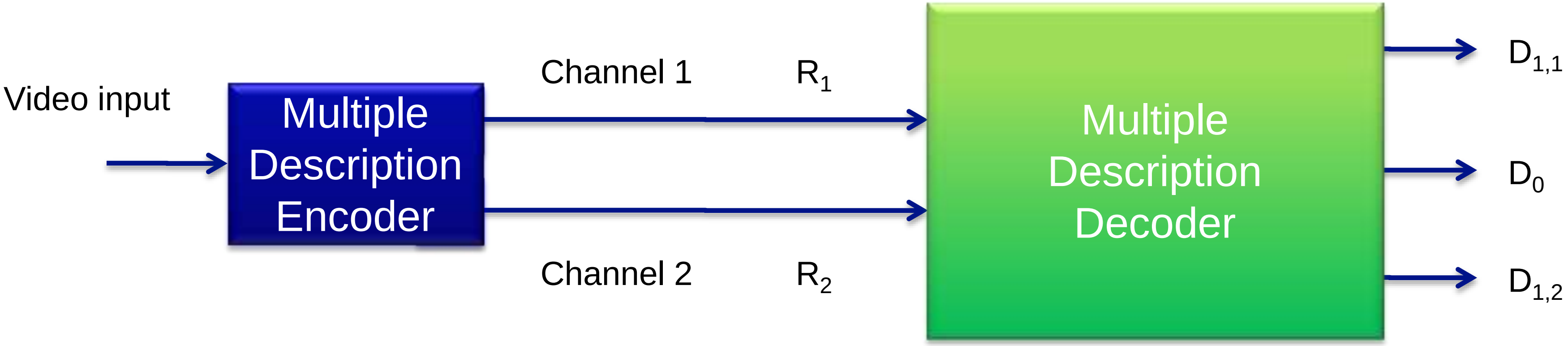
- **Excess rate: for the same quality, MDC demands more bits wrt single description coding**

MDC: conceptual scheme

Usually a balanced scheme is considered:

$R_1=R_2=R/2$

$D_{1,1}=D_{1,2}=D_1$

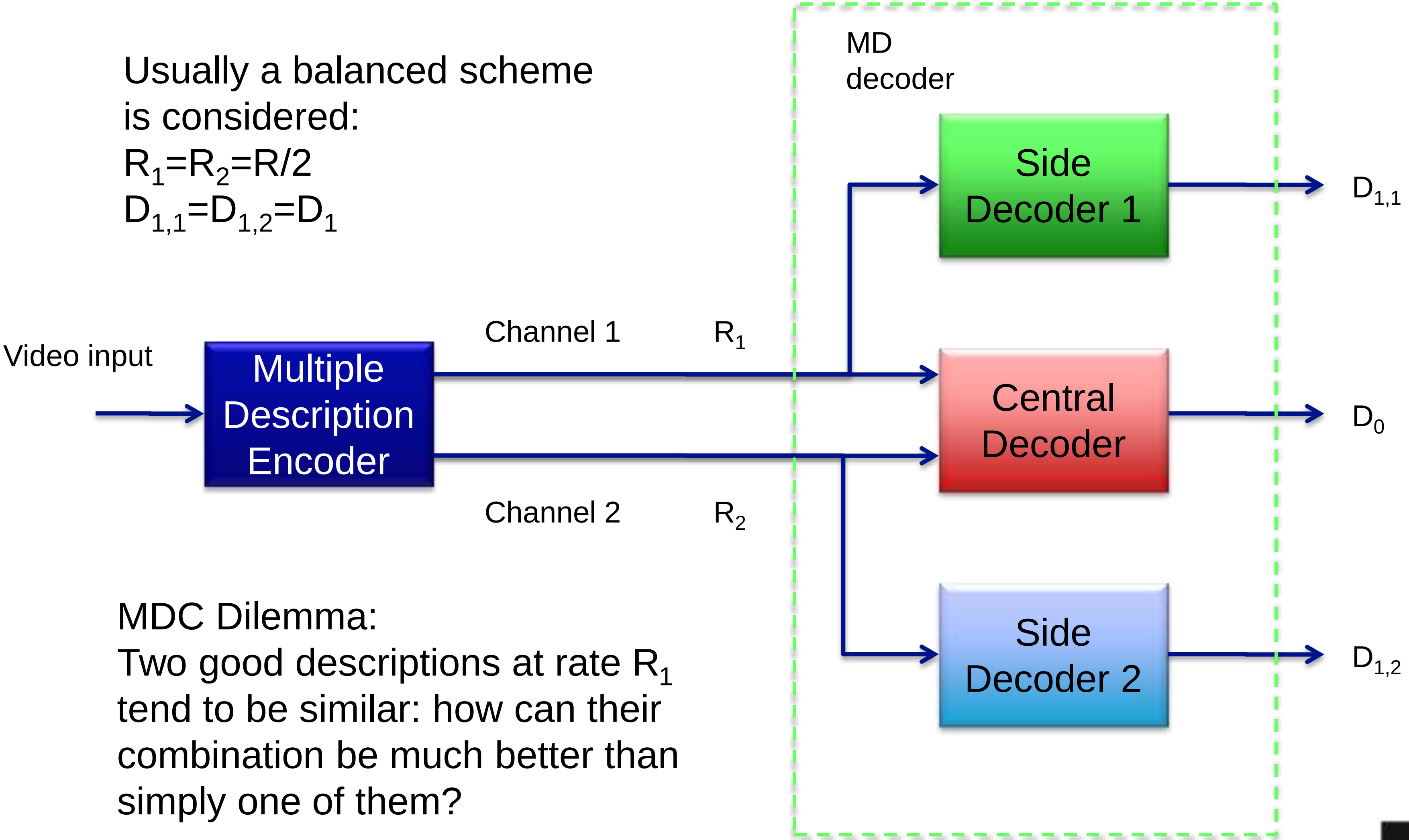


MDC: conceptual scheme

Usually a balanced scheme is considered:

$R_1=R_2=R/2$

$D_{1,1}=D_{1,2}=D_1$



MDC Dilemma:

Two good descriptions at rate R_1 tend to be similar: how can their combination be much better than simply one of them?

MDC problems

- **For SDC, one has to minimize D_0 for a given R**

- **For MDC, we consider the redundancy $\rho(D_0, D_1)$**
  - $\rho(D_0, D_1) = R - R^*$
  - $R = R_1 + R_2$: rate for MDC to assure distortion D_0 for the central decoder and D_1 for the side decoders
  - R^*: rate for the SDC to assure distortion D_0

MDC problems

- **Distortion bounds**
  - If $\rho \rightarrow 0^+$, then $\frac{\partial D_1}{\partial \rho} \rightarrow +\infty$
  - If the distortion bound is achieved for central decoder, use the additional rate for side decoder

- **Another approach uses the error probability p:**

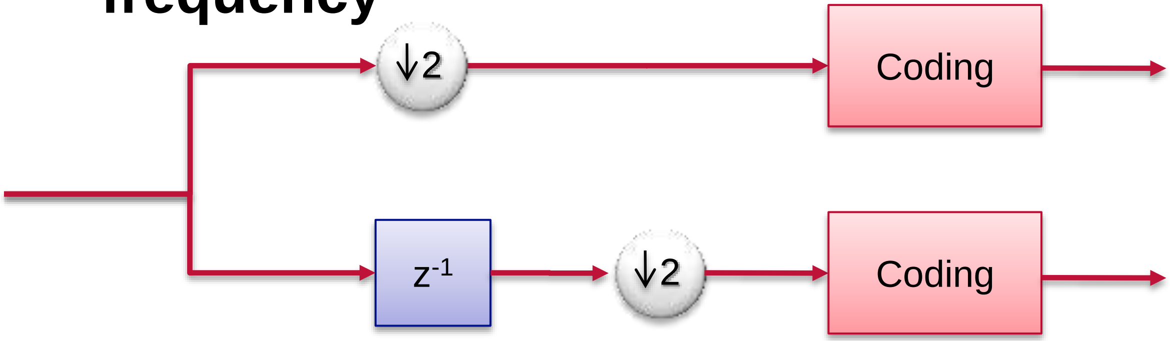
$$J = (1-p)^2 D_0 + 2p(1-p) D_1 + \lambda R$$

Overview

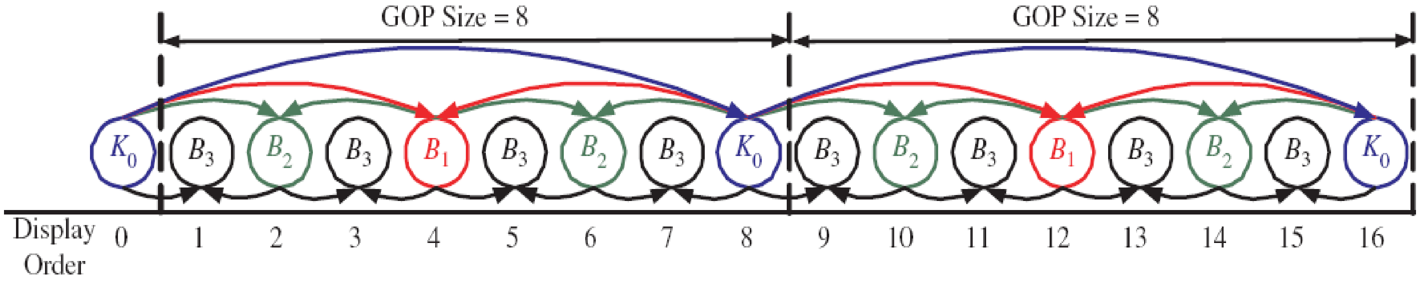
- **Examples and motivations**
- **Techniques for scalable coding**
- **Techniques for multiple description coding**
  - Principles
  - MDC techniques

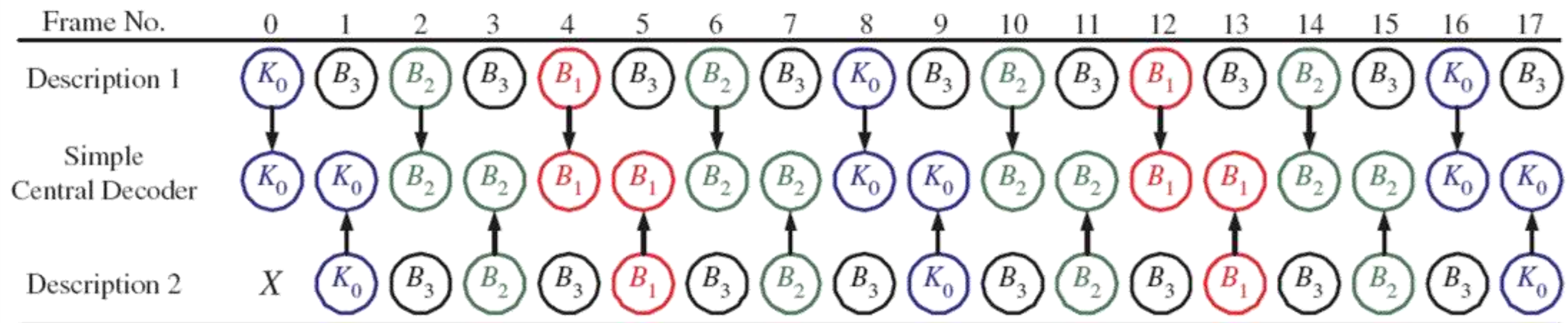
MDC techniques: Channel splitting

- Video data are split into 2 (or more channels) and processed separately
- Channel splitting relies only on channel redundancy
- The splitting can be performed in space, time, frequency



Temporal Channel splitting: An example

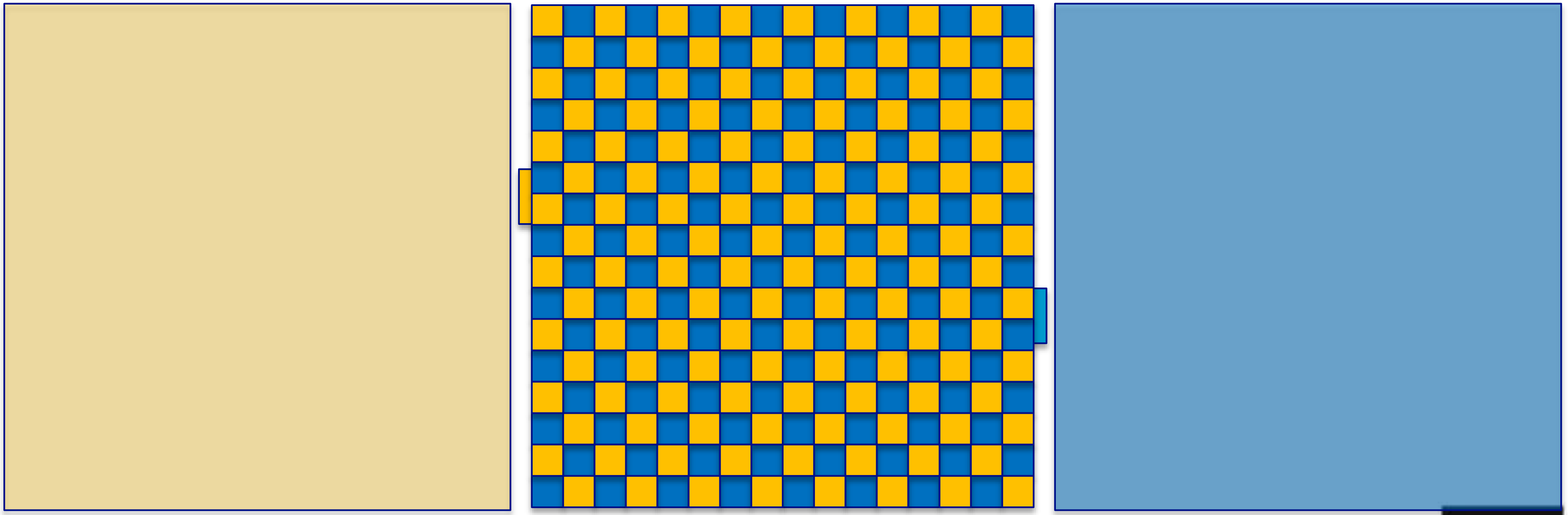




Frame No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Description 1	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3
Simple Central Decoder	K_0	K_0	B_2	B_2	B_1	B_1	B_2	B_2	K_0	K_0	B_2	B_2	B_1	B_1	B_2	B_2	K_0	K_0
Description 2	X	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0

Spatial channel splitting: FMO in H.264

- Images divided into slices
- Slices coded independently
- FMO allows to define arbitrarily shaped slices
- Slices correspond to descriptions



Channel splitting: example

- AR model for vocal signal

$$X_n = aX_{n-1} + Y_n$$

$$Y_n = N(0,\sigma)\mathrm{i.i.d.}, \mathrm{a} \in]0,1[$$

- Filter characterization ?

Channel splitting: example

- AR model for vocal signal

$$X_n = aX_{n-1} + Y_n$$

$$Y_n = N(0,\sigma)\mathrm{i.i.d.}, \mathrm{a} \in]0,1[$$

- Filter characterization ?

$$X(z) = az^{-1}X(z) + Y(z)$$

$$X_n = \sum_{m=0}^{+\infty} a^m Y_{n-m}$$

- Statistical characterization of X_n ?

Channel splitting: example

- **AR model for vocal signal**

$$X_n = aX_{n-1} + Y_n$$

$$Y_n = N(0,\sigma)\mathrm{i.i.d.},\mathrm{a} \in]0,1[$$

- **Filter characterization ?**

$$X(z) = az^{-1}X(z) + Y(z)$$

$$X_n = \sum_{m=0}^{+\infty} a^m Y_{n-m}$$

- **Statistical characterization of X_n? Zero-mean gaussian**

$$r_{xx}(n) = \frac{\sigma^2 a^{|k|}}{1-a^2}$$

For simplicity, we set $\sigma^2 = 1 - a^2 \rightarrow r_{xx}(k) = a^{|k|}$

Channel splitting: example

- **AR model for vocal signal**

$$X_n = aX_{n-1} + Y_n$$

$$Y_n = N(0,\sigma)\mathrm{i.i.d.}, \mathrm{a} \in]0,1[$$

$$X_n = N(0,1), \qquad r_{xx}(k) = a^{|k|}$$

- **The larger *a* , the more correlated the samples of *X***

- **Now we want to encode *X***
  - **Single description : DPCM coding**
  - **Multiple description: Channel splitting with DPCM on each channel and linear interpolation at side decoders**

Channel splitting: example

DPCM coding for AR Gaussian signal with correlation a

$$\tilde{x}_n = aX_{n-1}$$

$$E_n = Y_n$$

$$D_{\text{PCM}} = \sigma^2 2^{-2R} = (1 - a^2)2^{-2R}$$

- **For simplicity, we suppose a perfectly known, and we ignore the quantization error of the predictor (high-rate hypothesis)**
- **The formula for the distortion holds for the single description case, but also for the encoding of any AR Gaussian signal with correlation a**

Channel splitting: example

- **Channel splitting**

$$\boldsymbol{X_e(n) = X(2n)}$$
$$\boldsymbol{X_o(n) = X(2n+1)}$$
$$\boldsymbol{r_{x_e x_e}(k) =?}$$

Channel splitting: example

- Channel splitting

$$X_e(n) = X(2n)$$
$$X_o(n) = X(2n+1)$$
$$r_{x_e x_e}(k) = E[X(2n)X(2n+2k)] = r_{xx}(2k) = a^{2|k|}$$

- On each channel we have an AR Gaussian signal with correlation a^2

- We can use the previous formula for distortion

Channel splitting: example

Central decoder

- On both channel we receive an AR Gaussian signal

$$D_o = D_e = (1 - a^4)2^{-2R}$$

- The distortion (called full-rate distortion) is the average of the distortions on the two channels

$$D_F = \frac{1}{2}D_o + \frac{1}{2}D_e = (1 - a^4)2^{-2R}$$

- The central decoder has a constant factor increase in distortion (at high rate):

$$\frac{D_F}{D_{PCM}} = (1 + a^2) > 1$$

- The more correlated are the samples, the larger lossses will be caused by channel splitting

Channel splitting: example

- Side decoder (half-rate decoder)
- We only receive even samples, therefore:

$$D_e = (1 - a^4)2^{-2R}$$

- For odd samples: the quantization distortion (modulated by a term depending on quantization noise correlation, $\omega \in]0, 1[$) *plus* the interpolation error

$$D_o = \omega D_e + (1 - a)^2 + \frac{1}{2}(1 - a^2)$$

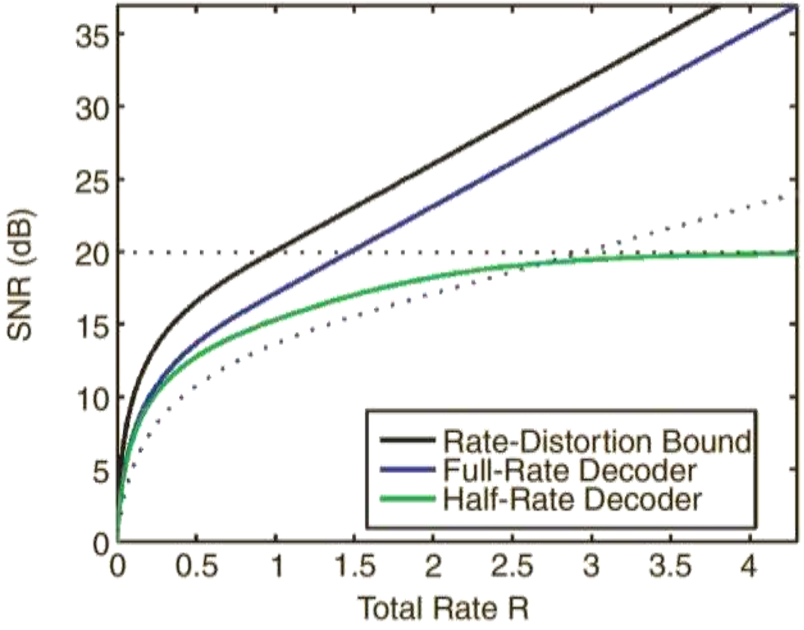
$$D_H = \frac{1}{2}D_o + \frac{1}{2}D_e = \frac{1}{2}\left[(1 - a)^2 + \frac{1}{2}(1 - a^2)\right] + \frac{1+\omega}{2}D_F$$

- In the worst case $(\omega = 1)$, $D_H = D_F + k_a$
- Actually at low bit rate we can have $D_H < D_F$

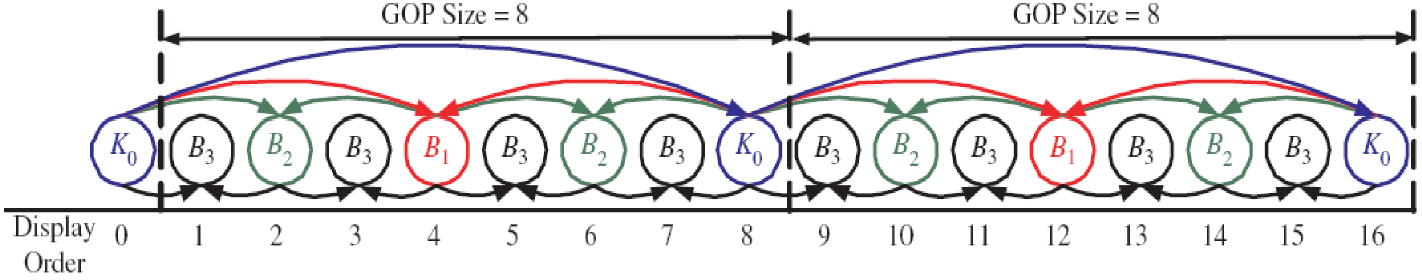
Channel splitting: example

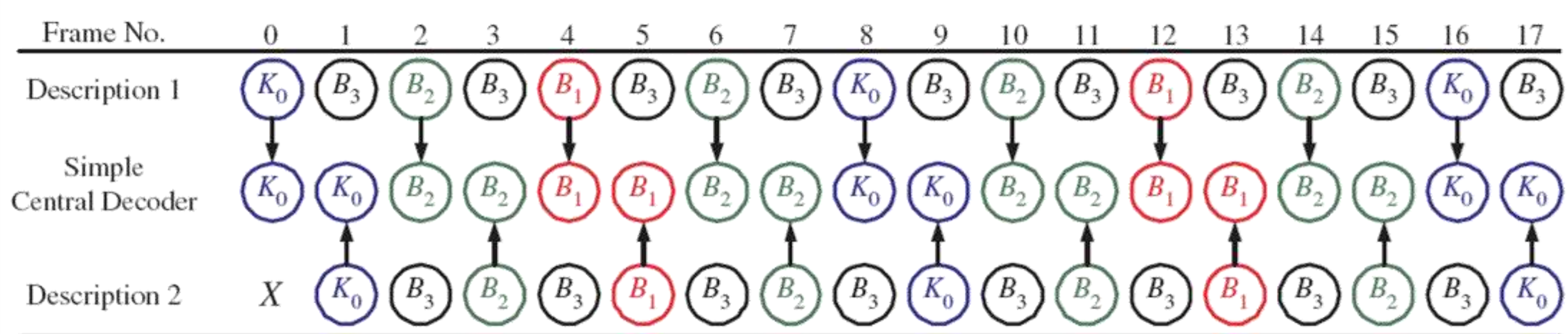
Summary:

- Central decoder: constant loss (in dB) wrt single description coding
- The side decoder performance saturates since it cannot get rid of interpolation error, no matter how high the rate is
- At low bit-rate, the side decoder can outperform the central one: our approximation do not hold longer; it could be better to use the few bits we have to encode pretty one only the even samples and interpolate the odd ones
- Results are coherent with experimentation in audio coding (see figure)



Channel splitting: an example for video coding

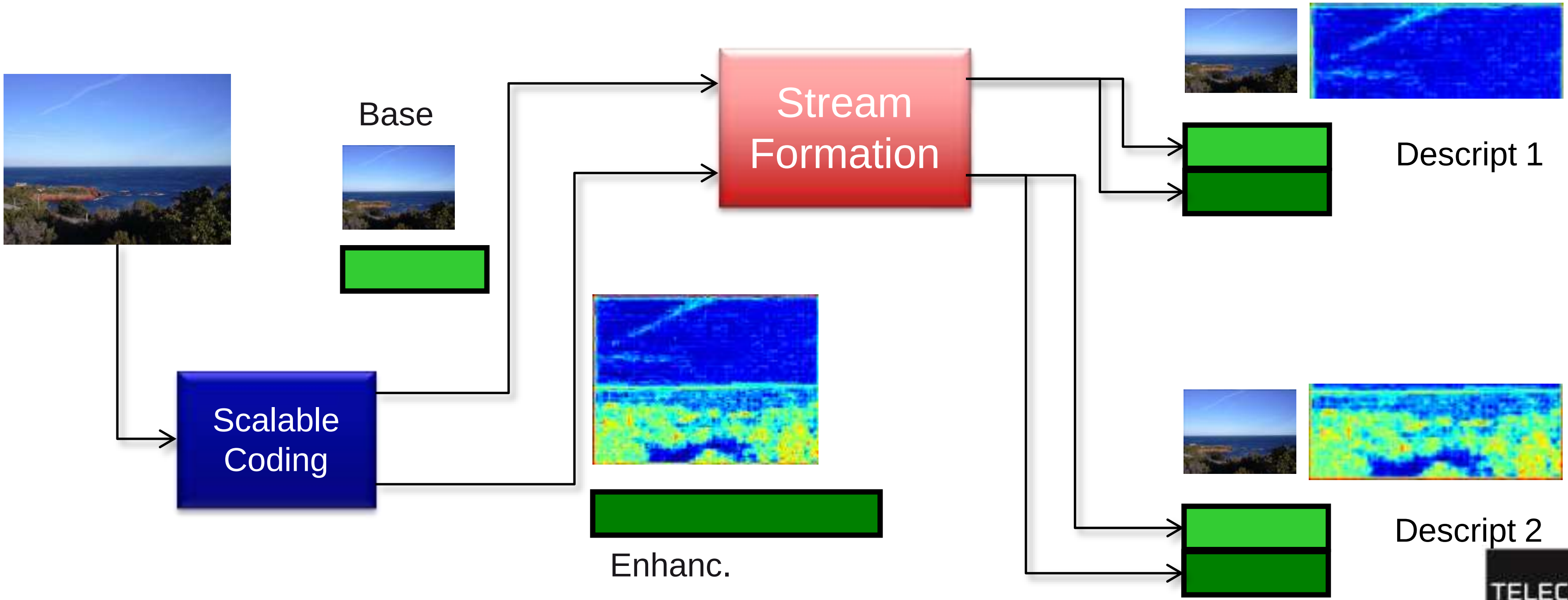




Frame No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Description 1	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3
Simple Central Decoder	K_0	K_0	B_2	B_2	B_1	B_1	B_2	B_2	K_0	K_0	B_2	B_2	B_1	B_1	B_2	B_2	K_0	K_0
Description 2	X	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0	B_3	B_2	B_3	B_1	B_3	B_2	B_3	K_0

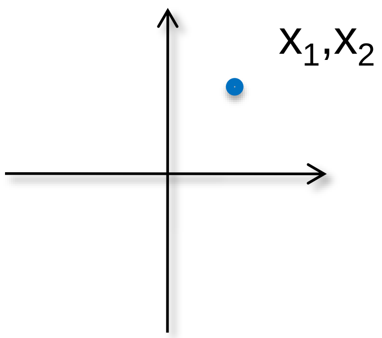
MDC techniques: Unequal Error Protection

- Produce a two-layers progressive representation of data
- Insert the base layer in both descriptions
- Split the enhancement layer in the two descriptions

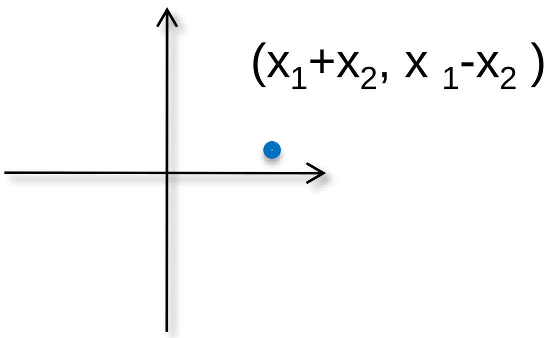


MDC techniques: Correlating transforms

- **Keep some correlation among coefficients**
- **Statistical dependencies are used to estimate transform coefficients that are in a lost description**



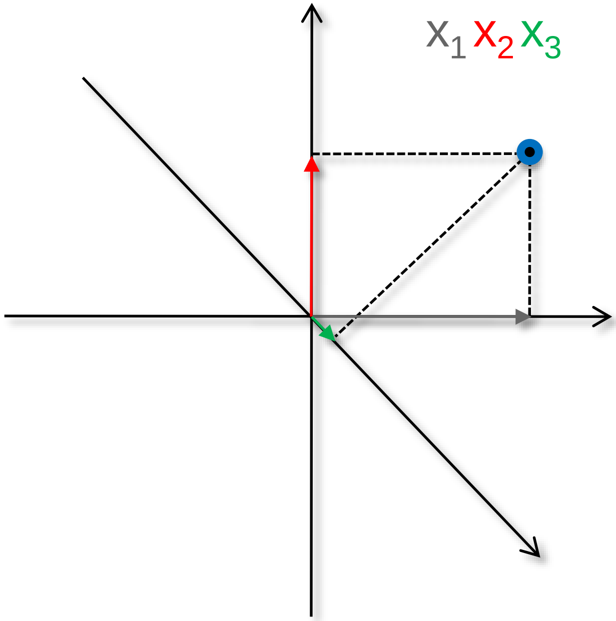
X_1, X_2: Independent Gaussian variables

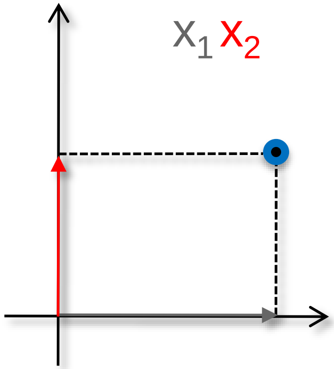


Y_1, Y_2: Correlated Gaussian variables

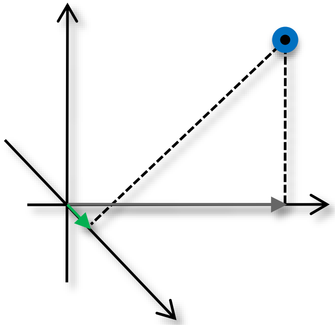
MDC techniques: Redundant transforms

- Project the input vector into a redundant set
- Use subsets of coefficients to form descriptions





Description 1

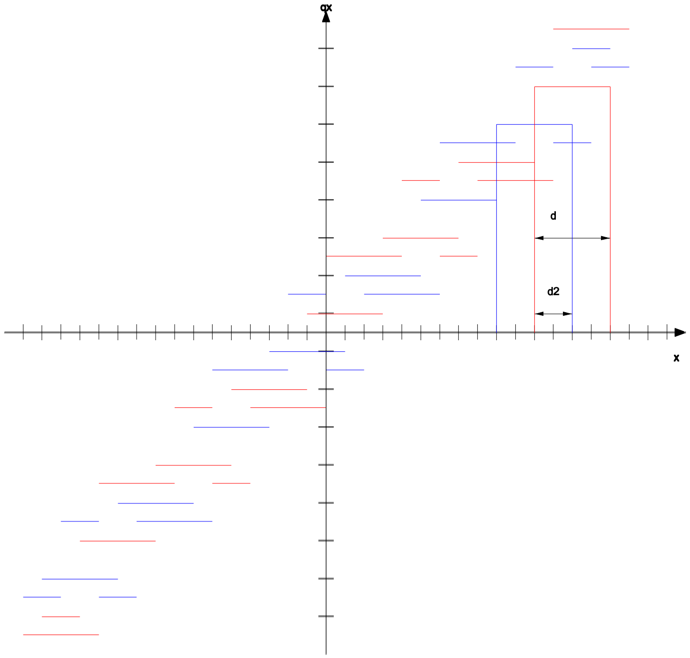


$x_1\ x_3$

Description 2

Other MDC approaches

- **Shifted intervals**
- **Non-convex quantizers**
  - One index corresponds to disjoint intervals
- **Oversampled filter banks**
  - Related to wavelets
  - Scalability+MDC



Mismatch for video MDC

- Mismatch (i.e. drift) if prediction is made with information not available at the decoder

- Mismatch is probable because of packet loss

- No mismatch: improves side decoders
- Mismatch: improves central decoder

Conclusions

- **Traditional approaches unfit to video delivery over networks**
- **Basic idea: split the stream into substreams**
  - Scalable coding: good RD performances, good use of resources (memory, complexity, bandwidth), some error resilience
  - MDC: acceptable RD performances, quite good use of resources, good error resilience
- **Cons:**
  - Compression capability
  - Complexity

Bibliography

Scalable video coding

[1] W. Li, “Overview of Fine Granularity Scalability in MPEG-4 Video Standard”. IEEE Trans. Circ. and Syst. for Video Tech., 11(3) Mar. 2001.

[2] J.-R. Ohm, “Advances in Scalable Video Coding”. Proceedings of the IEEE, 93(1), Jan. 2005.

[3] H. Schwarz, D. Marpe, T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard”. IEEE Trans. Circ. and Syst. for Video Tech., 17(9), Sept. 2007

Multiple Description Coding

[4] V. Goyal, “Multiple Description Coding: Compression Meets the Network”. IEEE Sign. Proc. Magazine, Sept. 2001

[5] Y. Wang, A. R. Reibman, S. Lin, “Multiple Description Coding for Video Delivery”. Proceedings of the IEEE, 93(1), Jan. 2005.