Digital Video Systems
ECE 634

engineering.purdue.edu/~reibman/ece634/index.html

Professor Amy Reibman
MSEE 356
reibman@purdue.edu

Multiple description coding

- Layered coding vs. MDC
  - Layered coding: Base layer is high priority; cannot reconstruct video without base layer. Enhancement layer is "bonus"
  - Multiple description coding: either of the two descriptions are OK; both are better
- (Most) multiple description coding are not exactly compatible with standards
  - Partitioning the video and reassembling can (sometimes) be done outside the context of the standards

Considerations for MD coding

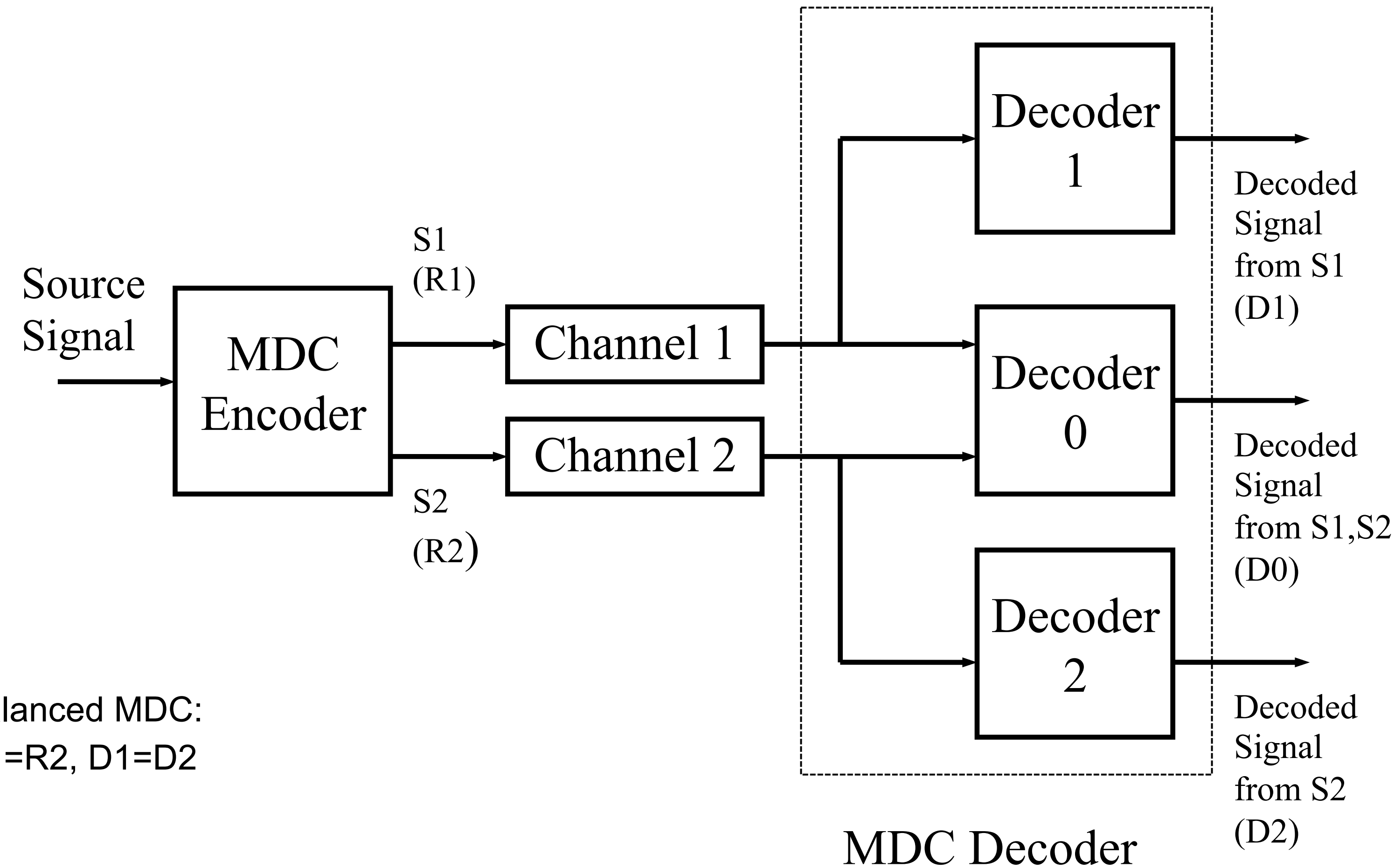
- Compression efficiency
- Decoder complexity
- Resilience to losses
- Flexible rate adaptation
  - Number of descriptions
  - Can you add redundancy flexibly?
- Compatibility with standards
- Ease of prioritization

- Prediction structure and method of partitioning controls most of these!

Multiple Description Coding

- MDC: Generate multiple correlated descriptions
  - Any description provides low but acceptable quality
  - Additional received descriptions provide incremental improvements
  - However: correlation → reduced coding efficiency
- Assumptions:
  - Multiple channels between source and destination
  - Independent error and failure events
  - Probability that all channels fail simultaneously is low
  - All are reasonable assumptions for the Internet and wireless networks, provided data are properly packetized and interleaved
- Design goal:
  - maximize the robustness to channel errors at a permissible level of redundancy

Generic Two Description Coder



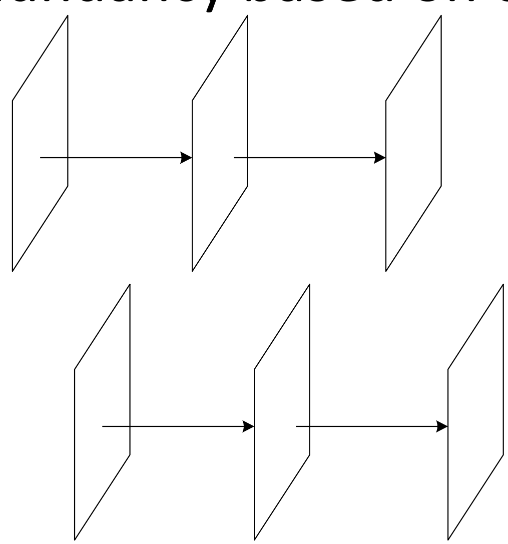
Balanced MDC:
R1=R2, D1=D2

Four approaches to MDC

- Duplication!
- Interleaved temporal sampling
- Multiple Description Scalar Quantizer
  - Basic idea: Use interleaving quantizers of a scalar value
- Pairwise Correlating Transforms
  - Basic idea:
    - Use *NxN* linear transform to generate two groups of coefficients with a desired amount of correlation, so that one group can be estimated from the other with a given accuracy.
  - Conceptually simple method:
    - Use a decorrelating transform followed by a pairwise correlating transform

Video Redundancy Coding in H.263+

- Code even frames and odd frames as separate threads
  - Good video quality at packet loss rates up to 20%
  - High redundancy (~30%) due to reduced prediction gain because of longer distance between frames
  - Hard to vary the redundancy based on channel loss characteristics



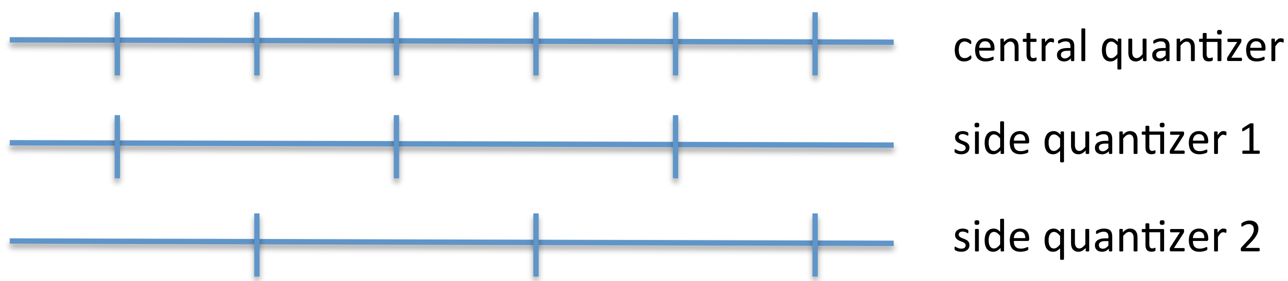
©Yao Wang, 2004

Three-thread three-pictures-per-thread VRC

- frame 0 is sync frame; appears in all threads
- frame 1 is in thread 1, predicted from 0
- frame 2 is in thread 2, predicted from 0
- frame 3 is in thread 3, predicted from 0
- frame 4 is in thread 1, predicted from 1
- frame 5 is in thread 2, predicted from 2
- frame 6 is in thread 3, predicted from 3
- frame 7 is sync frame; appears in all threads

Multiple description Scalar Quantization

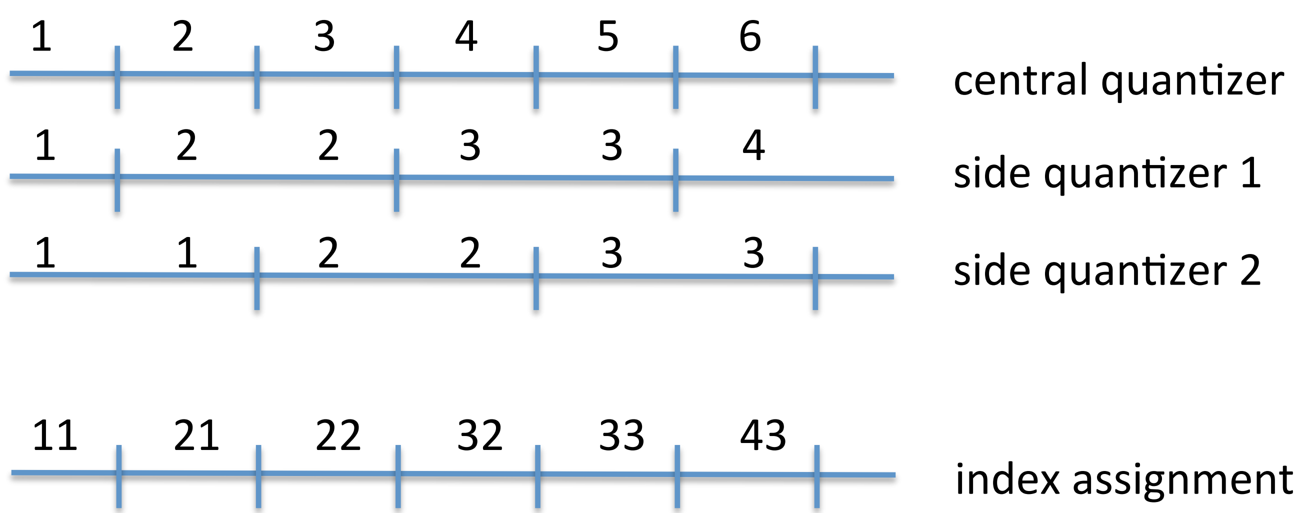
- Basic idea: Use interleaving scalar quantizers
  - a "central quantizer" when both descriptions received
  - a "side quantizer" when only one description received



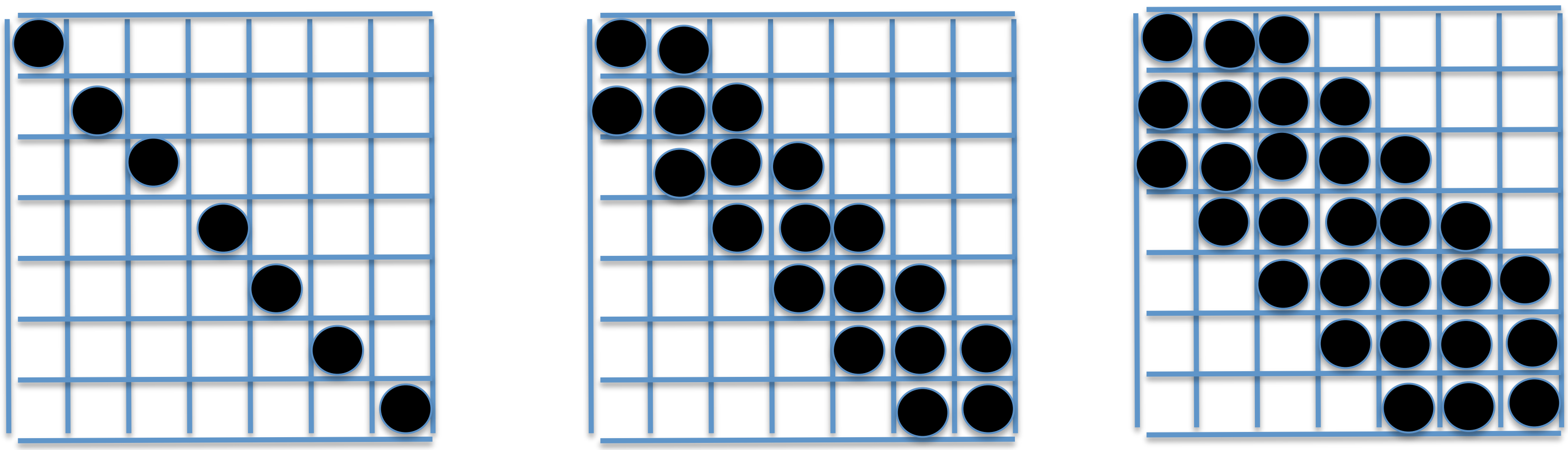
- Very elegant, but their design is complicated

MDSQ Index assignment

- Quantize once with desired quantizer
- Then assign each "bin" an index in each channel



Some MDSQ Index assignments



Less efficient

More robust

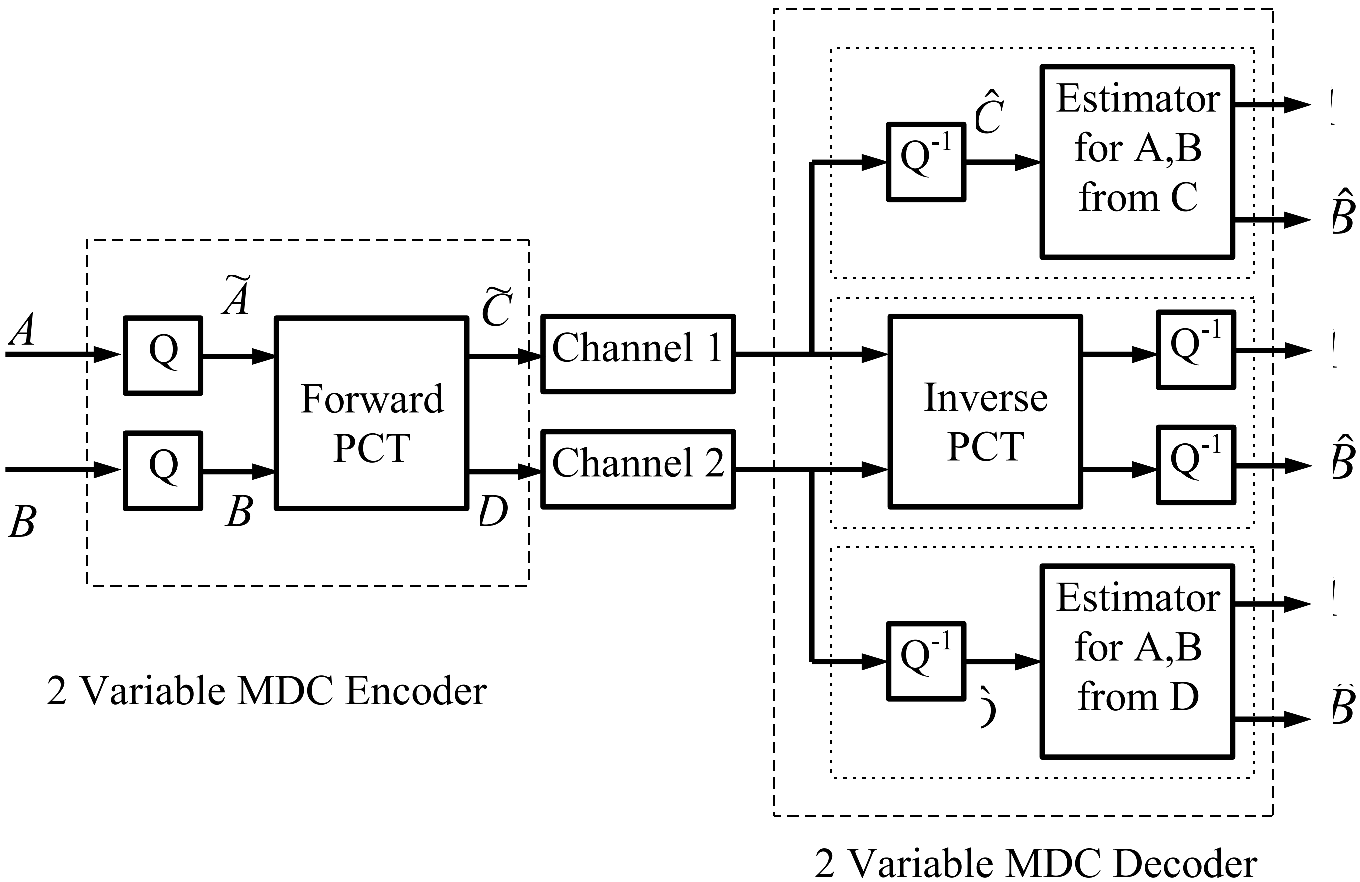
More efficient

Less robust
(single-channel reconstruction
error can be quite large)

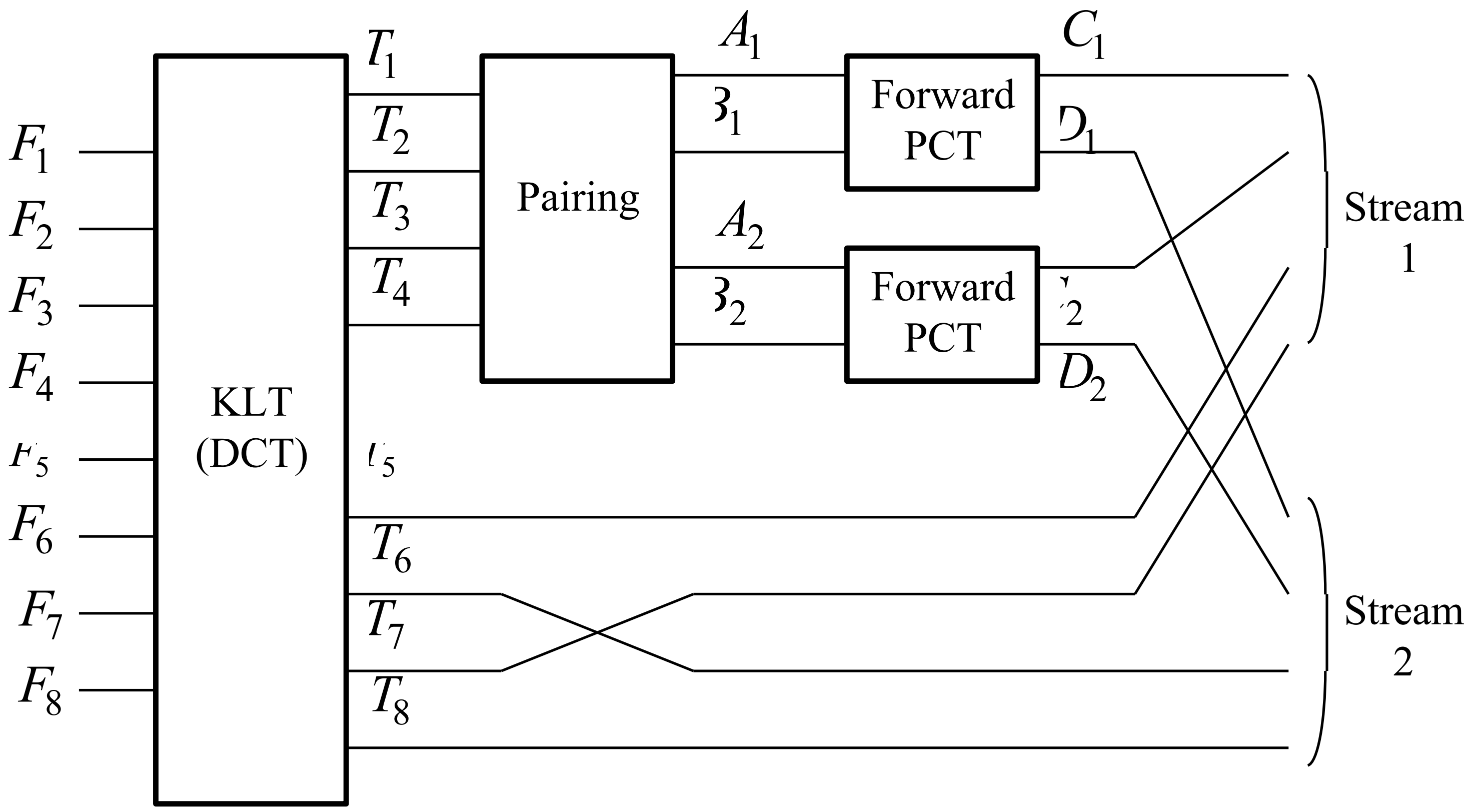
Pairwise correlating transforms

- Basic idea:
  - Generate two sets of transform coefficients, one for each description, such that each set is uncorrelated within the set, but correlated pairwise across sets
  - Take output of a decorrelating transform (KLT), and recorrelate pairs of coefficients;
    assign each coefficient of a pair to a different description
  - At decoder: if a coefficient is missing, estimate it using its pair

Coding of a Single Pair



MDC Using Pairwise Correlating Transform



$$\sigma_n^2 \prec D_{1,\max}, n \geq 5$$

Wang Orchard Reibman ICIP98

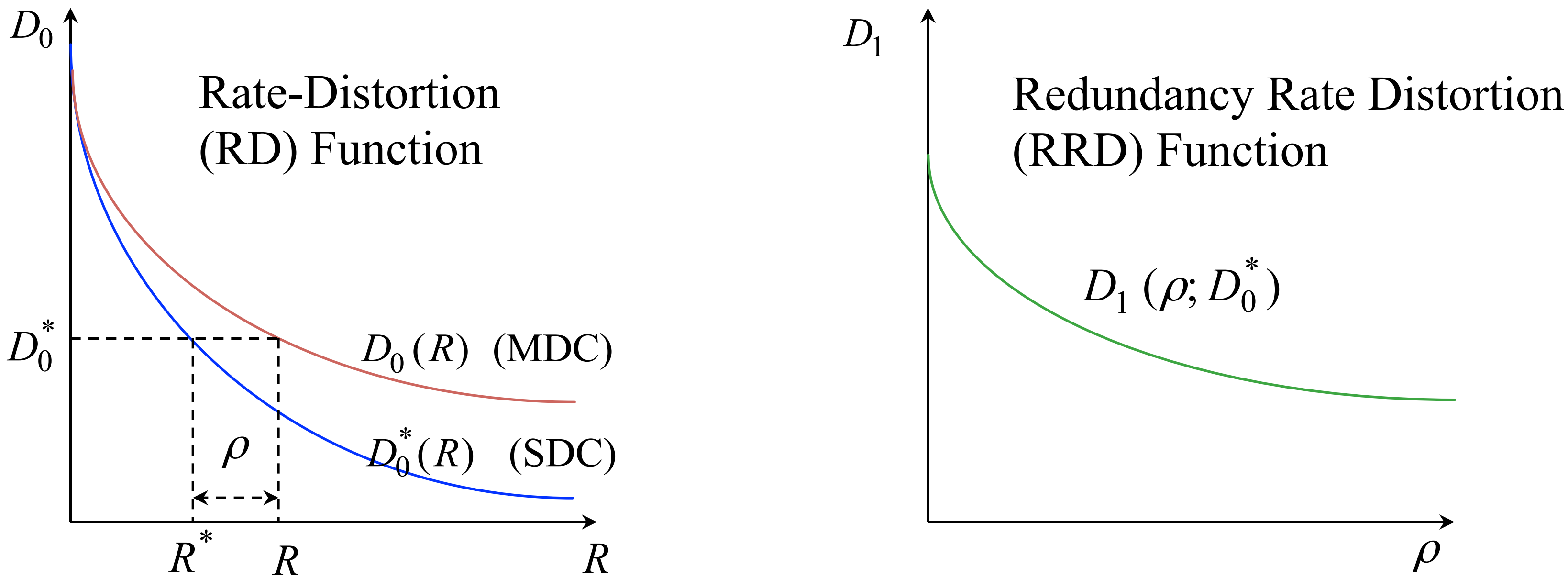
Performance evaluation of MD coders

- Definitions:
  - Distortion with both descriptions: D_0
  - Average distortion with one description: D_1
  - Single description coder R(D) function: $D_0^*(R^*), D_1^*(R^*)$
  - Multiple description coder rate-distortion: $D_0(R), D_1(R)$
- *Redundancy Rate-Distortion :*
  - To decrease D_1 one must introduce correlation between the two descriptions and thereby reduce the coding efficiency

$$\text{Let } \rho = R - R^*, \text{ when } D_0^*(R^*) = D_0(R) = D_0^*$$

$$\text{Redundancy Rate-Distortion Function: } \rho(D_1; D_0^*) \text{ or } D_1(\rho; D_0^*)$$

Redundancy Rate Distortion



- Design criteria for MD coders
  - Minimize D_1 for a given ρ , for fixed R^* or D_0^* (minimizing the average distortion given channel loss rates, for given total rate)
  - Can easily vary the ρ vs. D_1 trade-off to match network conditions

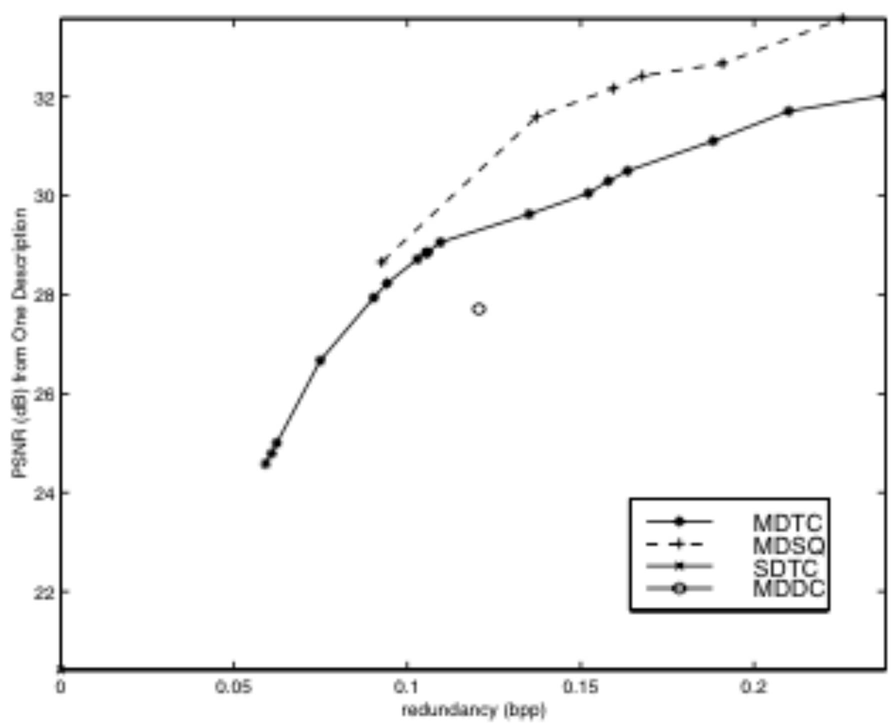
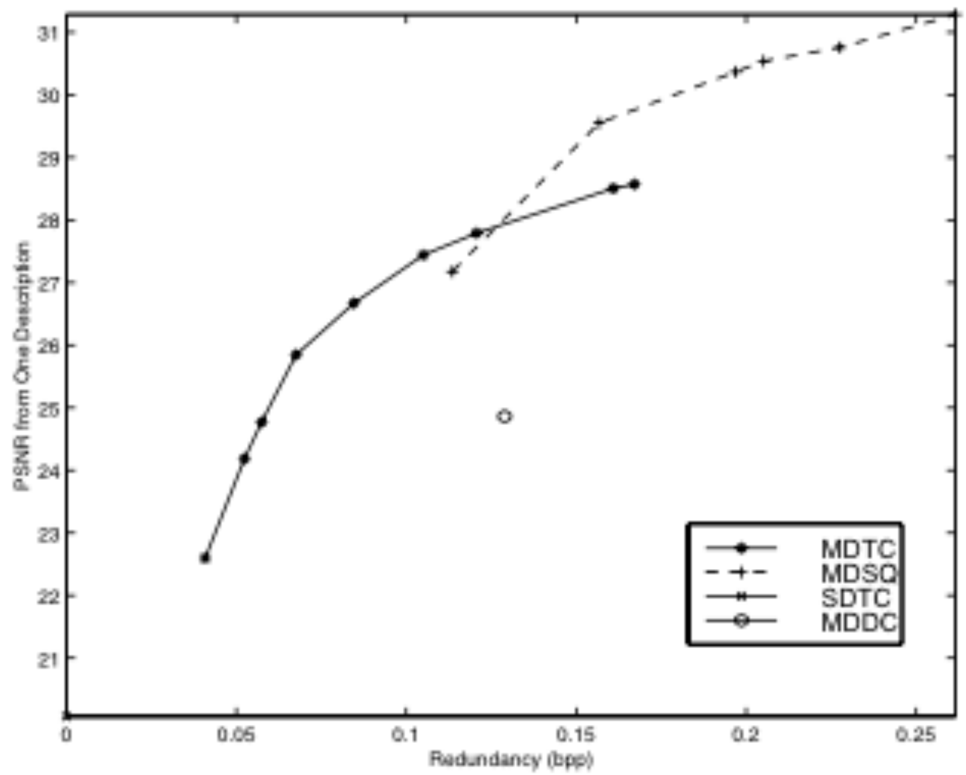


Fig. 9. RRD performance of different coders for image *Lena*. $R^* = 0.63$ bpp, $D_0 = 35.78$ dB.

Fig. 10. Image reconstruction results. (a) Reconstructed from both descriptions (R^{*}=0.60 bpp, D_0=35.78 dB); (b) MDTC, from a single description, ρ=0.088 bpp (15%), D_1= 27.94 dB; (c) MDTC, from a single description, ρ = 0.133 bpp (22%), D_1 = 29.63 dB; (d) MDSQ, from a single description, ρ = 0.090 bpp (15%), D_1 = 28.63 dB.



ig. 11. RRD performance of different coders for image *Horse*. $R^{*} = 0.70$ bpp, $D_0 = 33.51$

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

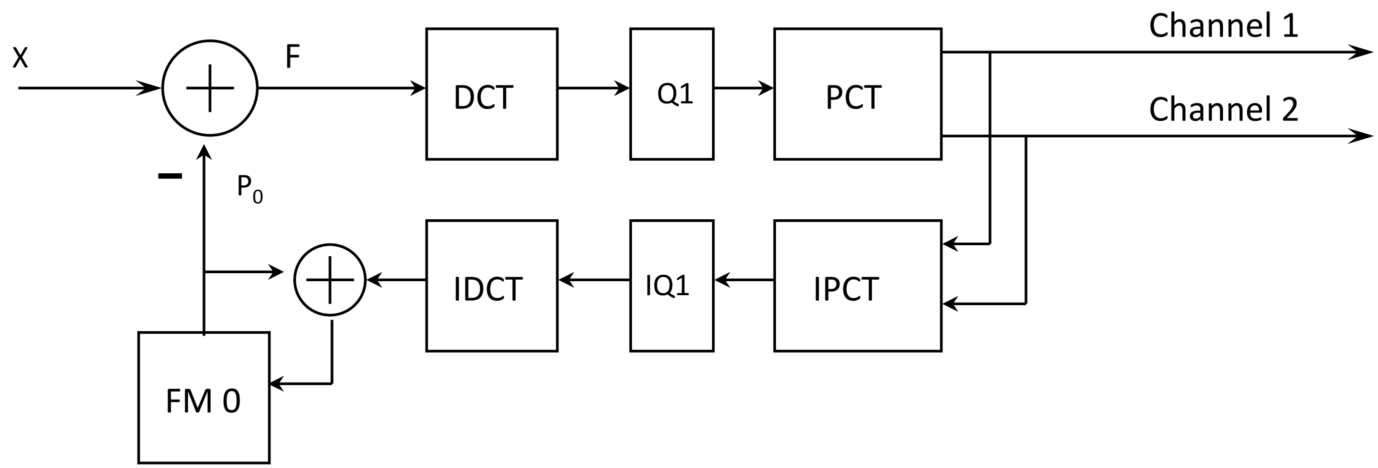
Fig. 12. Image reconstruction results. (a) Reconstructed from both descriptions (R^*=0.70 bpp, D_0^*=33.51 dB); (b) MDTC, from a single description, ρ=0.085 bpp (12%), D_1= 26.74 dB; (c) MDTC, from a single description, ρ=0.121 bpp (17%), D_1= 27.81 dB; (d) MDSQ, from a single description, ρ=0.110 bpp (16%), D_1=27.09 dB.

MD video questions

- How should the following elements of a hybrid video coder be adapted for a multiple description video coder?

- Side information (MB type)
  - (duplicate)
- Motion vector information
  - (duplicate)
- Prediction error (compared to two-channel reconstruction)
  - (pairwise correlating transforms)
- Mismatch (between two- and one-channel reconstructions)
  - (extra prediction-error signals)

Core MDTC-based video coder

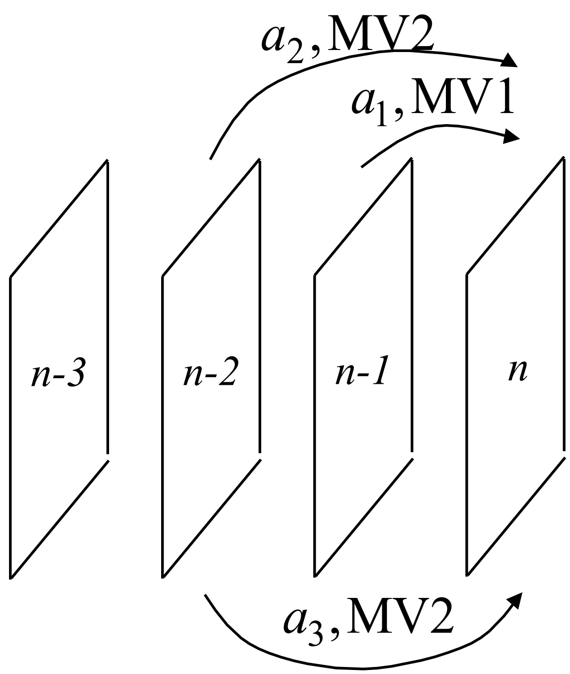


- Apply Pairwise Correlating Transform (PCT) to DCT coefficients inside main loop of hybrid video coder
- Mismatch if both channels are not received
- Mismatch *for this frame* is limited to the orthogonal subspace, but motion compensation may spread it

Multiple Description Motion Compensation (Wang and Lin, 2001)

- A description contains even (or odd) frames only, but each frame is predicted (central predictor) from both even and odd past frames
- Code the central prediction error
  - sufficient if both descriptions are received
- To avoid mismatch, a side predictor for even frames predicts only from the past even frame, and the mismatch signal (difference between central and side prediction) is also coded
- The quantization of the predictors and the mismatch error controls the redundancy of the coder, and can be designed based on the channel loss characteristics

Special Case: Two-Tap Predictor



$$\text{Central predictor: } \hat{\psi}_0(n) = a_1\tilde{\psi}_0(n-1) + a_2\tilde{\psi}_0(n-2)$$

$$\text{Central prediction error: } e_0(n) = \psi(n) - \hat{\psi}_0(n) \rightarrow \tilde{e}_0(n)$$

$$\text{Side predictor: } \hat{\psi}_1(n) = a_3\tilde{\psi}_1(n-2)$$

$$\text{Mismatch error: } e_1(n) = \hat{\psi}_0(n) - \hat{\psi}_1(n) - q_0(n) \rightarrow \tilde{e}_1(n)$$

$$\text{Send: } \tilde{e}_0(n), \tilde{e}_1(n), \text{MV1}, \text{MV2}$$

$$\text{Non-leaky predictor: } a_1 + a_2 = 1, a_3 = 1$$

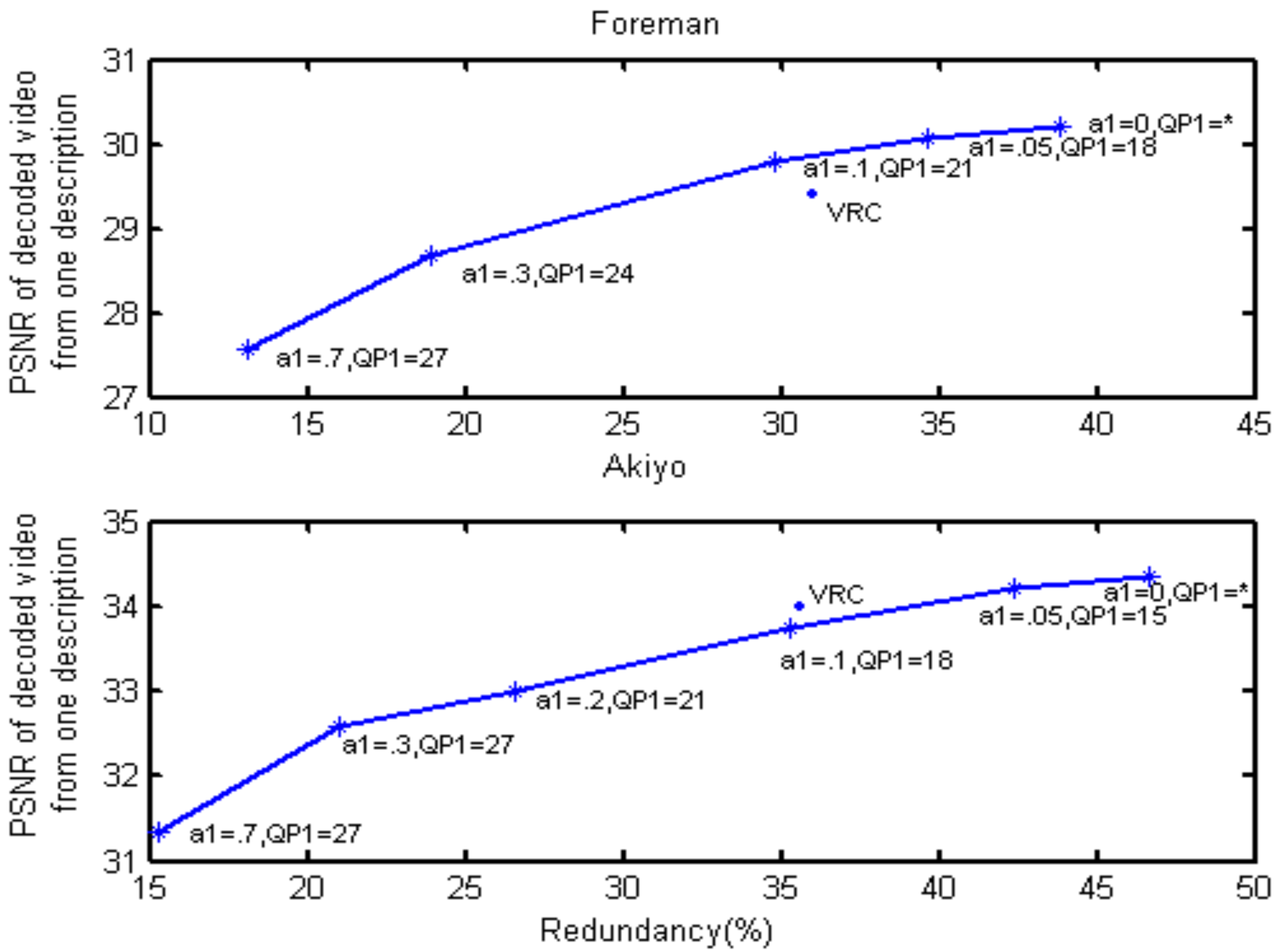
If both descriptions received (have both $\psi_0(n-1), \psi_0(n-2)$)

$$\psi_0(n) = \hat{\psi}_0(n) + \tilde{e}_0(n) = \psi(n) + q_0(n)$$

If one description is received (have only $\psi_1(n-2)$)

$$\psi_1(n) = \hat{\psi}_1(n) + \tilde{e}_0(n) + \tilde{e}_1(n) = \psi(n) + q_1(n)$$

RRD Performance of VRC and MDMC



©Yao W

Performance in Packet Lossy Networks

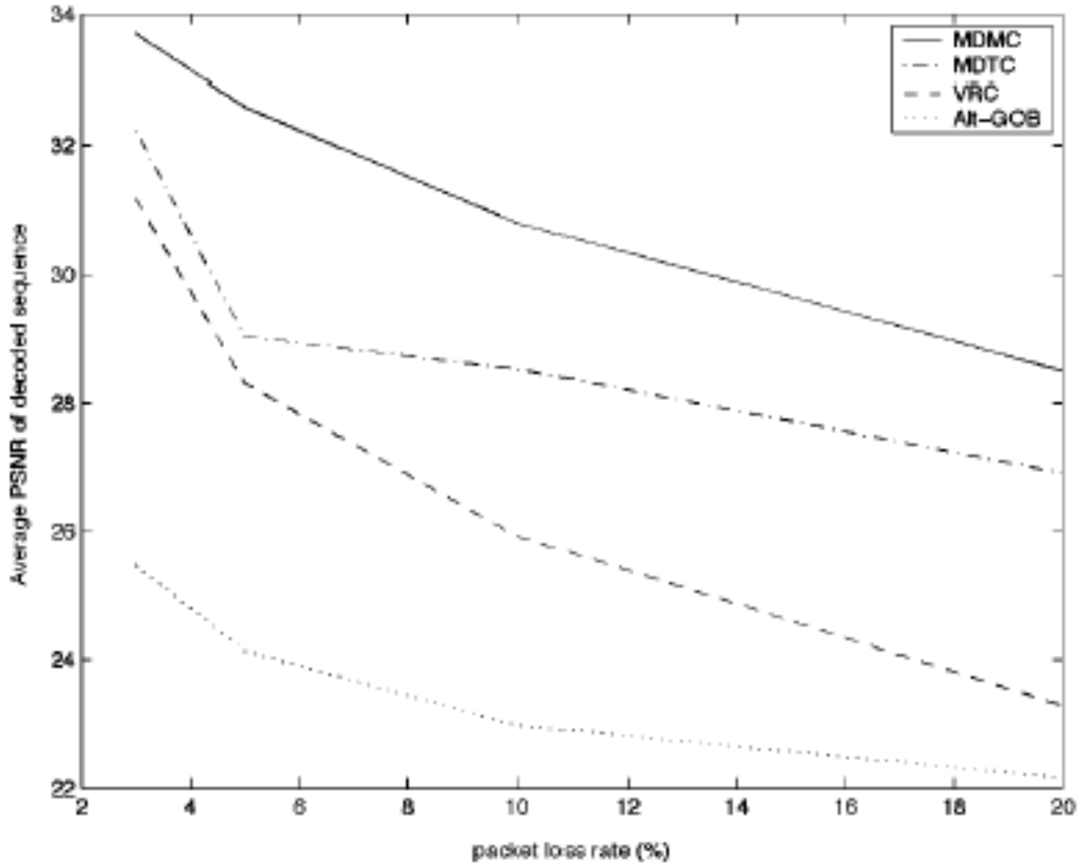


Fig. 13. PSNR of decoded sequences in different packet loss rates. Foreman, 7.5 fps, 144 kbps, two packets per frame.

Sample Reconstructed Frames

(10% Random Packet Loss, MDMC on top, VRC on bottom)

Summary: Challenges in Designing MD Video Coder

- To achieve high coding efficiency, the encoder should retain the temporal prediction loop
- Prediction strategies are key to control trade-off between added redundancy and reduced compression efficiency
  - Predict from two-description reconstruction, or one?
- Prediction based on two-description reconstruction
  - Higher prediction efficiency
  - Mismatch problem at the decoder
- Prediction based on single-description reconstruction
  - Lower prediction efficiency
  - No mismatch problem
- One design strategy
  - Predict based on two-description reconstruction, but explicitly code the mismatch error

References

- Y. Wang and Q. Zhu, “Error control in video communications – A review,” Proc. IEEE, 1998

- Y. Wang, A. R. Reibman, and S. Lin, “Multiple description coding for video delivery”, invited paper, Proc. IEEE, Jan. 2005.