Multimedia Networking ECE 599

Prof. Thinh Nguyen

School of Electrical Engineering and Computer Science

Video Communications

- Challenge for video communication
- Transport level error control
- Error resilient coding
- Error concealment
- Scalable coding
- Multiple description coding

Based on Y. Wang and M. van de Schaar’s lecture notes

Challenge for Video Communications

- Real networks are unreliable!
  - Wireless networks: random bit errors, long burst errors, and possibly link downs
  - Internet: packet loss and variable delay to network congestion
  - Excessive delay = loss for real-time applications
- Real networks are heterogenous in bandwidth and reliability
- Video data are delay-sensitive
  - One cannot relay on retransmission of lost packets because of stringent delay requirement.

Problems with compressed video in a lossy channel.

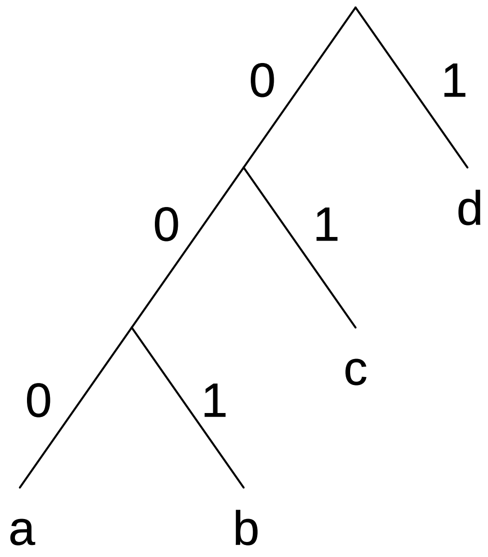
- Have poor reconstruction quality when certain coded (compressed) data is lost.

- Compressed video data is very sensitive to transmission errors
  - Variable length coding (huffman)
  - Temporal predictive coding
  - Spatial predictive coding
  - All contribute to error propagation within the same frame as well as following frames.

Problem with variable length coding

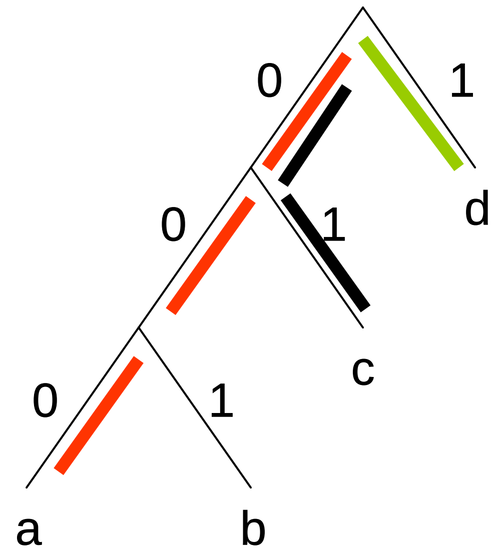
- Synchronization error!

Ex: a,b,d,c = 000001101



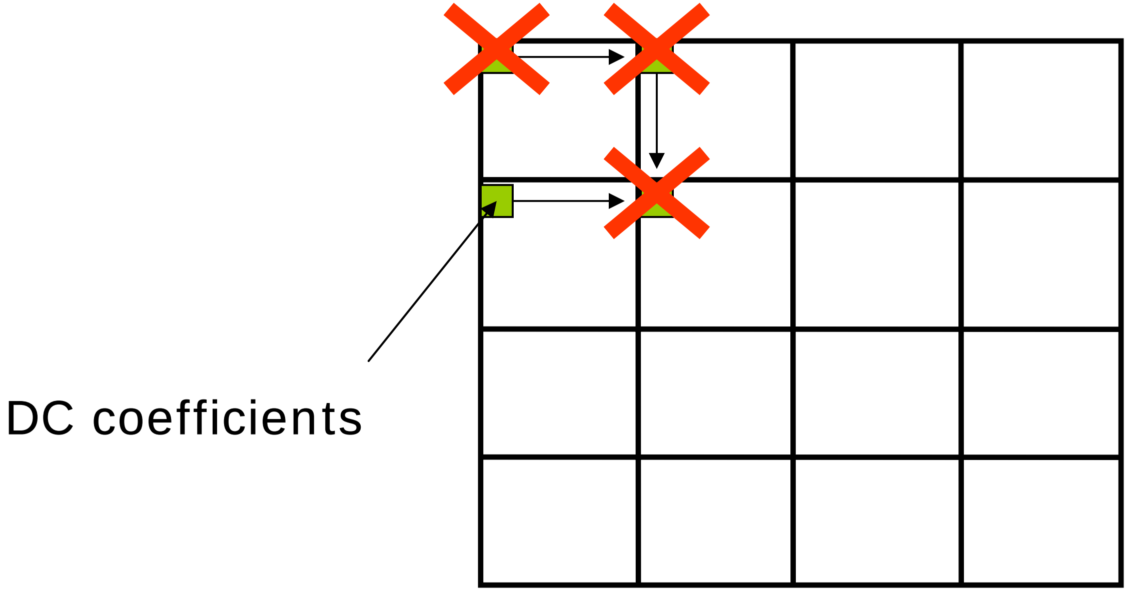
Suppose a bit in a is fliped

a = 100, abcd = 100001101

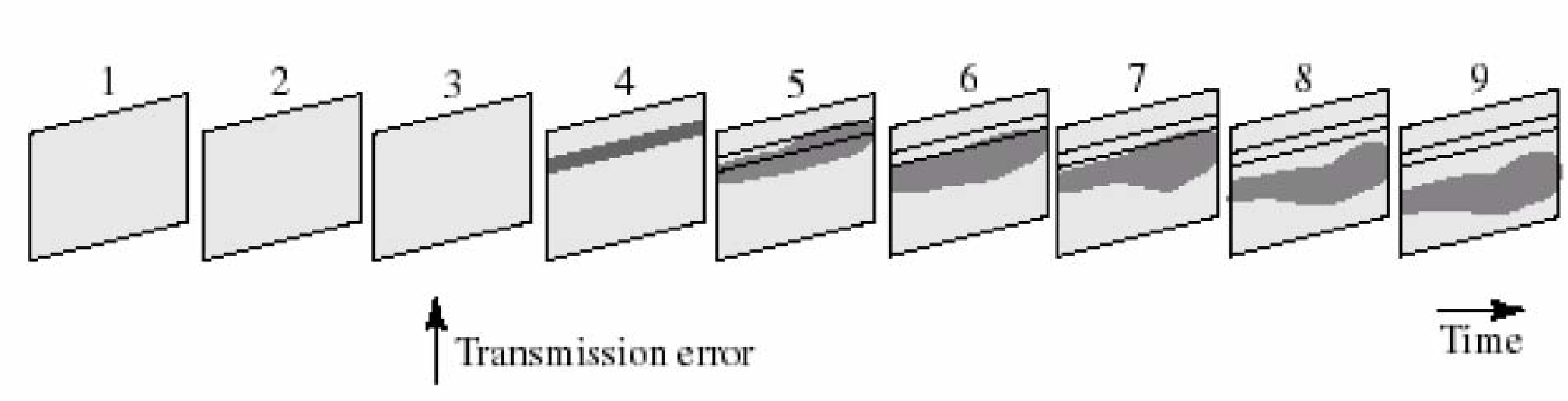


Decoder: d a c d c

Problem with spatial prediction



Problem with temporal prediction



Examples video frames resulted from transmission errors

No loss

3% loss

5% loss

10% loss

H.263 bit stream

QoS Requirements for Typical Video applications

- Interactive two-way visual communication
- One way video streaming
- One way video downloading
  - No different from downloading

Interactive two way video communication

- Ex: Video conferencing
  - Very stringent delay requirements
  - < 150 ms (two way) desired
  - < 400 ms (two way) can be acceptable
  - Audio and video must be in synch
  - Both encoding and decoding of the video bit stream must be in real time.
- Only intermmediate video quality is required.
  - QCIF at 5-10 fps acceptable for video telephony
  - CIF at 10-20 fps satisfactory for video conferencing.
  - Moderate amount of compression artifacts can be tolerated.
- Conference video tends to have limited motion – error can be concealed effectively (more later)

One way video streaming

- Ex. TV Broadcast, live video streaming
- Except for live broadcast, video can be precompressed, but the decoder need to decode in real time.
- Initial playout time can be up to several seconds
  - Receiver uses a large smoothing buffer to store several seconds of video frames before starting to playback.
- Bit rate video varies depending on the applications (typically higher than two-way interactive video applications)
- Recipients of the same video source may connect to network with different access links and may also have different computional power and display capabilities.
  - Scalable coding is desired (more later)

Major types of communication networks

- ISDN:
  - Circuit switching, very reliable, px64 kbps
- Broadband ISDN (ATM):
  - Virtual packet switching using fixed size small cells, quite reliable (cell loss rate 0.000001 -> 0.0001, bandwidth > 155 mbps)
- Internet
  - Datagram packet switching, unreliable with variable delay and loss
  - Packet loss rate depends on network congestion
- Wireless networks
  - Cellular networks
  - Wireless LAN

Video Applications

Application and standard family	Multiplex protocol	Video coding standard	Typical video bit rate	Packet size	Error characteristics
ISDN video phone (H.320)	H.221	H.261 and H.263	64–384 kbps	N/A	Practically error free ($BER = 10^{-10}$–10^{-8})
PSTN video phone (H.324)	H.223	H.263	20 kbps	100 bytes	Very few bit errors and packet losses
Mobile video phone (H.324 wireless)	H.223 with mobile extensions	H.263	10–300 kbps	100 bytes	$BER = 10^{-5}$–10^{-3}, occasional packet loss
Video phone over packet network (H.323)	H.225/RTP/UDP/IP	H.261, H.263, H.262	10–1000 kbps	≤1500 bytes	BER = 0, 0–30% packet losses
Terrestrial/cable/satellite TV	MPEG-2 system	MPEG-2 video	6–12 mbps	188 bytes	Almost error free, depending on weather
Video conferencing over “Native” ATM (H.310, H.321)	H.222.0	H.262	1–12 mbps	53 bytes (ATM cell)	Almost error free ($CLR = 10^{-6}$–10^{-4})

Error Control Techniques for Video

- Transport level error control
  - Error detection and correction using FEC
  - Retransmission of lost packets
- Error resilient encoding
  - Adding redundancy in the video bitstream to help the decoder recover from transmission error
- Error concealment
  - Recover lost/damaged regions based on image/video characteristics and human visual system properties at the decoder
- Encoder-decoder-network interactive error control
  - Adapt the encoder operations based on feedback about damaged/lost frames/blocks.

Transport Level Error Control

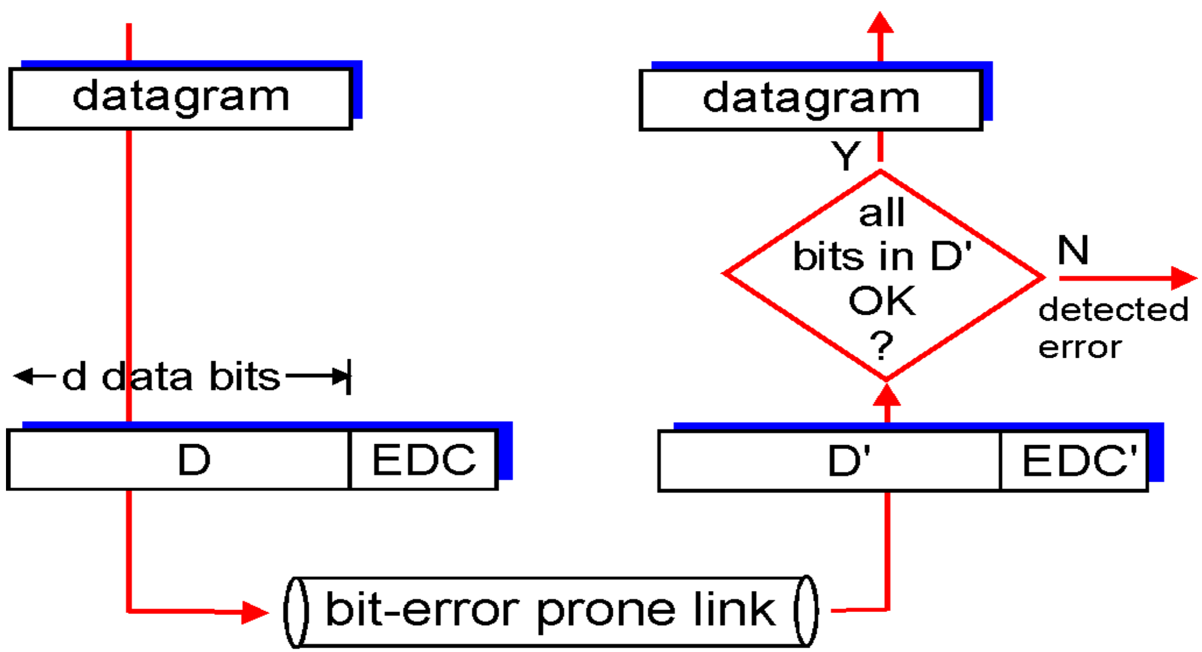
- Forward Error Correction (channel coding)
  - Adding redundancy to the video bit stream for error detection and correction

- Retransmission of lost packet ARQ (automatic retransmision request)

Error Detection

EDC= Error Detection and Correction bits (redundancy)
D = Data protected by error checking, may include header fields

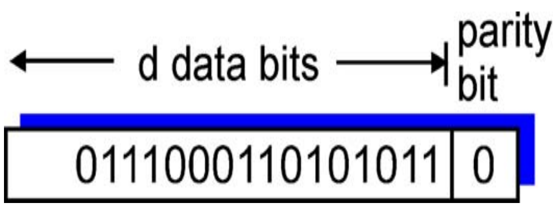
- Error detection not 100% reliable!
  - protocol may miss some errors, but rarely
  - larger EDC field yields better detection and correction



Error Correction

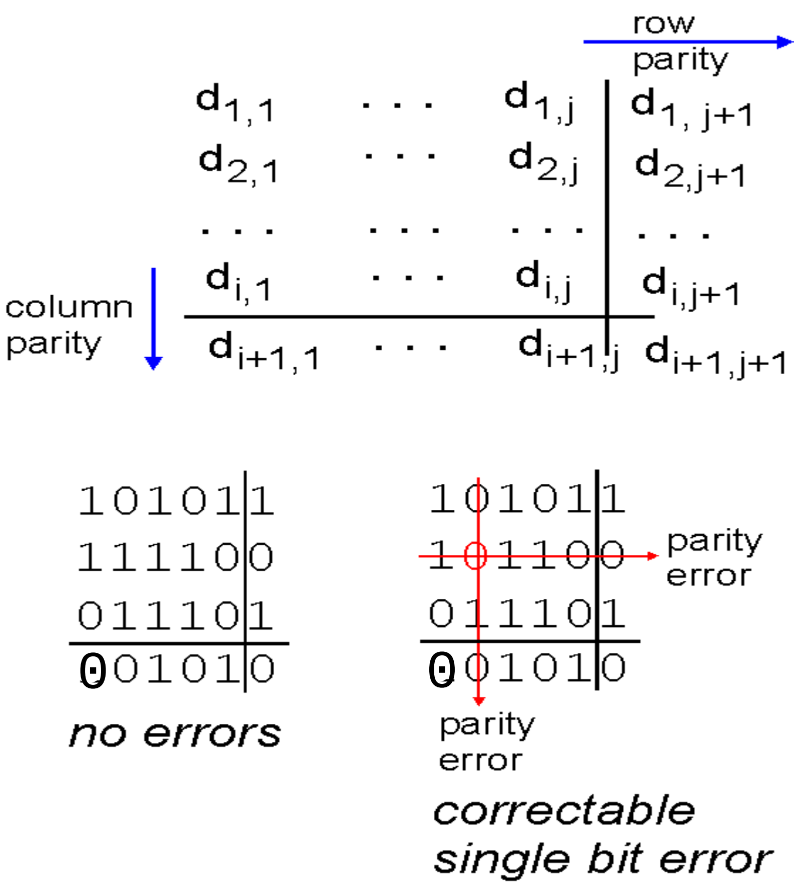
Single Bit Parity:

Detect single bit errors



Two Dimensional Bit Parity:

Detect *and correct* single bit errors

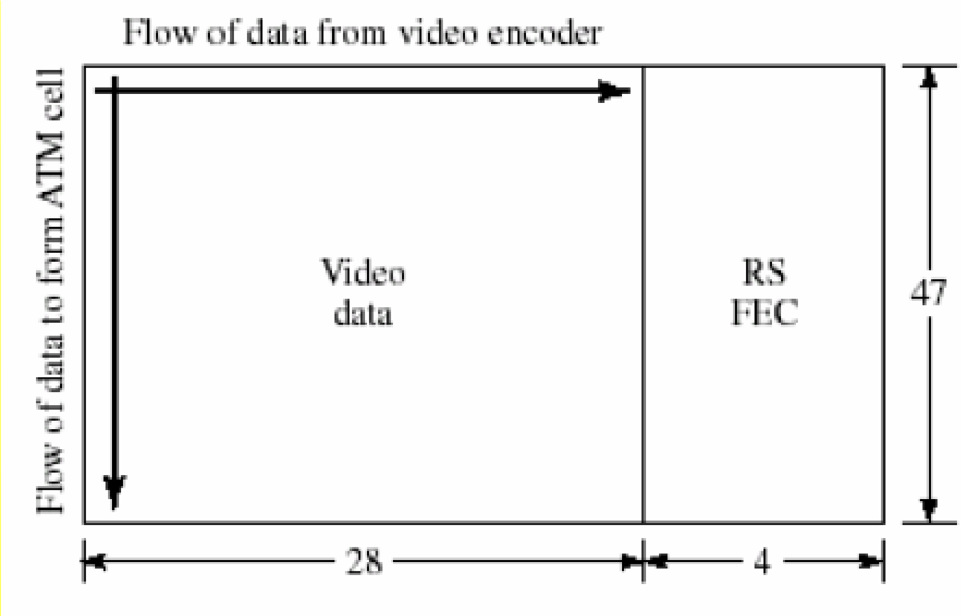


Popular FEC code

- Reed-Solomon code
  - For every n data bits, put in addititional k bits.
  - If the receiver receives any n bits, the entire n data bits can recovered.
- can also use packets in place of bits.

FEC for video transmission

- For wireless networks, FEC is necessary to reduce raw bit error rates
- For the Internet, errors are mainly due to congestion which causes pa ss packets



•Unequal error protection (UEP): using stronger protection for important bit stream (more later)

Delay-constrained ARQ

- ARQ basics:
  - Receiver requests retranmission of lost or erroneoulsy delivered packets
- For delay-insensitive (file transfer) application, ARQ is effective.
- For video applications, ARQ must be carefully designed to meet delay requirements
  - Number of round trip time
  - Apply ARQ to important packets only?
- For broadcast application, ARQ is not appropriate. Why?

Error-Resilient Video Coding

- Basic idea: intentionally insert redundancy in source coding to help recover from transmission errors

- Design goal: miminize the redundancy to achieve a desired level of resilience

- Error isolation (part of H.263/MPEG-4)
  - Insert sync markers
  - Data partition

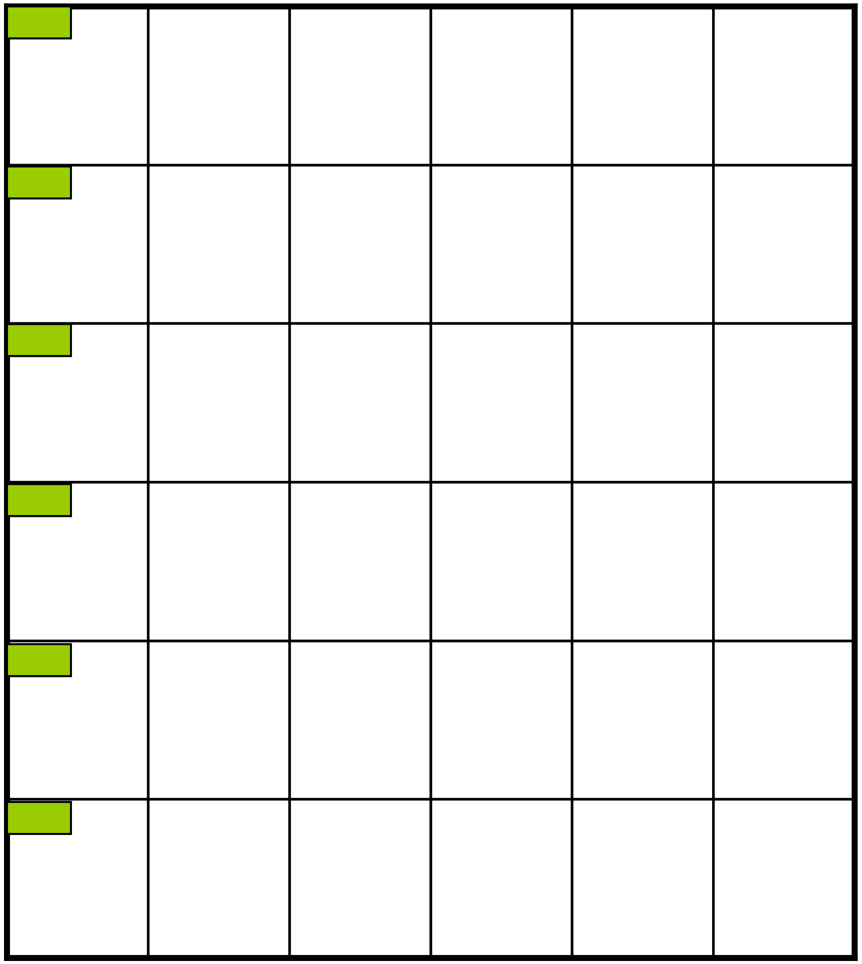
- Robust binary encoding
  - Fixed length code
  - Reversible VLC (RVLC) (part of H.263/MPEG-4)

- Error resilient prediction
  - Insert intra-mode periodically (accommodate by standard)
  - Independent segment prediction (part of H.263/MPEG-4)
- Layered coding with unequal error protection

- Multiple description coding

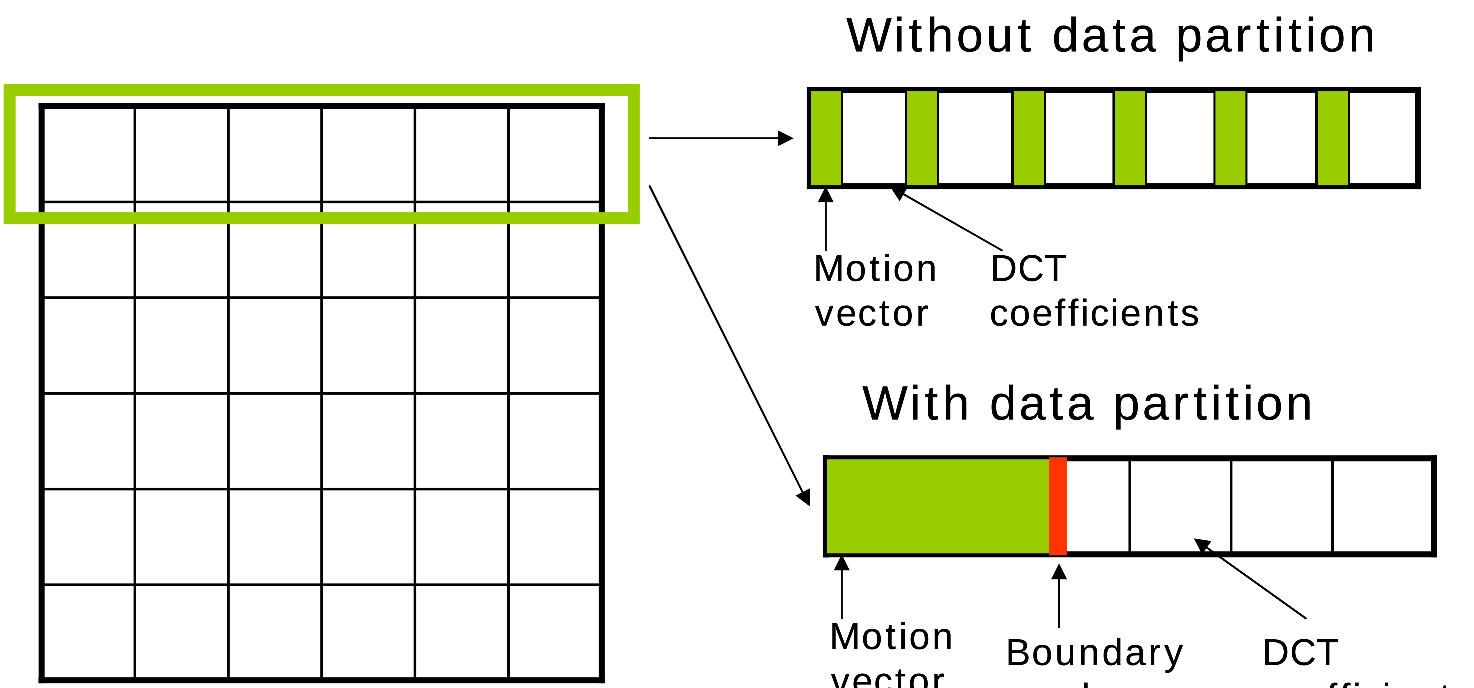
Sync marker (H.263/MPEG-4)

- Insert sync markers
  - Uniquely identifable in the bitstream.
  - Sync markers ensure synchronization at the decoder (additional fields inserted after the sync word for crticial information such as the block address, quantizer value, etc)
  - Sync words may be inserted at various locations, either at a uniform spatial interval in the coded frame or at a uniform bit interval in the bitstream
  - In H.263, sync can be inserted at the beginning of every row.
  - Header extension is duplicated at the slice level.



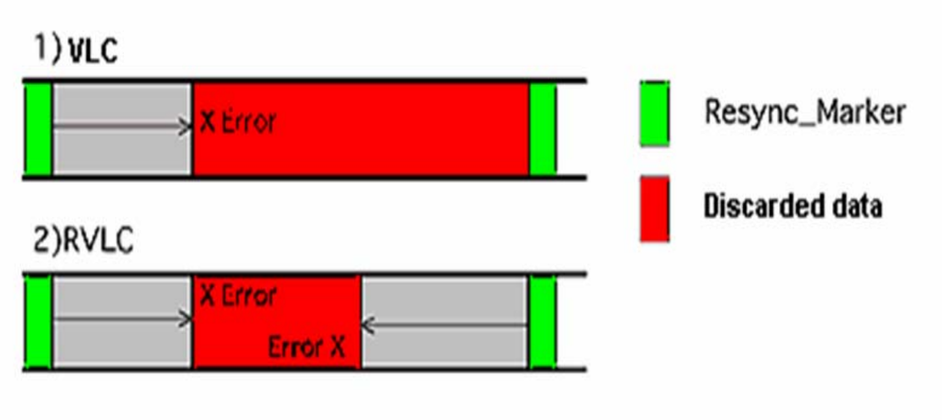
Data partitioning (H.263/MPEG-4)

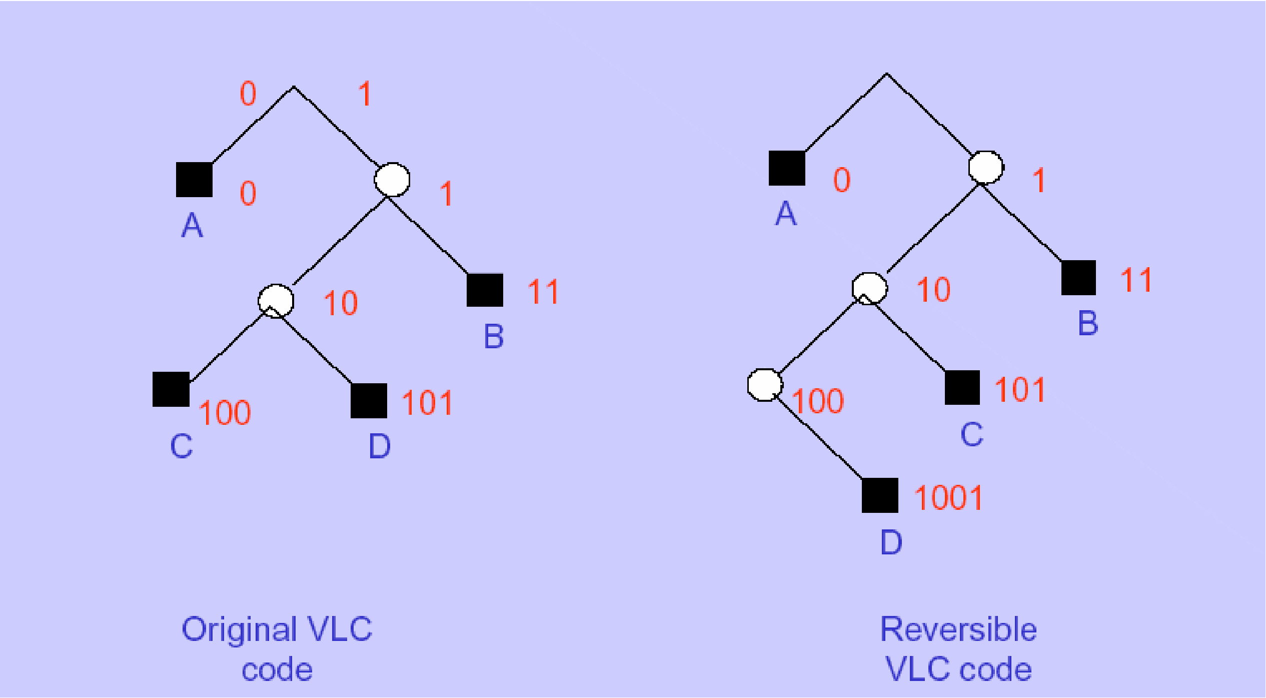
- Without DP, motion vectors and DCT coefficients are grouped together for every macroblock
- With DP, motion vectors and DCT coefficients within a slice are grouped into two partitions separated a boundary marker.
- Advantage: error concealment



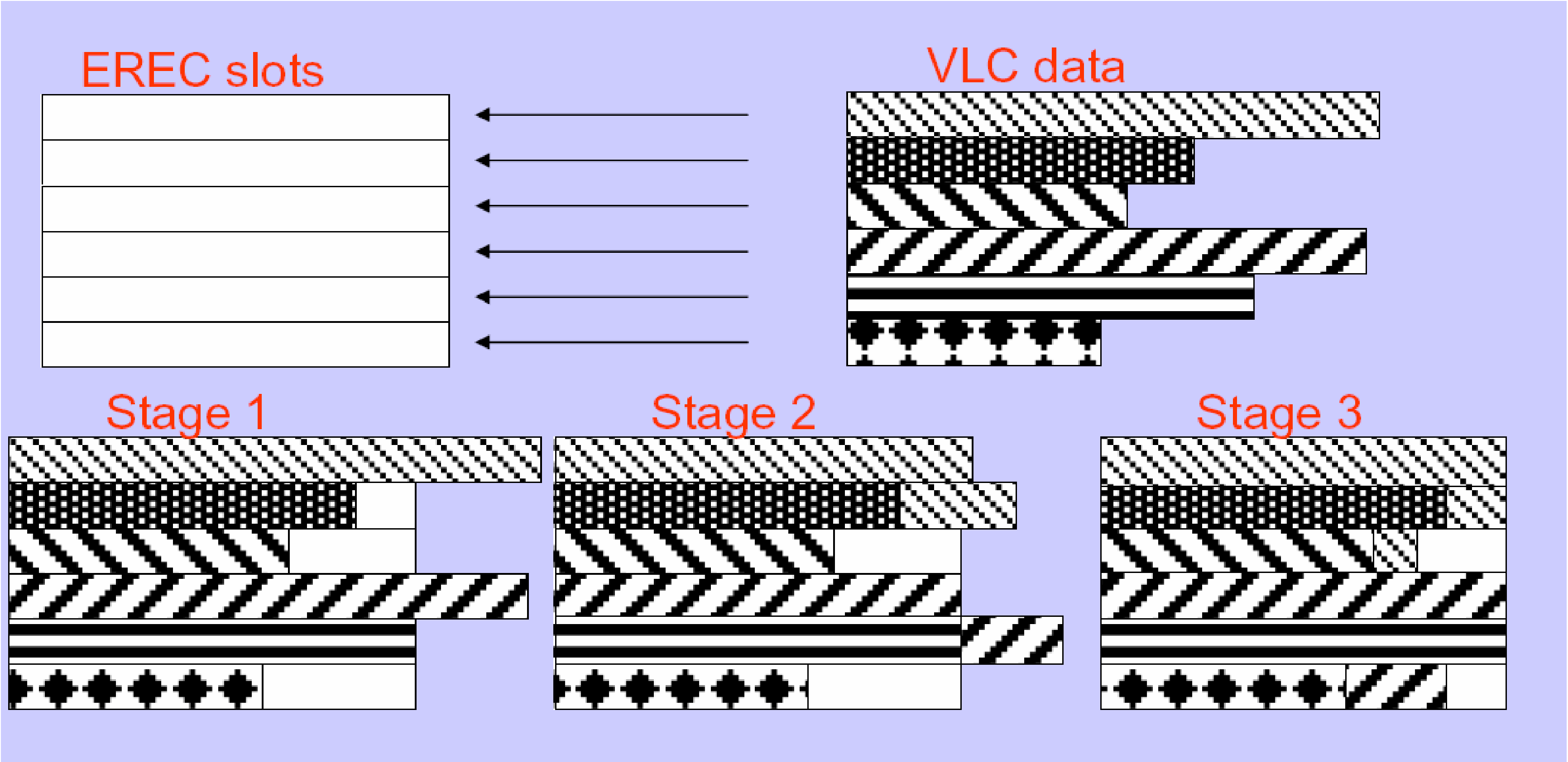
Robust Binary Encoding

- Fixed length code
- Reversible VLC
- EREC (error resilient entropy coding)

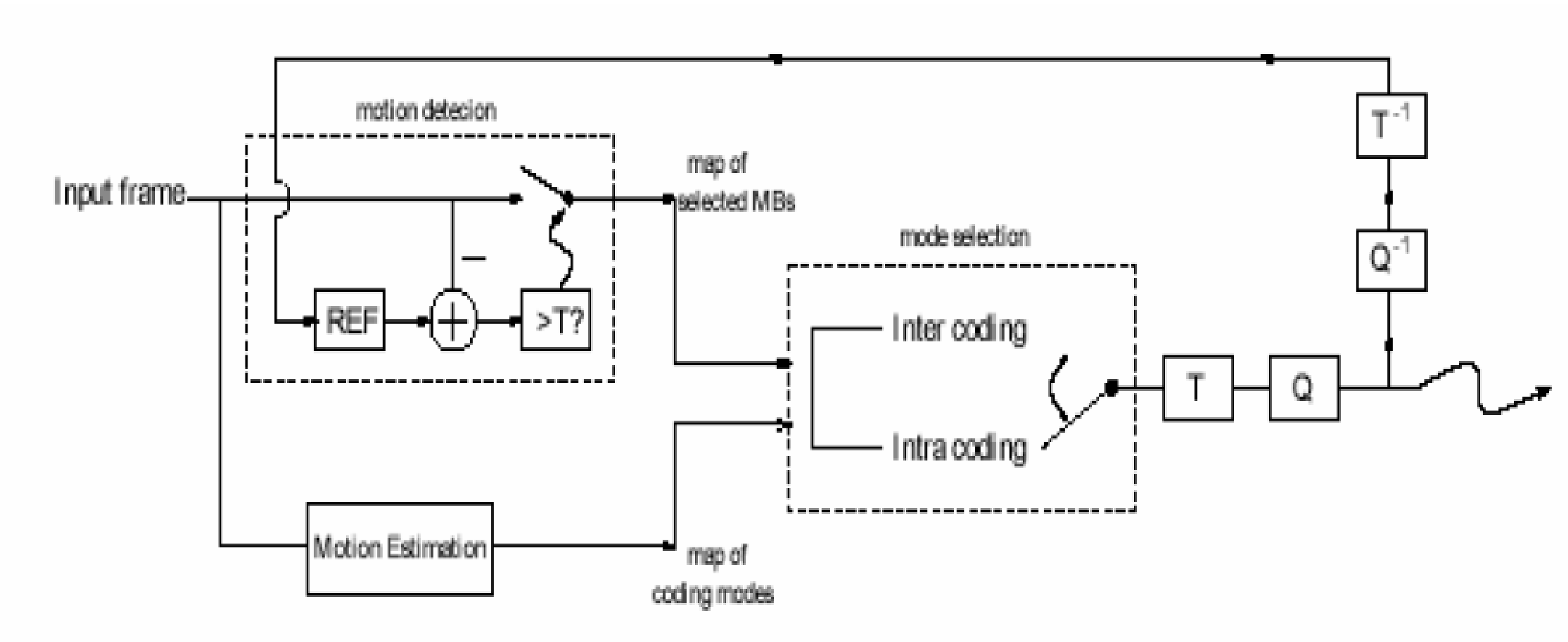




Error-resilient entropy coding



Random intra coding



Encoder-Decoder interactive Error Control

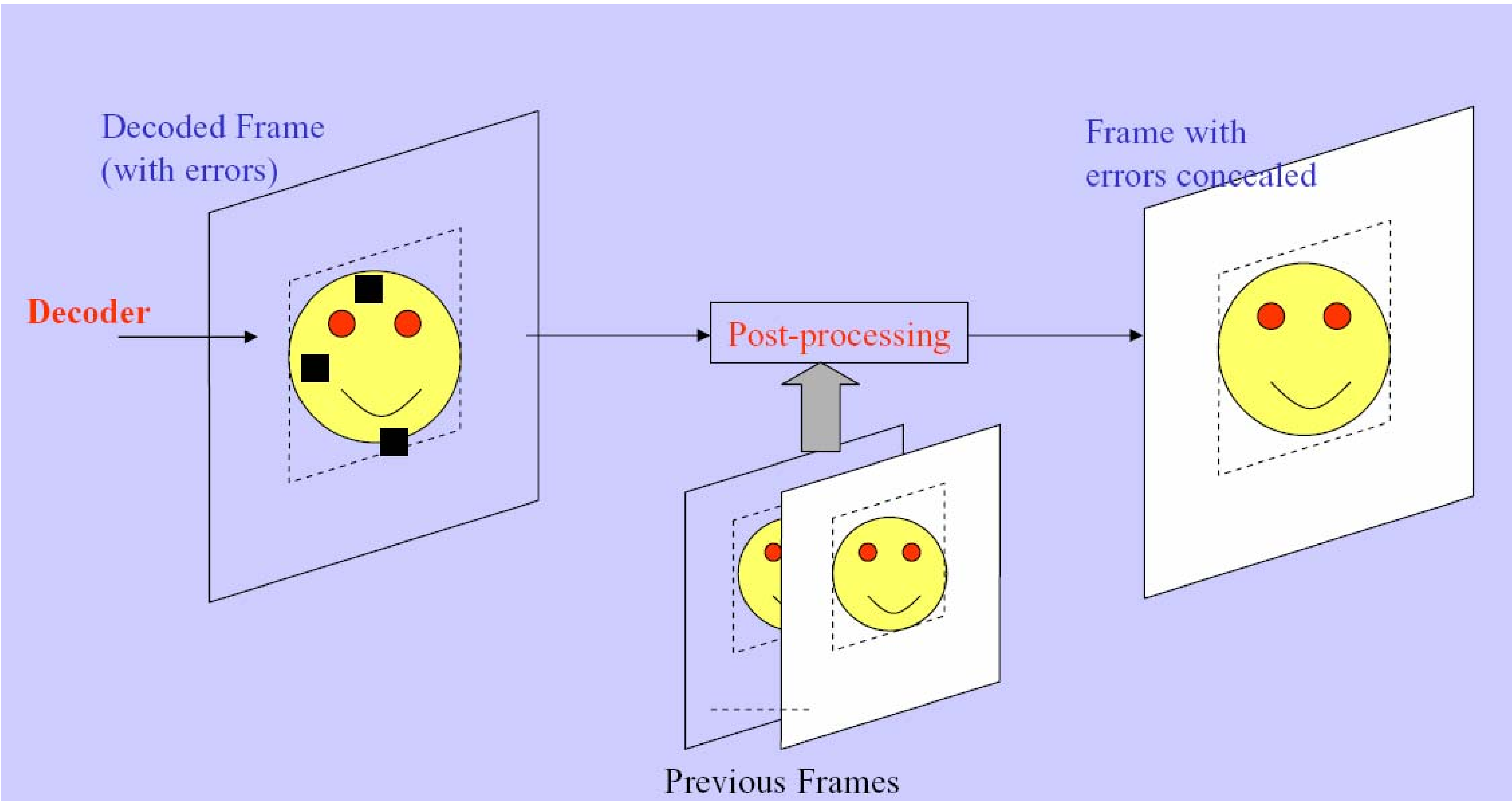
- Coding parameter adaptation based on channel condition
  - Change intra-rate based average loss rates
- Reference picture selection (H.263/MPEG-4)
  - Following a damaged frame (feedback info from receiver), use undamaged previous frame as reference frame for temporal prediction)
- Error tracking
  - Determine which MB are damaged and void using them for prediction
- Require feedback channel

Error Concealment

- Basic idea:
  - Recover damaged regions by interopolating from surrounding regions (in the same frame or nearby frames)

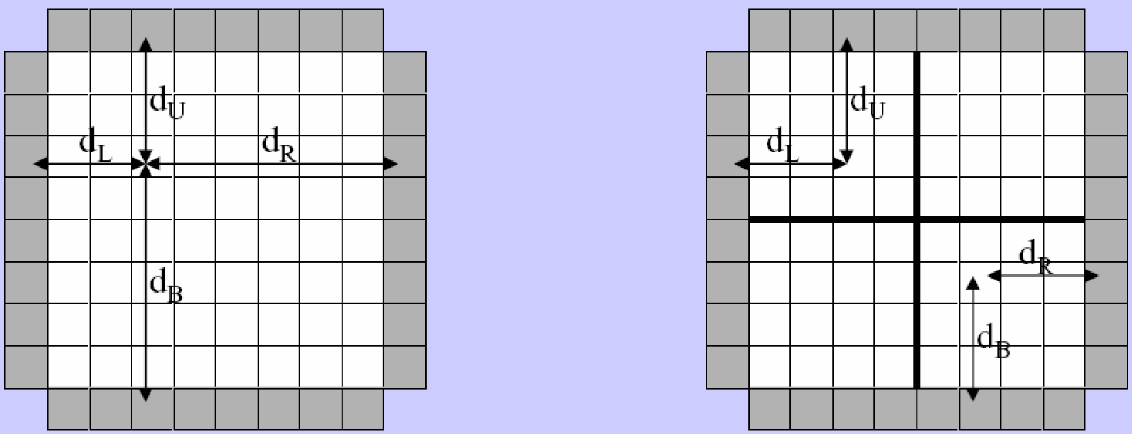
- Not in the standards. Different decoder may implement different error concealment schemes.

Error Concealment



Spatial interpolation

- at the block or macroblock level



Aign and Fazel [1995]

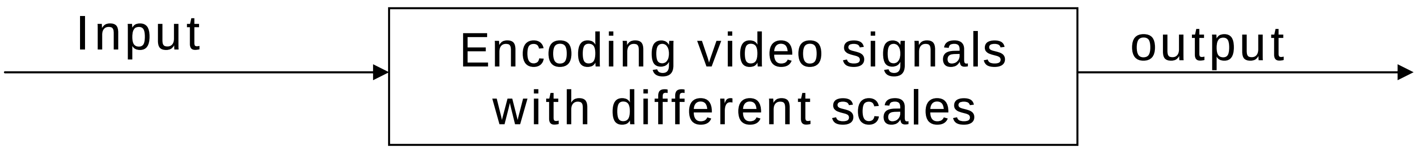
Motion vector interpolation

- Interpolate the motion vectors from the surrounding good blocks
- Use the interopolated motion vectors to obtain the block from the previous frame to replace the lost block in the current frame.

Which solutions?

- Channel feedback
- Channel characteristics
- QoS requirement
- Complexity at encoder/decoder

Scalable Video Coding

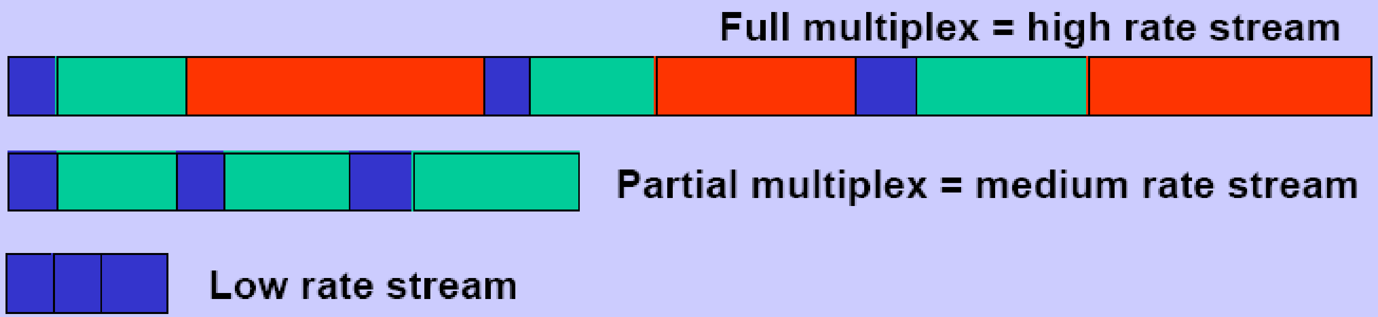


Different Methods:

- SNR scalability: reduce bit rate by introducing quantization error
- Spatial scalability: reduce bit rate spatially subsampling the image
- Temporal scalability: reduce bit rate by reducing frame rate
- Content scalability: Select certain the content for coding.

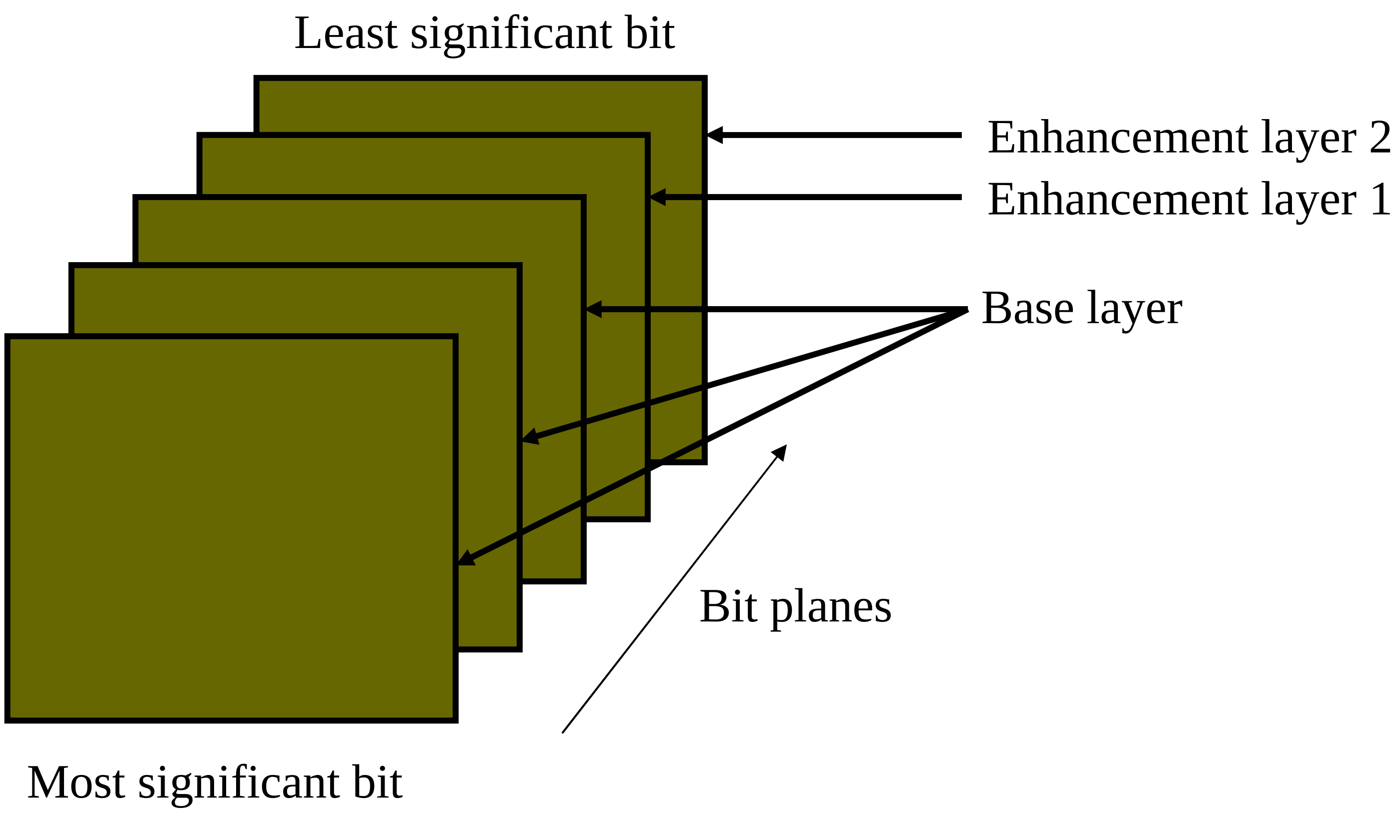
Layered Coding

- Layered coding supports embedded streams
  - Re-configuration of bit stream for reconstruction with different spatial/temporal/quality resolution

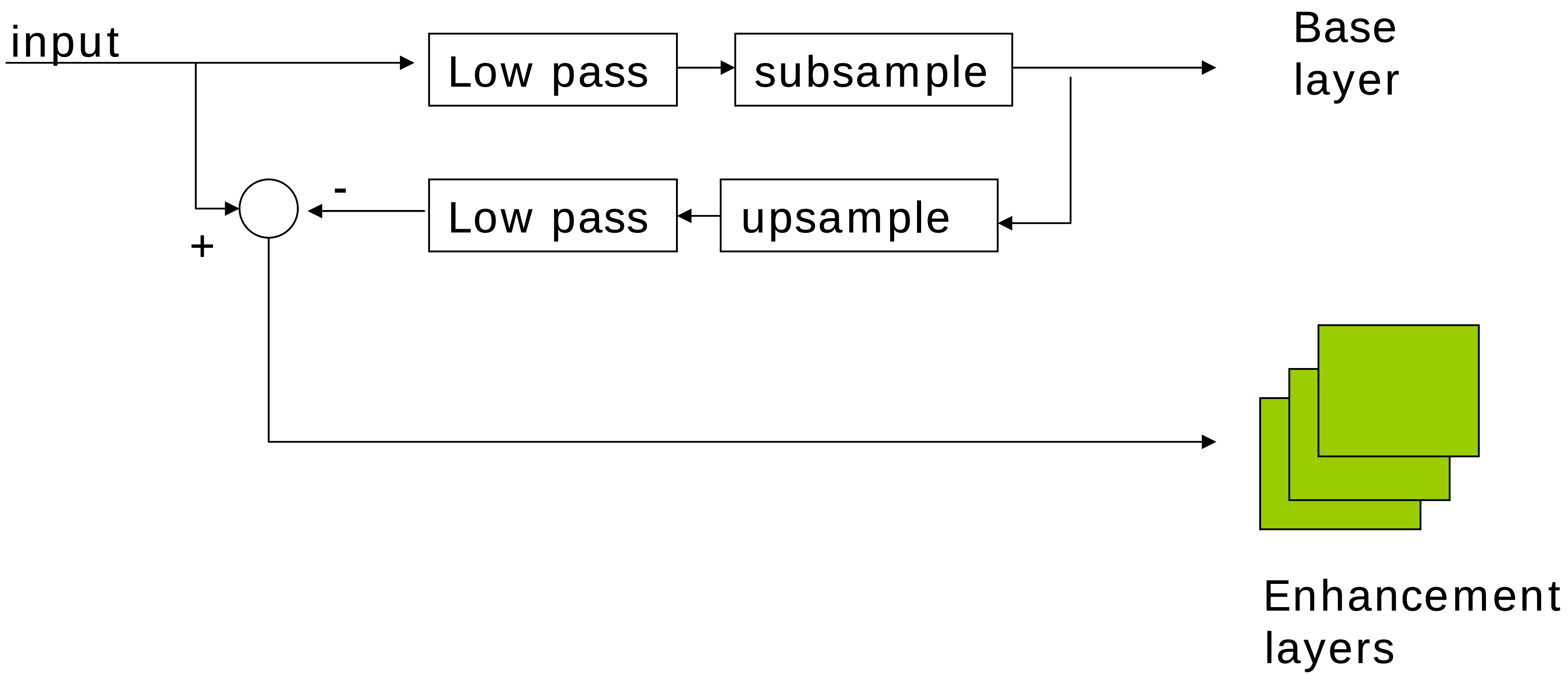


- Possible loss in efficiency depends on coding scheme
- In theory, arbitrary number of scales could be achieved

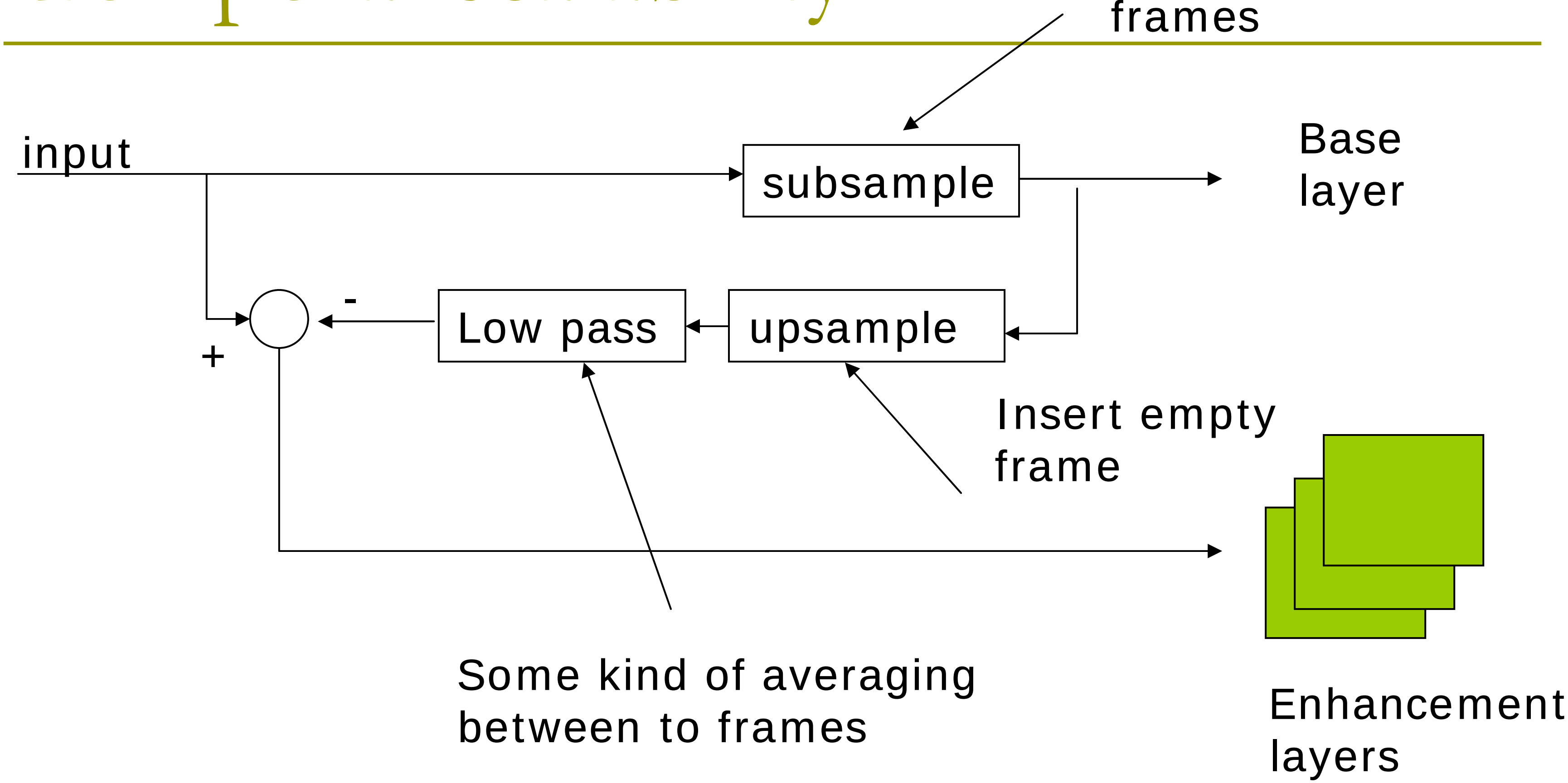
SNR Scalability



Spatial Scalability



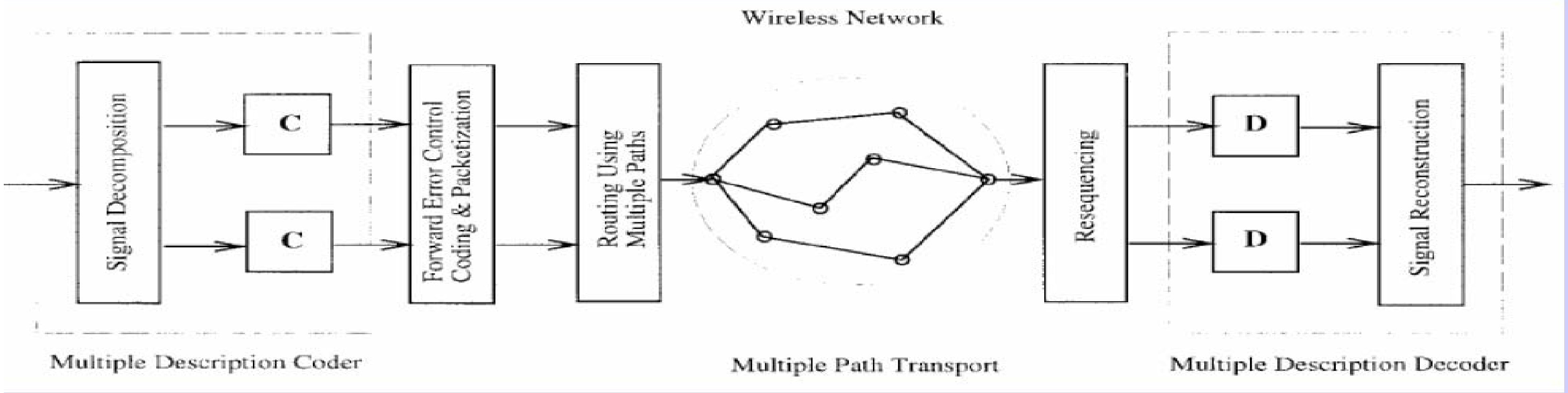
Temporal scalability



Multiple Description Coding

Multiple Description Coding

- Set of techniques to improve the robustness of video to losses by creating correlated coded representations of the video and transmitting them on separate channels.



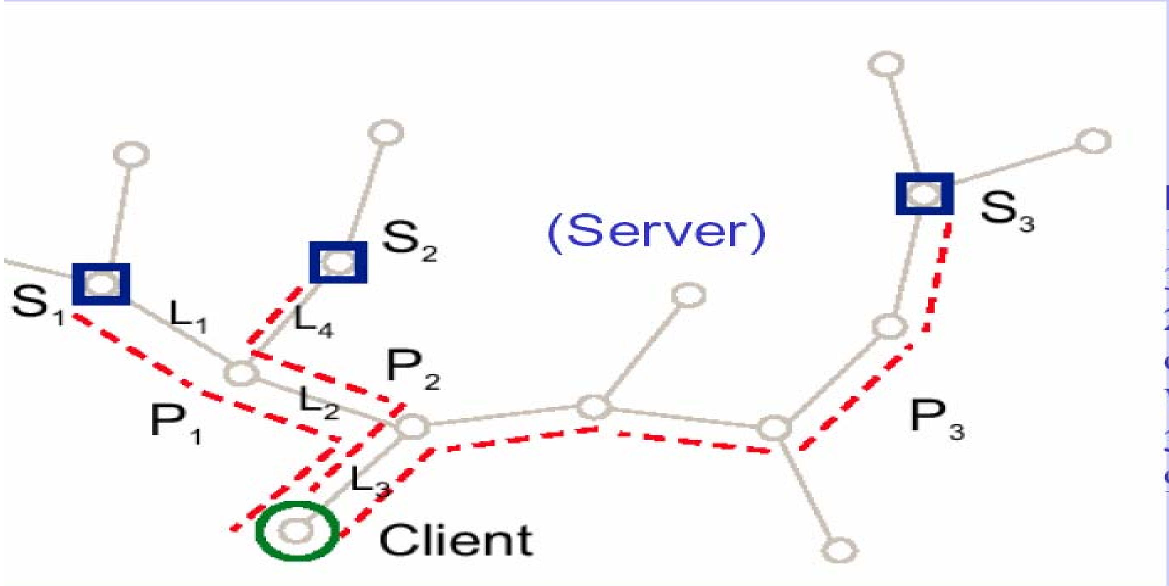
Multiple Description Coding

Multiple Description Coding

- When useful?
  - One or more users sometime fail to receive one or more descriptions
  - Various quality levels are acceptable and distinguishable
- What to use on?
  - Multimedia
    - Audio, Images, Video, Graphics
- Applications
  - Packet networks
    - Internet, wireless
  - Distributed storage and transmission
  - Peer-to-peer networks
  - Frequency-hopping wireless systems
  - Hybrid digital broadcast

Multiple Description Coding

Application Scenario: Content Delivery Networks (CDNs)



L = Link:
1. Wireless: GSM, GPRS, 3G, wireless LAN, bluetooth
2. Wired: dial-up, ISDN, cable, xDSL, fiber, LAN, WAN
3. Connection bandwidth 9.6kbps –100Mbps

Client = PDA; laptop; desktop; set-top box; TV; HDTV

Using CDNs and exploiting path and server diversity for heterogeneous networks with heterogeneous devices

Multiple Description Coding

Media Content Delivery Networks (CDN)

- CDN : set of multiple servers that are managed together
- Steps for Media streaming over CDNs
  - Use MDC to code media stream
  - Distribute multiple descriptions across servers in CDN.
  - When a client requests the media, select multiple servers with complementary descriptions of the stream.
  - Stream from these multiple servers through different paths to client
- Issues
  - Server placement problem
  - Content distribution problem
  - Server selection problem

Multiple Description Coding

Multiple Description Coding

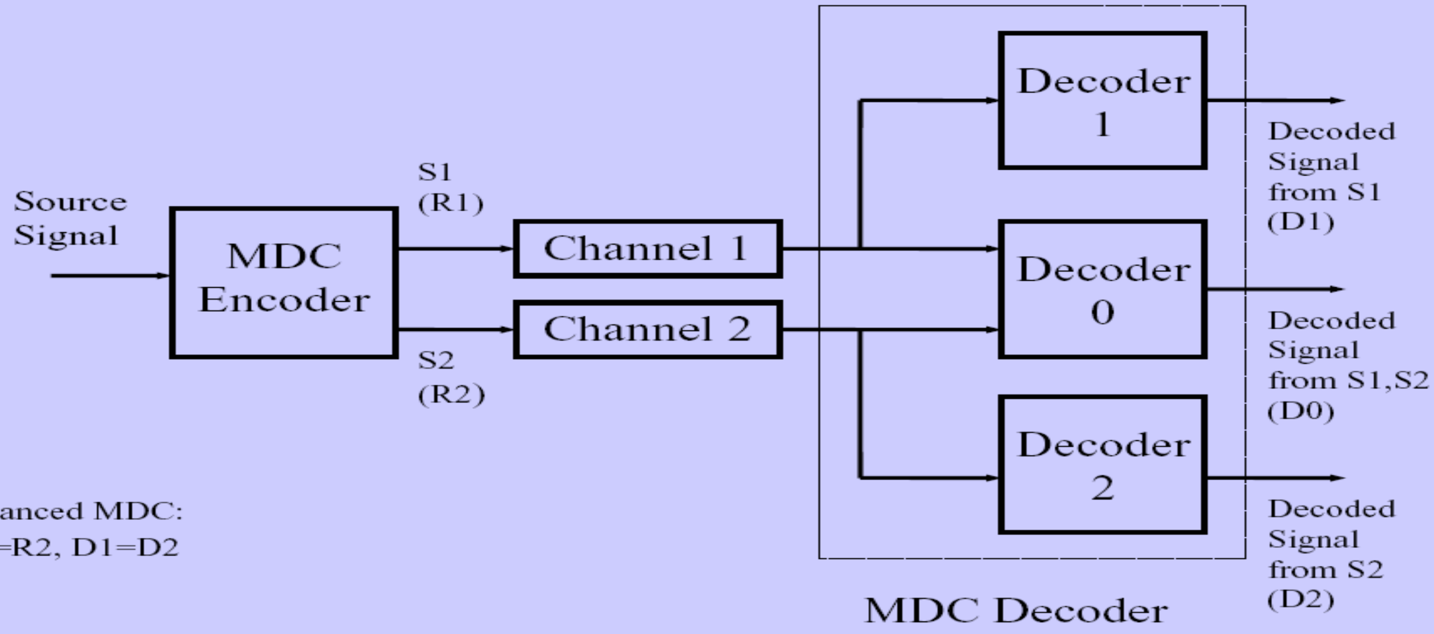
Assumptions:
- multiple channels between source and destination
- independent error and failure events
- probability that all channels fail simultaneously is low
- good model for the Internet and wireless networks
  - when data properly packetized and interleaved

Generate multiple correlated descriptions
- any description provides low but acceptable quality
- additional descriptions provide incremental improvements
- No retransmission required (low delay)
- Reduced efficiency

Design goal:
- maximize the robustness to channel errors at a permissible level of redundancy

Multiple Description Coding

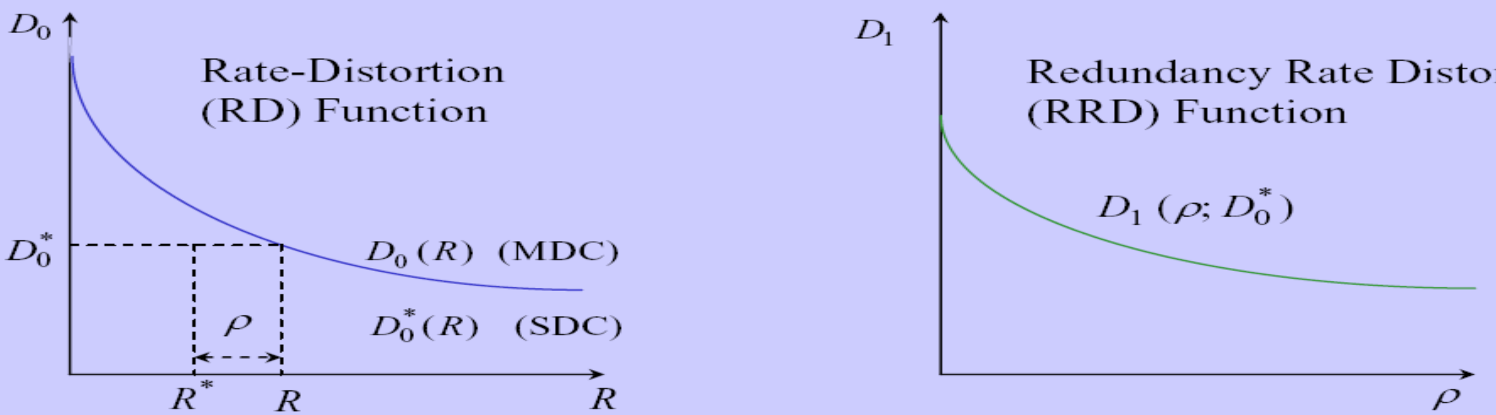
Example: Two Description Coder



Balanced MDC:
R1=R2, D1=D2

Multiple Description Coding

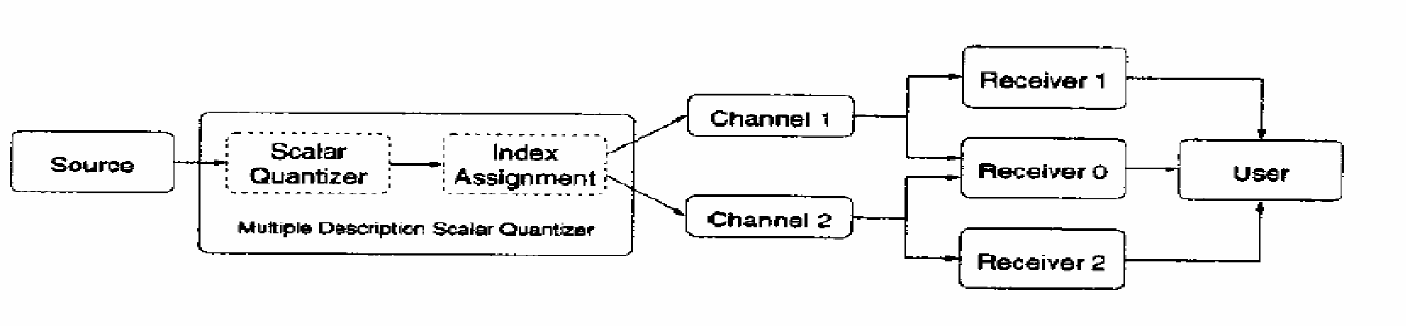
Redundancy Rate Distortion

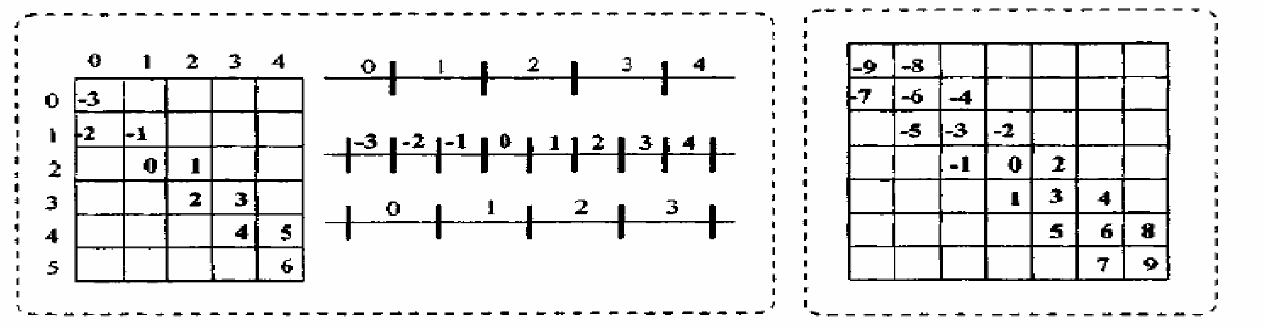


- Design criteria for MD coders
  - Minimize D_1 for a given ρ , for fixed R^* or D_0^* (minimizing the average distortion given channel loss rates, for given total rate)
  - Can vary the ρ vs. D_1 trade-off to match network conditions

Multiple Description Coding

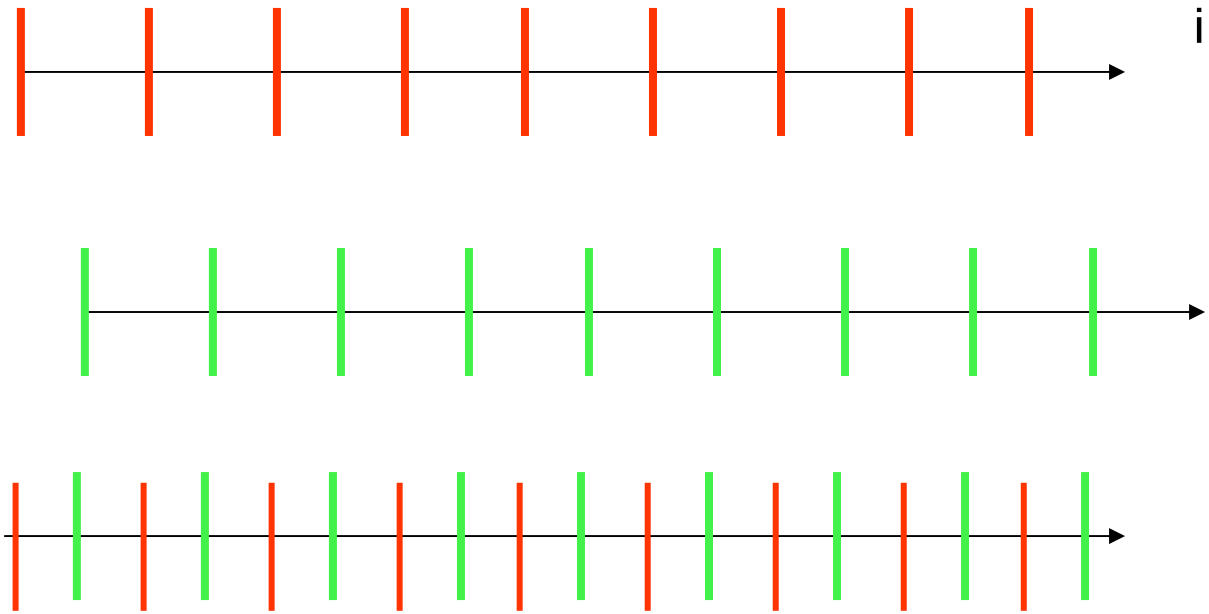
MD Quantization





Multiple Description Coding

Where is the redundancy?



3 bits = 8 intervals

3 bits = 8 intervals

6 bits

But we can use only 4 bits = 16 intervals

Multiple Description Coding

MD Correlating Transforms

Correlating Transform adds redundancy between transform coeffs making them easier to estimate when some are lost

Consider x_1 and x_2: independent Gaussian transform coeffs.

Average error when only one is received $\frac{1}{4}(\sigma_1^2 + \sigma_2^2)$

Instead if we transform these into y_1 and y_2

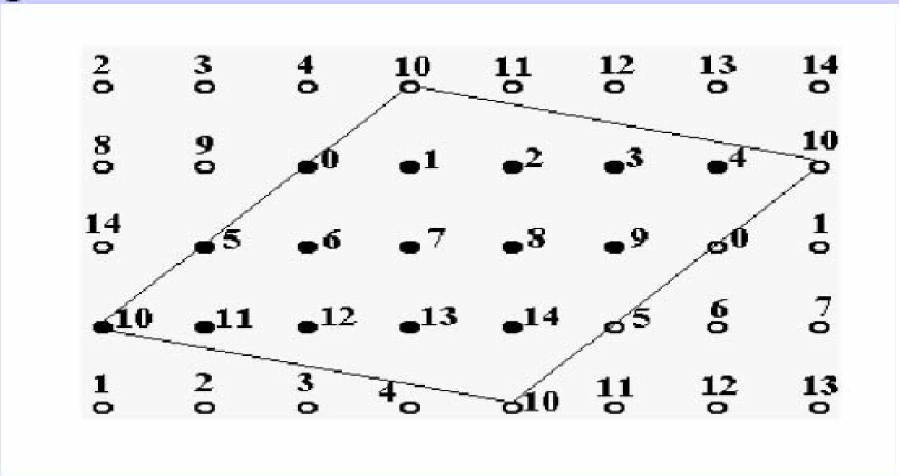
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \theta & (2\theta)^{-1} \\ -\theta & (2\theta)^{-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Average error when only one is received ranges between $\frac{1}{4}\sigma_1^2$ and $\frac{1}{4}\sigma_2^2$

Multiple Description Coding

MD Splitting: Spatial

How to partition space to allow best recovery of coefficients?
Analogous to sphere packing problem:
Maximize the minimum distance between samples within each partition



Example: 15 partitions

Multiple Description Coding

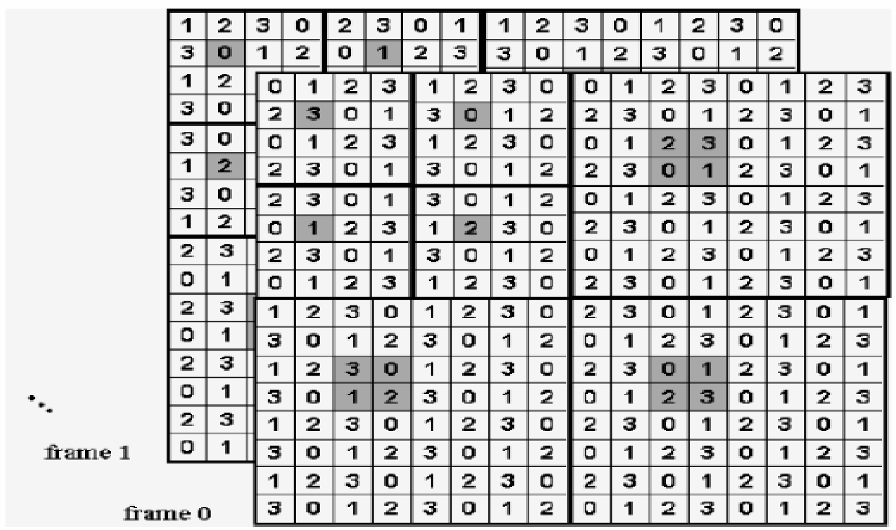
- **Wavelet-based MDC**

0	1	2	3	1	2	3	0	0	1	2	3	0	1	2	3
2	3	0	1	3	0	1	2	2	3	0	1	2	3	0	1
0	1	2	3	1	2	3	0	0	1	2	3	0	1	2	3
2	3	0	1	3	0	1	2	2	3	0	1	2	3	0	1
2	3	0	1	3	0	1	2	0	1	2	3	0	1	2	3
0	1	2	3	1	2	3	0	2	3	0	1	2	3	0	1
2	3	0	1	3	0	1	2	0	1	2	3	0	1	2	3
0	1	2	3	1	2	3	0	2	3	0	1	2	3	0	1
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3

Packetize coefficients into different descriptions
Disperse samples from the spatial orientation tree across descriptions

Multiple Description Video Coding

Wavelet video based MDC



Packetization of coefficients from successive frames

Multiple Description Video Coding

- **Comparisons with one description**

17.5 dB
0.48 bpp

Split transform blocks

29.9 dB
0.57 bpp

MD Correlating Transform

31.9 dB
0.68 bpp

MD Quantizer

26.5 dB
0.59 bpp

Split Transform Coefficients