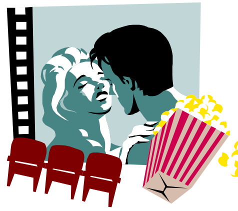
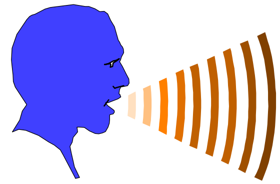





# Video Processing & Communications

## Error Control in Video Communications

Yao Wang
Polytechnic University, Brooklyn, NY11201
http://eeweb.poly.edu/~yao

Based on: Y. Wang, J. Ostermann, and Y.-Q. Zhang, Video Processing and Communications, Prentice Hall, 2002.

# Outline

- Necessity/challenge for error control
- Characteristics of typical applications and networks
- Overview of techniques
- Error resilient encoding
- Error concealment
- Encoder/decoder interactive error control

# Steps involved in a Communication Session

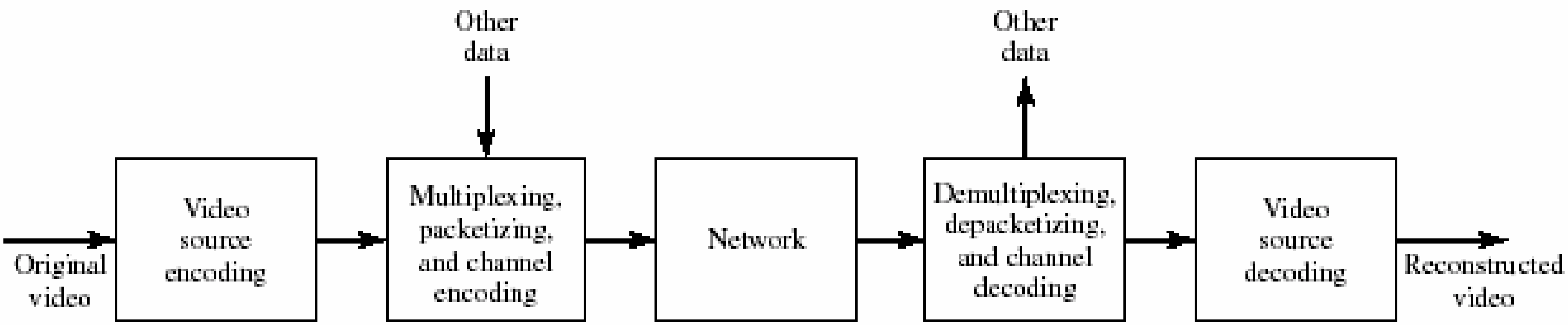
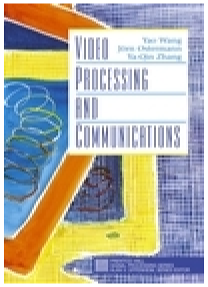



**Figure 14.1** A typical video communication system.




# End-to-End Delay

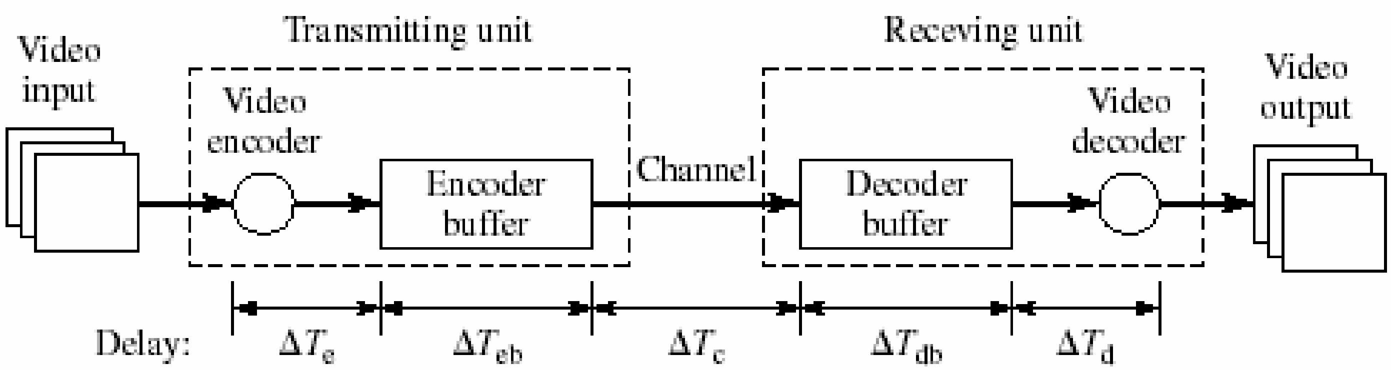
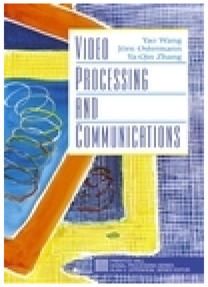
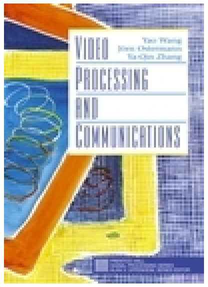



**Figure 14.4** Factors contributing to the end-to-end delay in a video communication system. Adapted from A. Ortega and K. Ramchandran, Rate-distortion methods for image and video compression, *IEEE Signal Processing Magazine* (Nov. 1998), 15:23–50. Copyright 1998 IEEE.




# Challenge for Video Communications

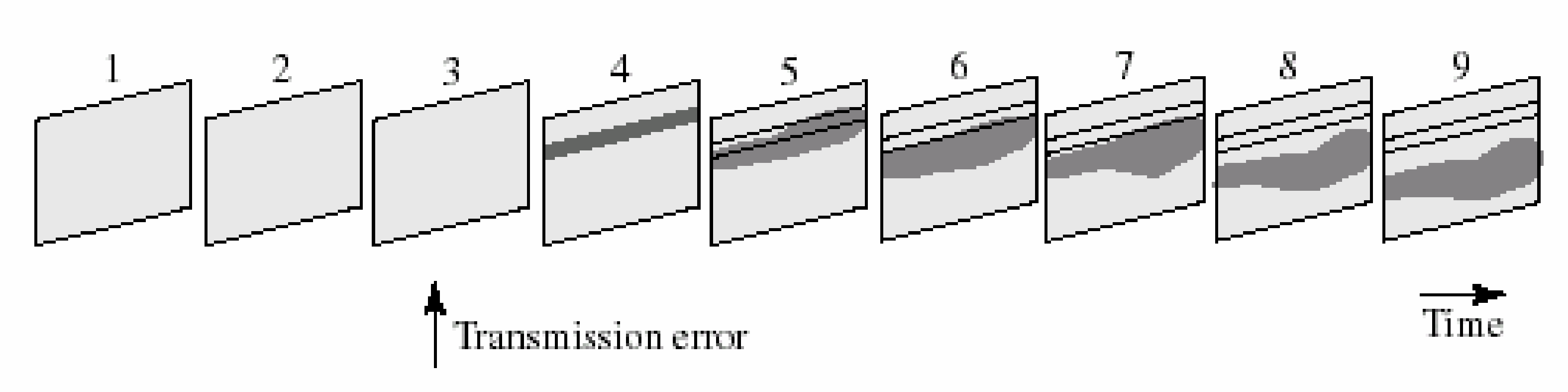
- Real networks are unreliable ☹
  - Wireless networks: random bit errors, long burst errors, and possibly link downs
  - Internet: packet loss and variable delay due to network congestion
  - Excessive delay = loss for real-time applications
- Real networks are heterogeneous in bandwidth and reliability
- Video data are delay-sensitive
  - One cannot rely on retransmission for error control because of the stringent delay requirement!




# Conventional source coding technique is not good enough ☹

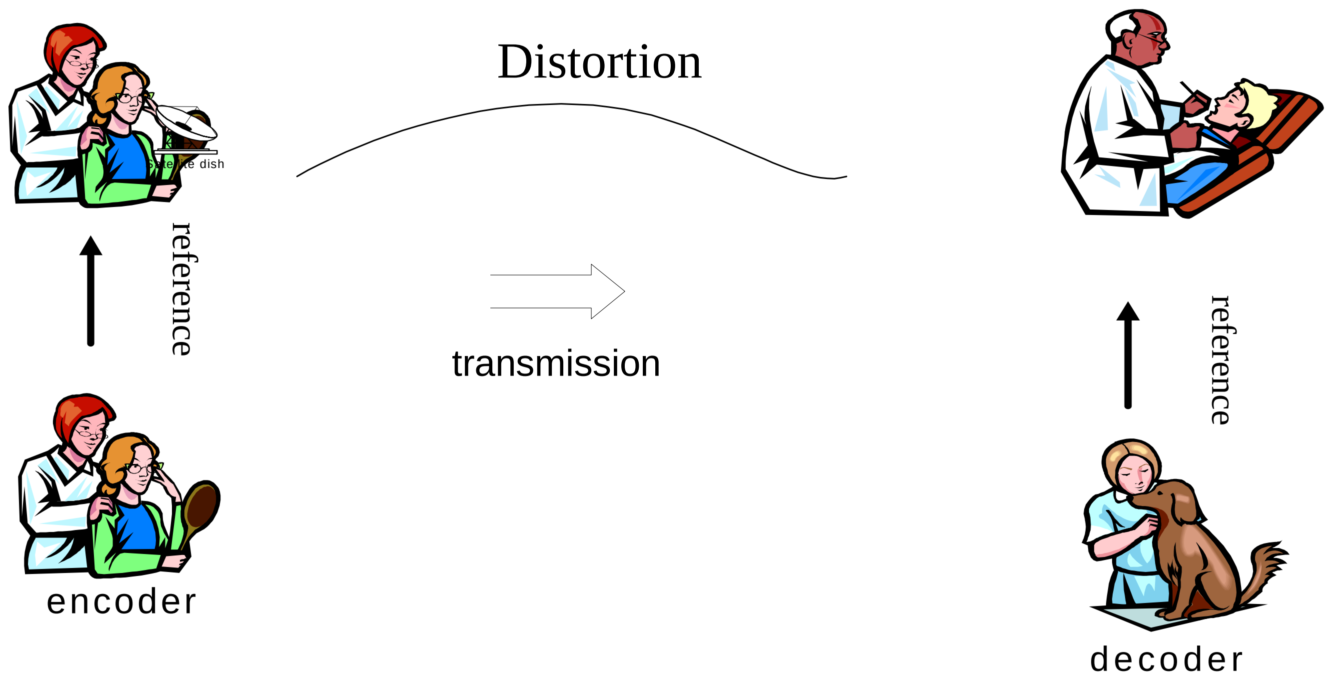
- Optimal performance only for fixed rate and perfect channel
- Have poor reconstruction quality when parts of coded data are lost
- Compressed video data is very sensitive to transmission errors
  - Variable length coding
  - Temporal predictive coding
  - Spatial predictive coding
  - All contribute to error propagation within the same frame as well as in following frames: 1 bit error or packet loss can render following received data useless

# Spatial/Temporal Error Propagation



# Drift (Reference Mismatch) Problem

- Motion compensated temporal prediction should be retained to preserve the coding efficiency
- Loss in a previous frame can cause mismatch between the reference frame used in the encoder and that in the decoder
  - Encoder and decoder out of sync



# Effect of Transmission Errors

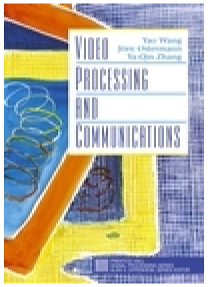
Coded, No loss

3%

5%

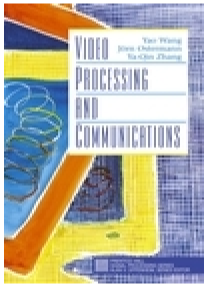
10%

Example reconstructed video frames from a H.263 coded sequence, subject to packet losses



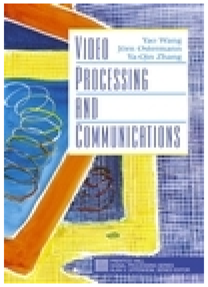

# QoS Requirements of Typical Video Applications

- Interactive two-way visual communications
- One-way video streaming
- One-way video downloading
  - No difference from file downloading



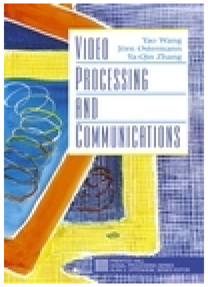

# Interactive two-way visual communications

- Ex. Teleconferencing, video telephony, virtual classroom
- Very stringent delay requirement
  - <=150 ms (one way) desired
  - 150-400 ms can be acceptable
  - >400 ms not acceptable
  - Audio and video must be in sync to maintain lip sync.
  - Both encoding and decoding must be completed in real-time.
- Only low to intermediate 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/transmission artifacts can be tolerated.
- Raw video has limited motion -> easier to code and conceal errors




# One-Way Video Streaming

- Ex. TV broadcast, Multicast of a conference/event, Video streaming from Internet
- Except for live broadcast/multicast, can pre-compress the video, but decoding must be done in real-time
- Initial playout delay can be up to a few seconds
  - Receiver uses a large smoothing buffer to store several seconds of video frames before starting to display the first received frame
- Bit rate/video quality can vary widely depending on the applications
- Recipients of the same video source may be connected to the network with different access links (e.g. wireless modem to 100 mbps fast ethernet) and the receiving terminal may have varying computing power (palm vs. laptop vs. desktop)
  - Scalable coding desired

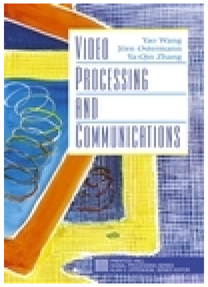



# Major Types of Communications Networks

- ISDN:
  - circuit switching, very reliable, px64 kbps
- Broadband ISDN (ATM):
  - virtual packet switching using fixed size small cells, quite reliable (cell loss rate 10^{-6}-10^{-4}), >=155 mbps
- Internet
  - Datagram packet switching, unreliable with variable delay and packet loss
  - Packet loss rates and delay depend on network congestion
  - New protocols have been developed for real-time data transport (RTP, RSVP, RTSP, SIP, etc.)
  - Most widely used Ethernet LAN has rate 10/100 mbps
- Wireless networks
  - Cellular networks -> 3G/4G
  - Wireless LAN (IEEE 802.11b: 11 mbps, 802.11a: 56 mbps)

# Characteristics of Major Video Communications 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 | $\leq$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 = $10^{-6}$–$10^{-4}$) |

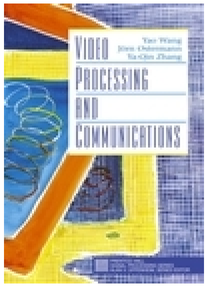



# Error Control Techniques for Video

- Transport level error control
  - Error detection and correction through FEC
  - Retransmission of lost packets
- Error resilient encoding
  - Adding redundancy in the 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
  - Joint source/channel coding
    - Ex: layered coding with unequal error propagation
  - Feedback-based adaptive encoding
    - Ex. Reference picture selection, Selective intra update
  - Path diversity
    - Different bitstreams sent through separate paths

# Transport Level Error Control

- Forward Error Detection and Correction (Channel Coding)
- Retransmission (Automatic Retransmission Request or ARQ)
- Error resilient packetization and multiplexing
- Unequal error protection

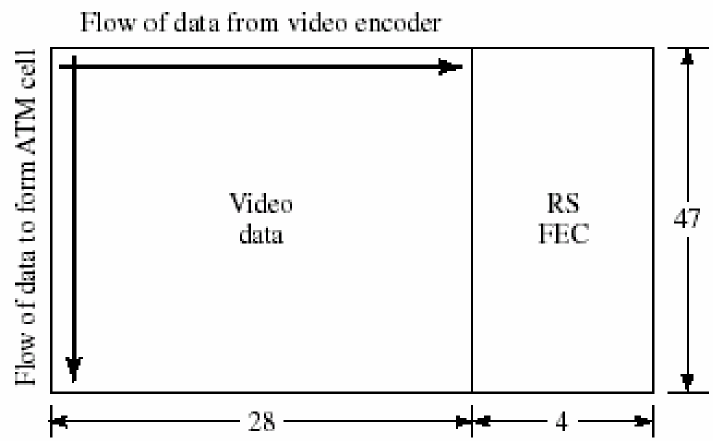



# Channel Coding Basics

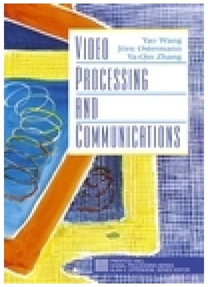
- Channel coding (FEC)
  - Adding redundancy bits on compressed source bits to enable error detection and correction
  - Simple example: Add a parity check bit at the end of a block of datastream, can detect all single bit errors
  - Channel coding rate:
    - For every k source bits, add l channel bits, to create n=k+l bits -> channel coding rate r=k/n
    - Well designed code (e.g. Reed-Solomn code) can correct t=l/2 error bits in each n-bit block
- Shannon theorem for communication:
  - Source and channel codes can be designed separately:
    - Source coding minimizes the bit rate necessary to satisfy a distortion criterion (Shannon rate-distortion theory)
    - Channel coding adds just enough redundancy bits to reduce the raw channel error rate to the permitted level
  - Only valid for stationary source and channel and requires processing of infinitely long blocks of data (delay = infinity!)

# FEC for Video Transmission

- For wireless networks, FEC is necessary to reduce raw bit error rates
- For Internet, errors are mainly due to congestion-caused packet losses, FEC can be applied across packets to correct/detect packet losses




- Unequal error protection (UEP): using stronger channel codes (r=k/n smaller) for more “important” bitstreams (base-layer). Best implemented with RCPC (rate compatible punctured convolutional) code.

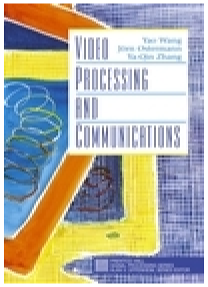



# Delay-Constrained ARQ

- ARQ Basics:
  - receiver requests retransmission of a lost or erroneously delivered packet, incorporated in TCP
- For data transmission, ARQ is an effective mechanism for error control
- For video applications, ARQ must be limited to within the delay constraint of the application
  - How many retransmission attempts are acceptable depends on the round-trip time (RTT)
  - Should only apply ARQ to “important” packets (base-layer) (another way to achieve UEP)
- For broadcast/multicast applications, ARQ is inappropriate in general, although it can be deployed at the link layer

# Error-Resilient Encoding

- Basic idea: intentionally insert redundancy in source coding to help recover from transmission errors
- Design goal: minimize the redundancy to achieve a desired level of resilience
- Error isolation (part of H.263/MPEG4 standard)
  - Inserting sync markers
  - Data partition
- Robust binary encoding
  - Reversible VLC (RVLC) (part of H.263/MPEG4 standard)
- Error resilient prediction
  - Insert intra-mode periodically (accommodated by the standard)
  - Independent segment prediction (part of H.263/MPEG4 standard)
- Layered coding with unequal error protection
- Multiple description coding




# Reversible Variable Length Coding

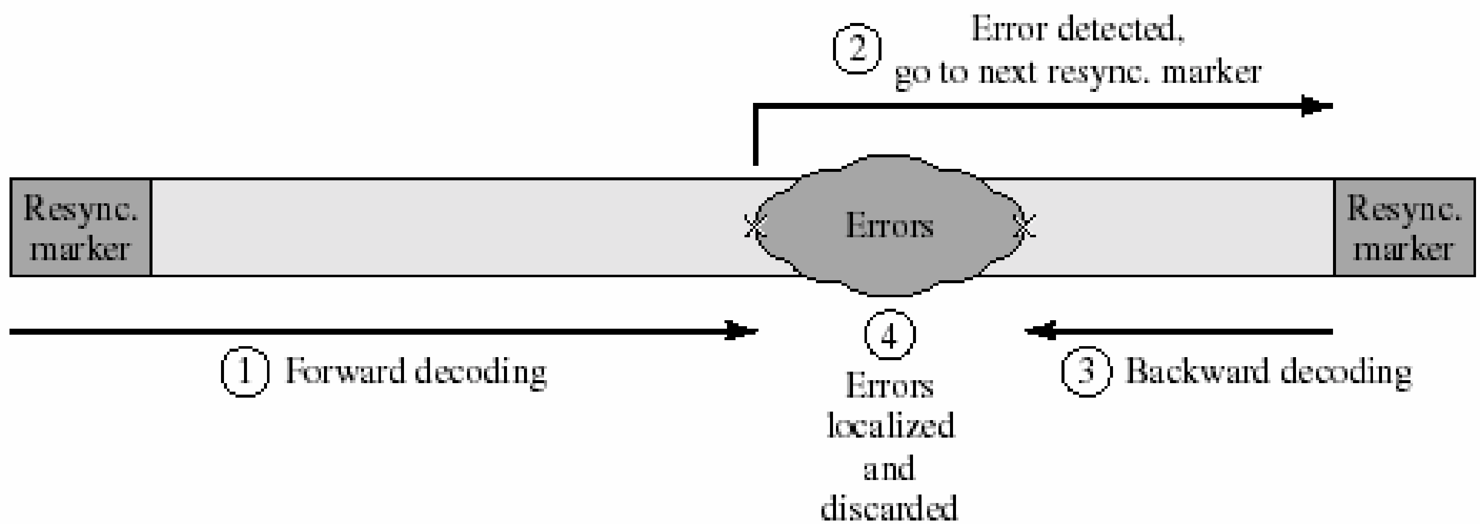
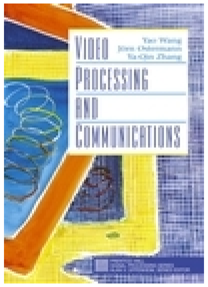


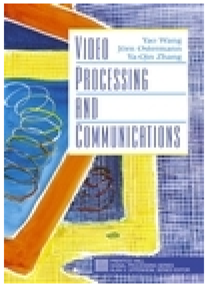

**Figure 14.7** RVLC codewords can be parsed in both the forward and backward direction, making it possible to recover more data from a corrupted data stream.




# Mode Selection Based on Network Conditions

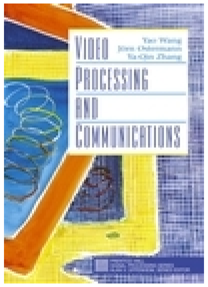
- Recall R-D optimized mode selection in source coding
- When the network is unreliable, more Intra mode is needed
- How often and when/where to insert intra-blocks?
- Extend the R-D framework to consider effect of channel errors:
  - Minimize the decoder distortion, subject to total rate (source rate plus channel redundancy due to FEC/retransmission)
  - Several papers considered this problem: e.g. the ROPE method

Zhang, R., S. L. Regunathan, and K. Rose.Video coding with optimal inter/intramode switching for packet loss resilience. *IEEE Journal on Selected Areas in Communications* (June 2000), 18(6):966–76.




# Layered Coding with Unequal Error Protection

- LC+UEP is until recently the most promising approach for combating channel errors
  - Base layer provides acceptable quality, enhancement layer refines the quality
  - Base layer stream is delivered through a reliable channel (by using ARQ and strong FEC)
  - Good for a network with differentiated service (Do NOT exist today over Internet, may become part of emerging wireless standards)
- Problems:
  - Any error in the base layer causes severe degradation
  - Repetitive ARQ may incur unacceptable delay, strong FEC may be too complex or cause extra delay
  - The enhancement layer is useless by itself



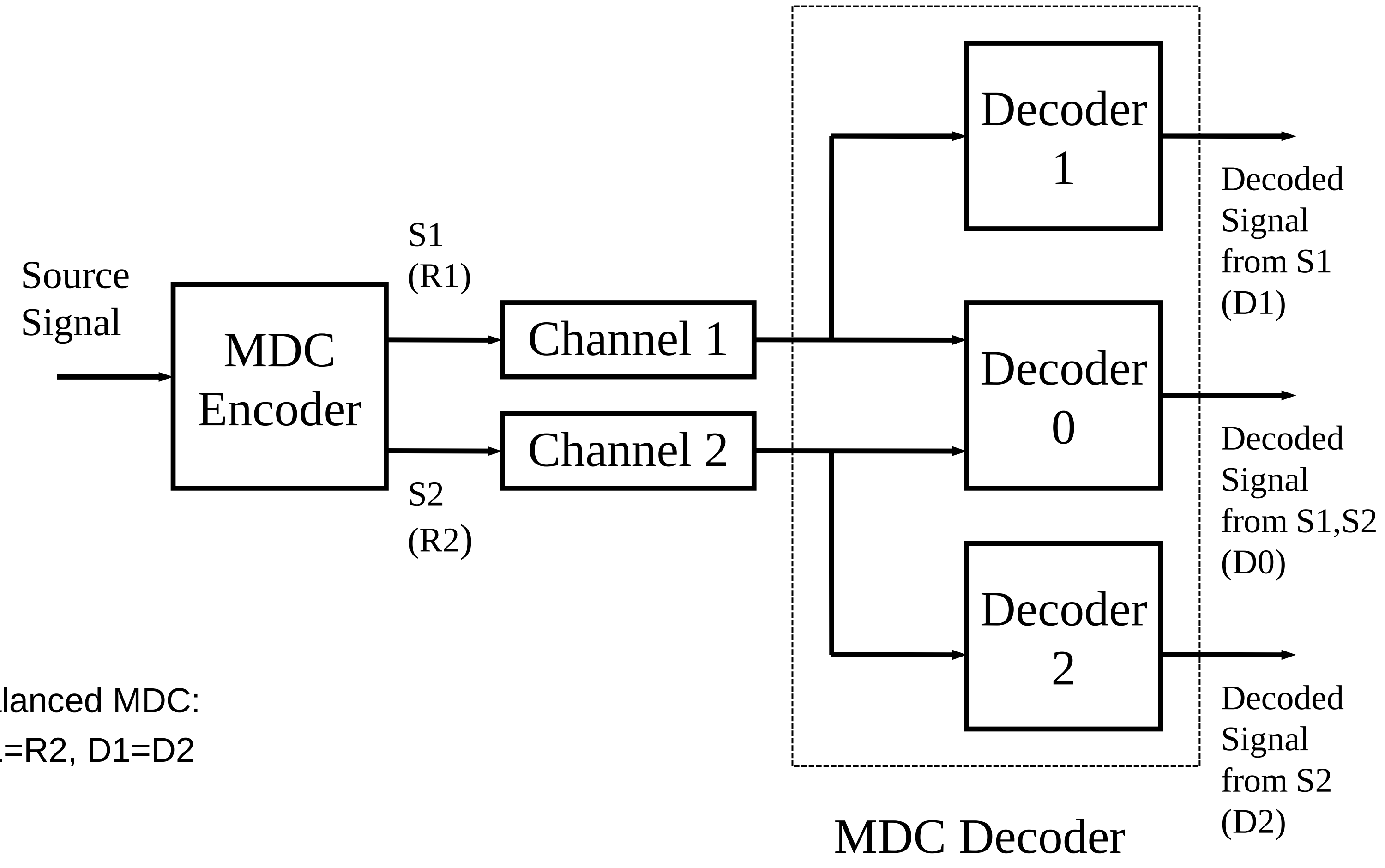

# Multiple Description Coding

- 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 are properly packetized and interleaved

- MDC: Generate multiple correlated descriptions
  - any description provides low but acceptable quality
  - additional descriptions provide incremental improvements
  - No retransmission required -> low delay ☺
  - However: correlation -> reduced coding efficiency ☹

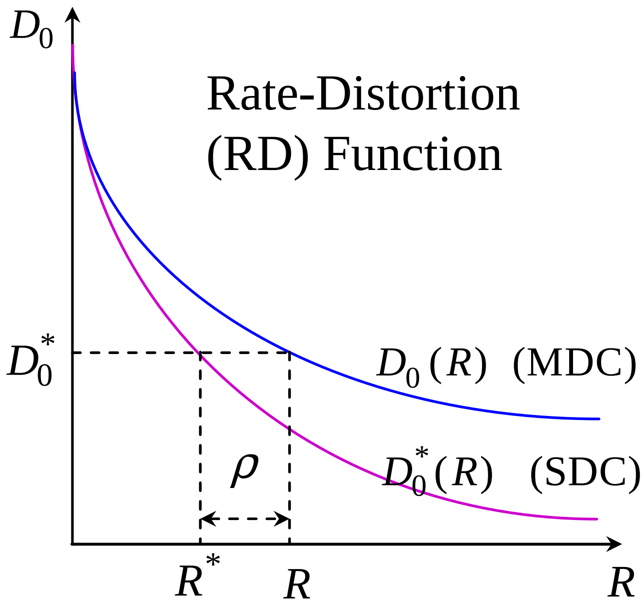
- Design goal:
  - maximize the robustness to channel errors at a permissible level of redundancy

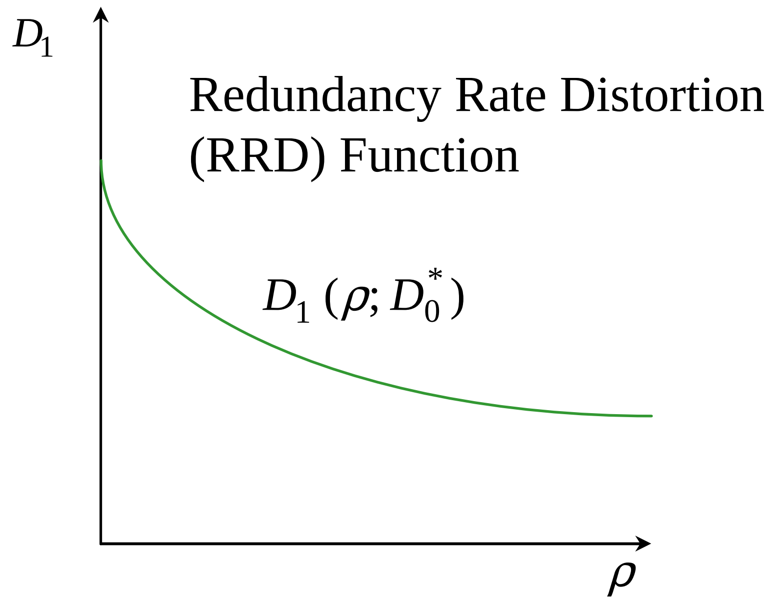
# Generic Two Description Coder



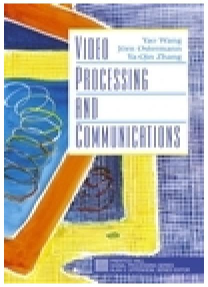

Balanced MDC:
R1=R2, D1=D2

# 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




# Challenge in Designing MD Video Coder

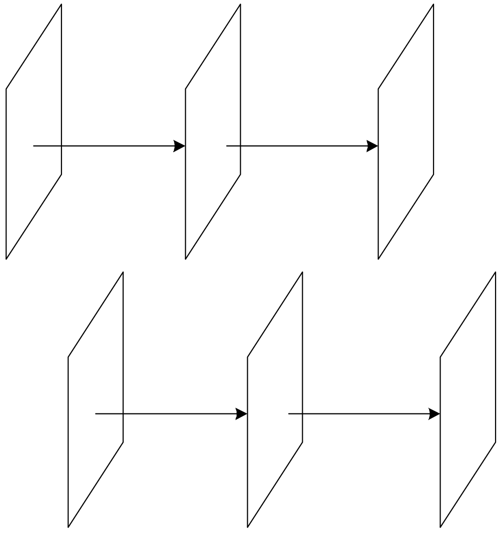
- To achieve high coding efficiency, the encoder should retain the temporal prediction loop
- Should the prediction be based on reconstruction from both descriptions or individual descriptions?
- Prediction based on two-description reconstruction
  - Higher prediction efficiency
  - Mismatch problem at the decoder
- Prediction based on single-description reconstruction
  - Lower prediction efficiency
  - No mismatch problem
- How to provide a proper trade-off between prediction efficiency and mismatch
  - Predict based on two-description reconstruction, but explicitly code the mismatch error

# Video Redundancy Coding in H.263+

- Coding even frames and odd frames as separate threads
  - High redundancy (~30%) due to reduced prediction gain because of longer distance between frames
  - Hard to vary the redundancy based on channel loss characteristics

even frames

odd frames



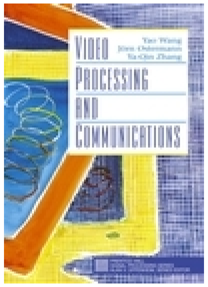
# Multiple Description Motion Compensation (Wang and Lin, 2001)

- A description contains even (or odd) frames only, but each frame is predicted (central predictor) from both even and odd past frames

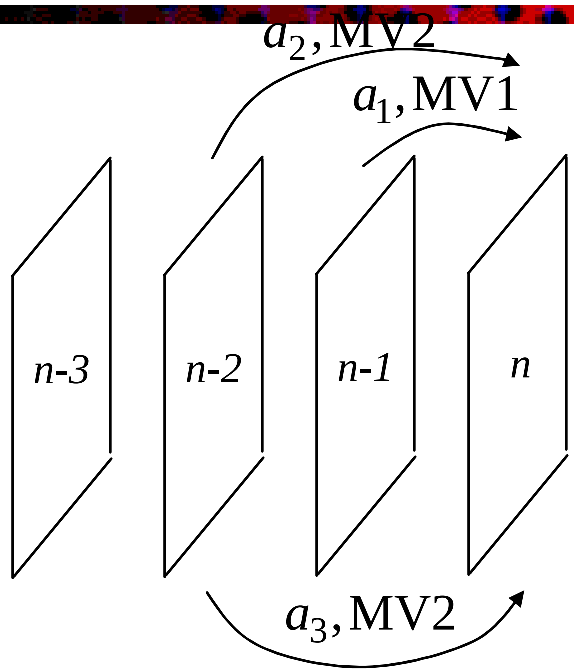
- Code the central prediction error
  - sufficient if both descriptions are received

- 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 predictors and the mismatch error quantizer control the redundancy of the coder, and can be designed based on the channel loss characteristics




# Special Case: Two-Tap Predictor




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

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

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

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

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

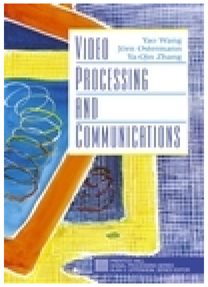
Non-leaky predictor: $a_1 + a_2 = 1, a_3 = 1$

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

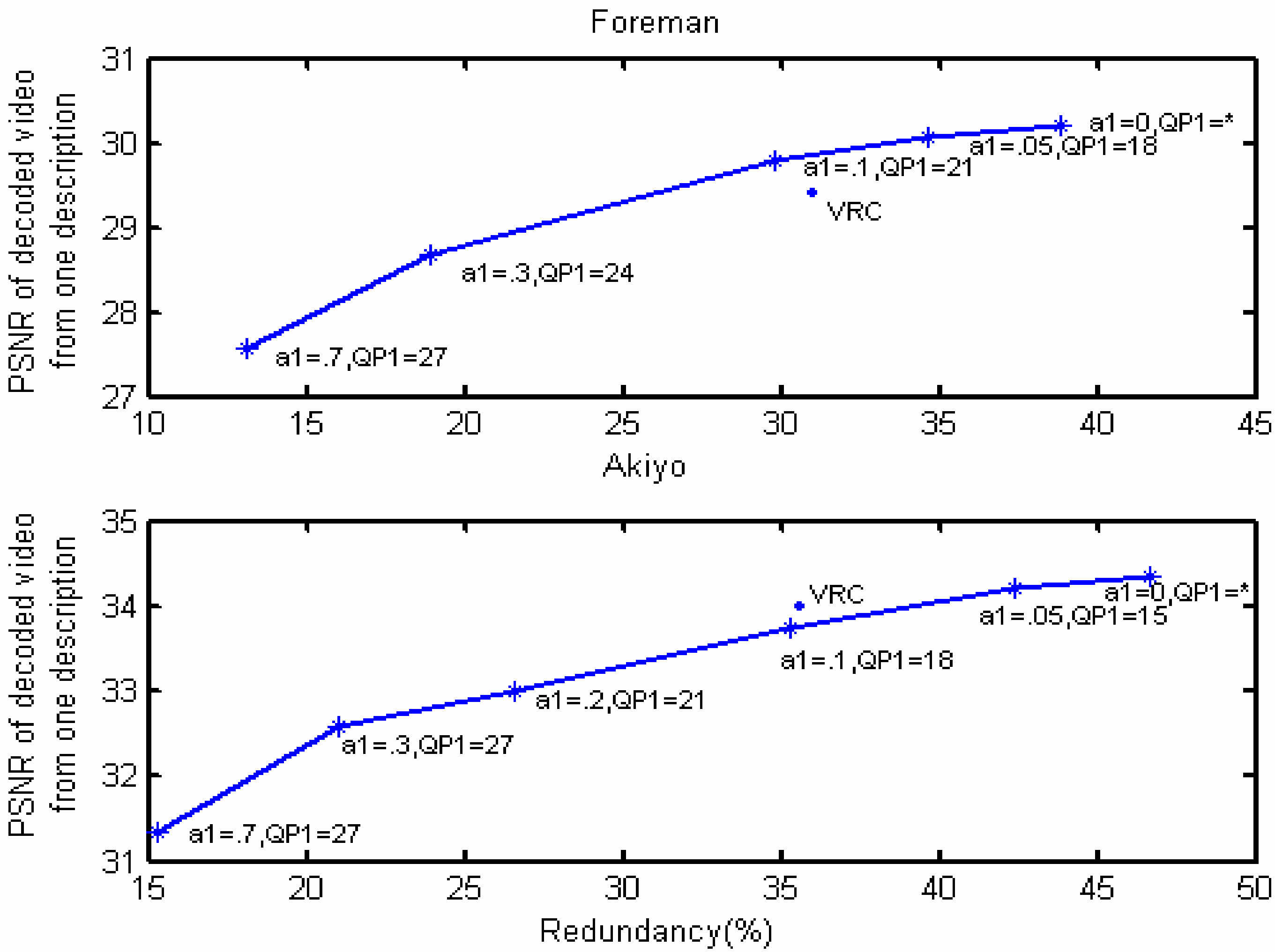
$$\psi_0(n) = \hat{\psi}_0(n) + \tilde{e}_0(n) = \psi(n) + q_0(n)$$

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



# Performance in Packet Lossy Networks

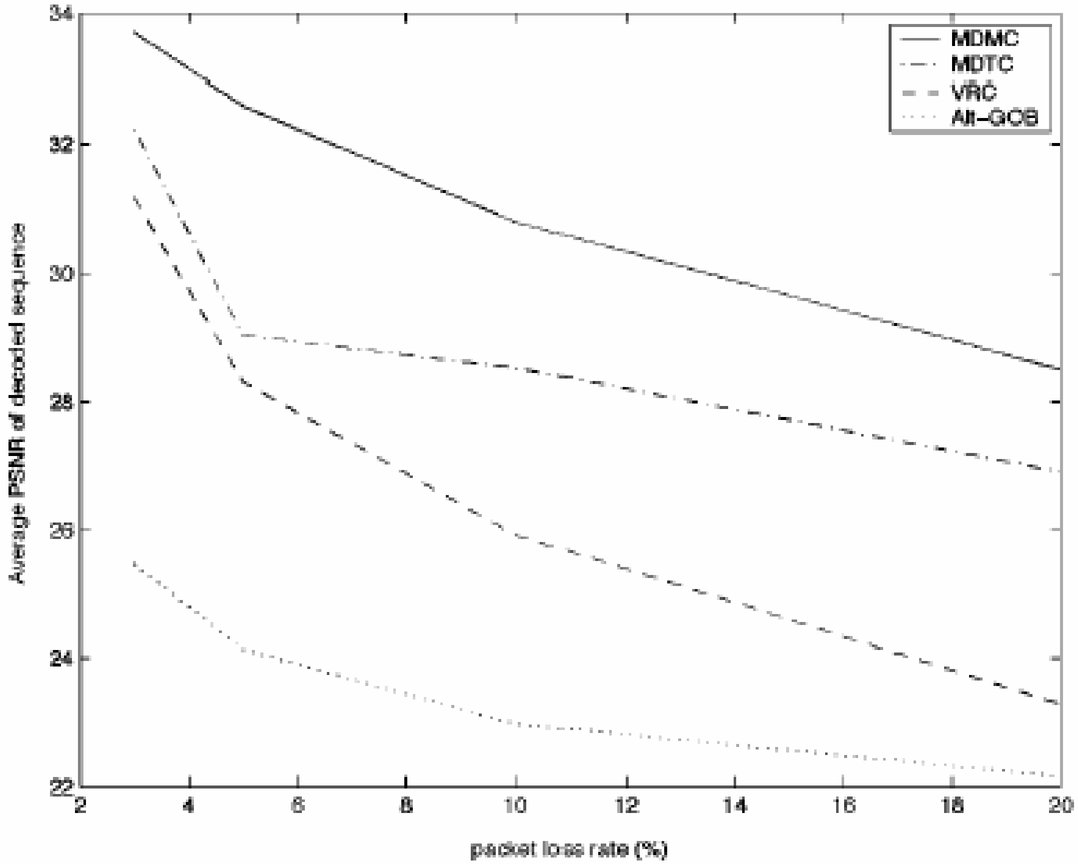
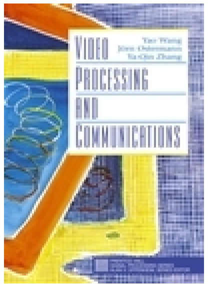



Fig. 13. PSNR of decoded sequences in different packet loss rates. Foreman, 7.5 fps, 144 kbps, two packets per frame.

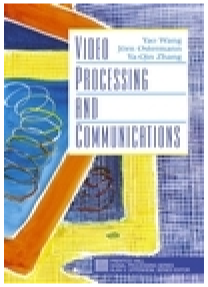
# Sample Reconstructed Frames

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




# Decoder Error Concealment

- With proper error-resilience tools, packet loss typically lead to the loss of an isolated segment of a frame
- The lost region can be “recovered” based on the received regions by spatial/temporal interpolation → Error concealment
- Decoder optimization issue, not part of video coding standard!
- Decoders on the market differ in their error concealment capabilities

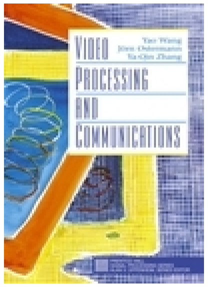



# Error Concealment Techniques

- Basic idea:
  - Recover damaged regions by interpolating from surrounding (in the same frame and in nearby frames) regions
- Motion-compensated temporal interpolation
  - Replace a damaged MB by its corresponding MB in the reference frame
  - If the MV is also lost, has to estimate the MV first, typically by copying the MV of the MB above
  - Simple and quite effective, if the data were appropriately partitioned
- Maximally smooth recovery (Wang/Zhu, 1993)
  - Estimate the missing DCT coefficients in a block so that a combination of spatial and temporal smoothness measure is maximized

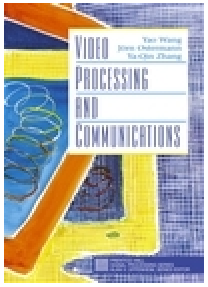
# Sample Error Concealment Results

Without concealment With concealment

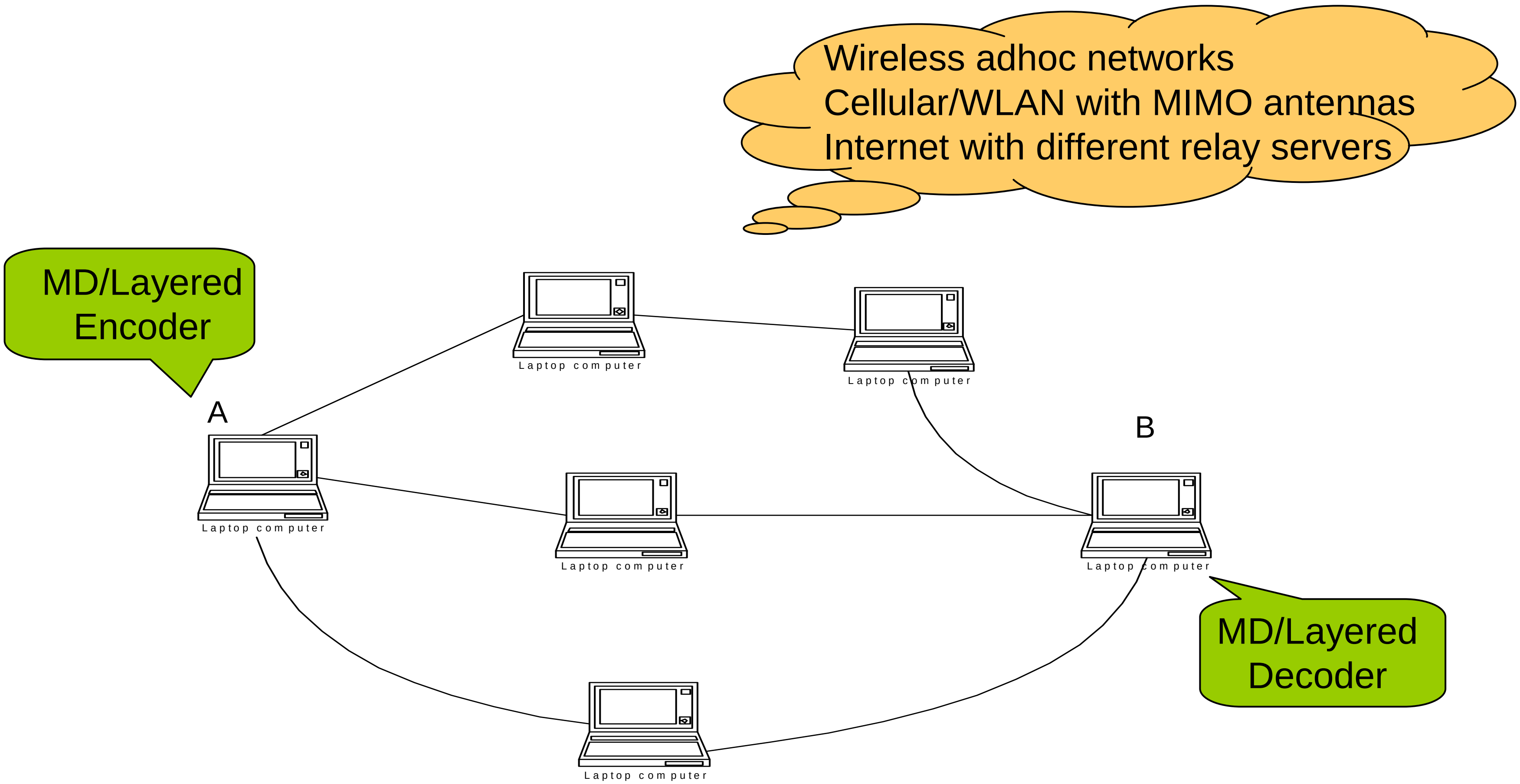


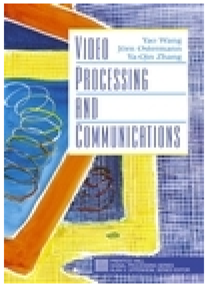

# Encoder-Decoder Interactive Error Control

- Coding parameter adaptation based on channel conditions
  - Change intra-rate based on average loss rates
- Reference picture selection (part of H.263/MPEG-4 standard)
  - Following a damaged frame (feedback info from receiver), use undamaged previous frame as reference frame for temporal prediction
- Error tracking
  - Determine which MBs are affected following a lost MB (feedback info), avoid using those MBs as reference pixels
- Requires a feedback channel, not necessarily involving extra coding delay
- Multiple path transport with multiple stream coding



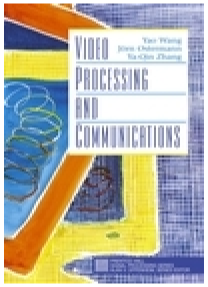
# Video Transport Using Path Diversity






# Why Using Multiple Paths

- Enable sending high bandwidth signal (video) that exceeds the capacity of a single path (e.g. low bandwidth wireless links)
- Can withstand individual path failures in wireless networks or excessive delay due to congestion on a particular path in the Internet
- Enable traffic dispersion and load balancing, which in turn help to reduce congestion and consequently packet losses
- If one path is known to be better than the other, or if it is feasible to set up such a path, use layered coding
- If the paths are symmetric in QoS and no reliable path can be set up, use MDC



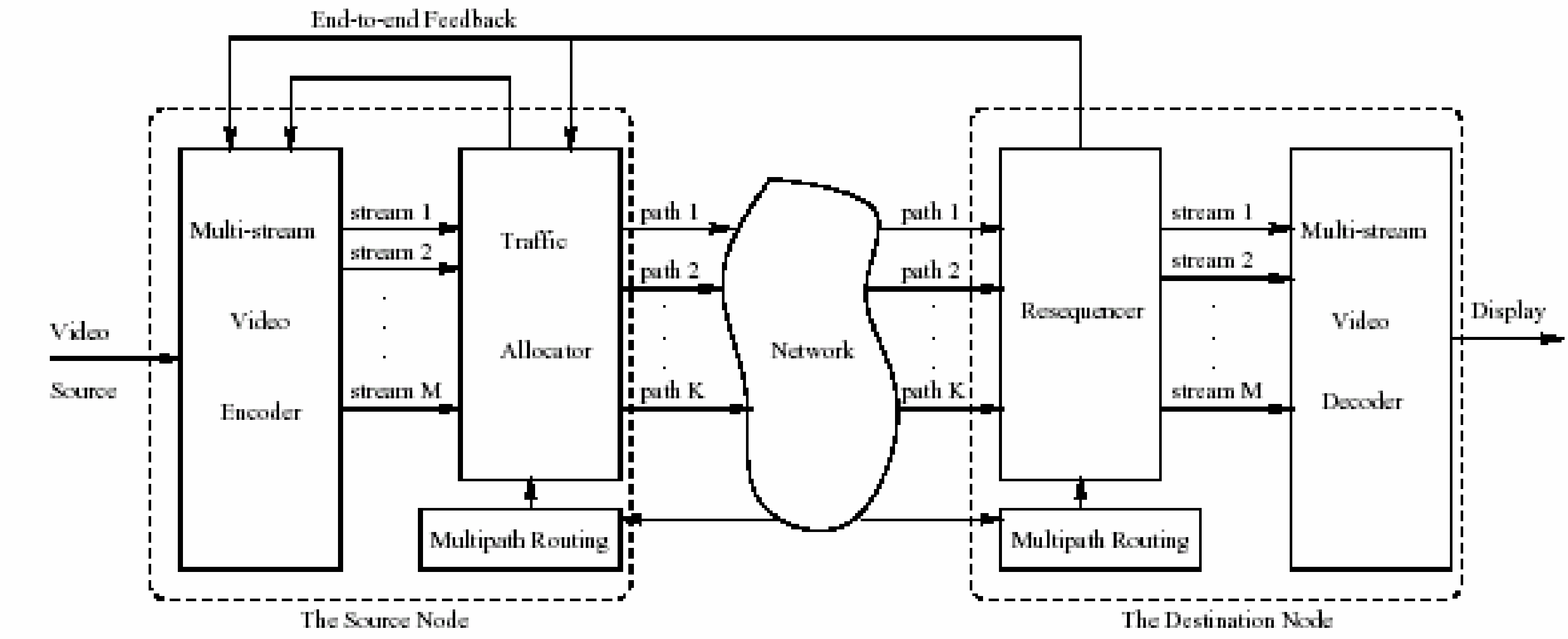

# Video Over Adhoc Networks

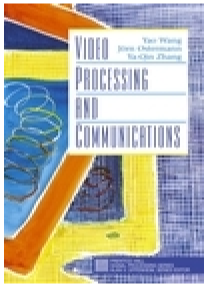
Supported by NSF ITR Program

Joint work with Shivendra Panwar, Shunan Lin, Shiwen Mao

- Wireless adhoc networks:
  - Has no fixed infrastructure, peer to peer
  - Neighboring nodes are continuously changing
  - Multiple paths exit between two end users and many proposed routing protocols return multiple paths to destination
  - A path may become invalid during a connection due to a link down
- Proposed solution: Integration of multistream coding with multipath transport
- Scope of  the Project
  - Set-up/Update an active path set (Haas, Cornell)
  - QoS monitoring of existing paths (Panwar, Poly)
  - MDC/LC based on QoS parameters of active paths (Wang, Poly)
  - Testbed development

# Proposed Solution:
# Multistream Coding + Multipath Transport

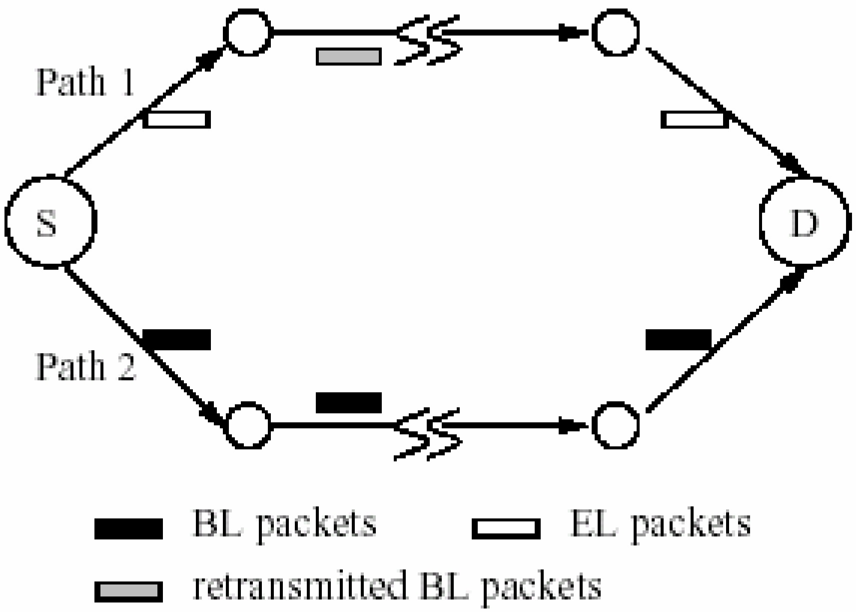





# How to generate multiple source streams?

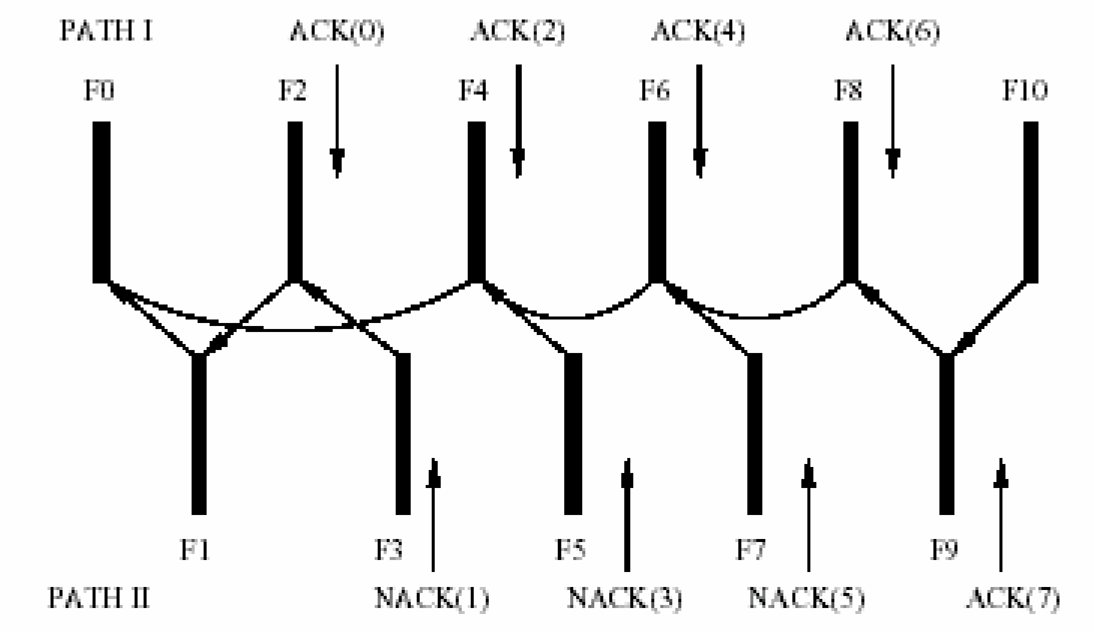
- Trade-off between coding efficiency and error-resilience
- We examine three types of coding and transport schemes
  - Multiple description coding
  - Layered coding with selective ARQ
  - Reference picture selection based on channel feedback
  - The three schemes differ in
    - Requirement for a feedback channel
    - Delay
    - Buffer requirement
    - Video quality under different channel loss patterns

# Layered Coding with Selective ARQ



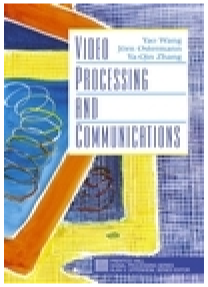

Base layer sent on better path, and lost base layer packets are retransmitted over the enhancement layer path, while dropping the corresponding enhancement layer packets. Redundancy due to prediction using base layer reconstruction only.

# Reference Picture Selection



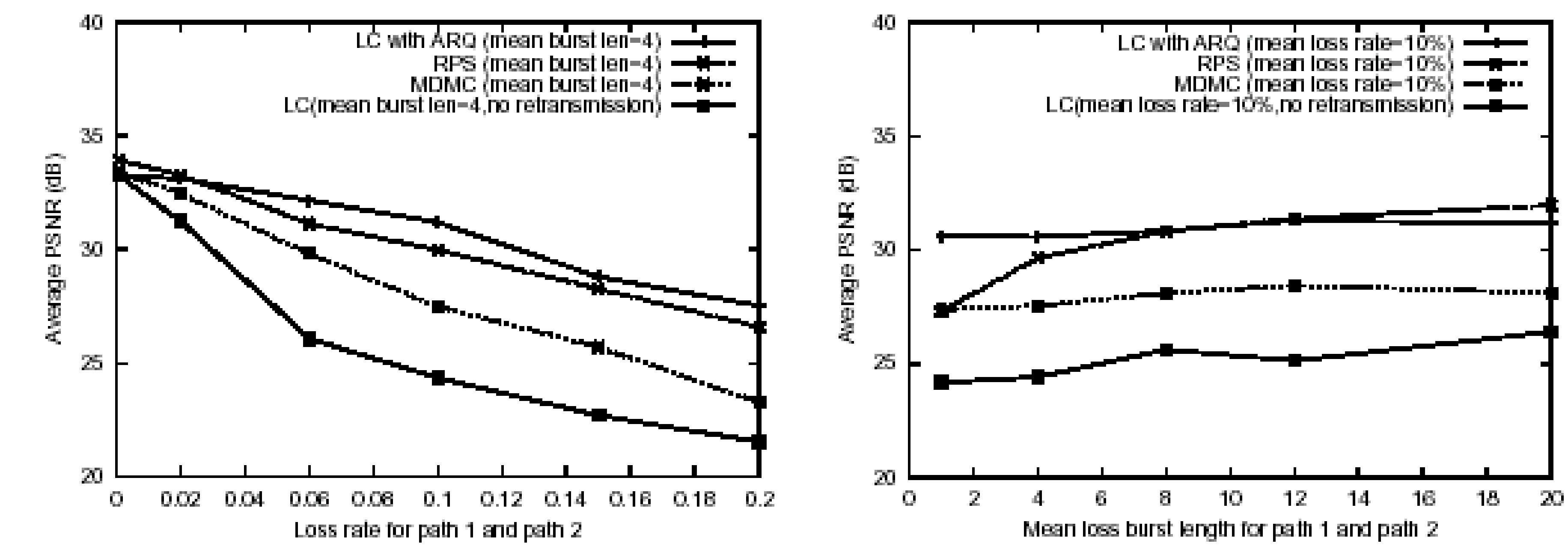

Even/odd frames sent on separate paths. Predict damaged frames based on NACK on each path, and use undamaged frames as reference pictures.

Compatible with the RPS option in H.263+. Standard. ☺

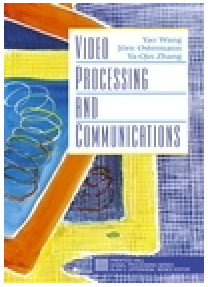



# Simulation Results Based on Markov Link Models

- Path model: Each path consists of 3 links, chosen randomly from an available link pool. Each link is modeled by a 3-state Markov chain (down, poor, good). The path is updated frequently.
- We run a large set of tests to compare the three schemes under different packet loss patterns
- We investigate the influence of the following channel characteristics
  – Mean loss rate
  – Mean burst length
  – Symmetric vs asymmetric paths

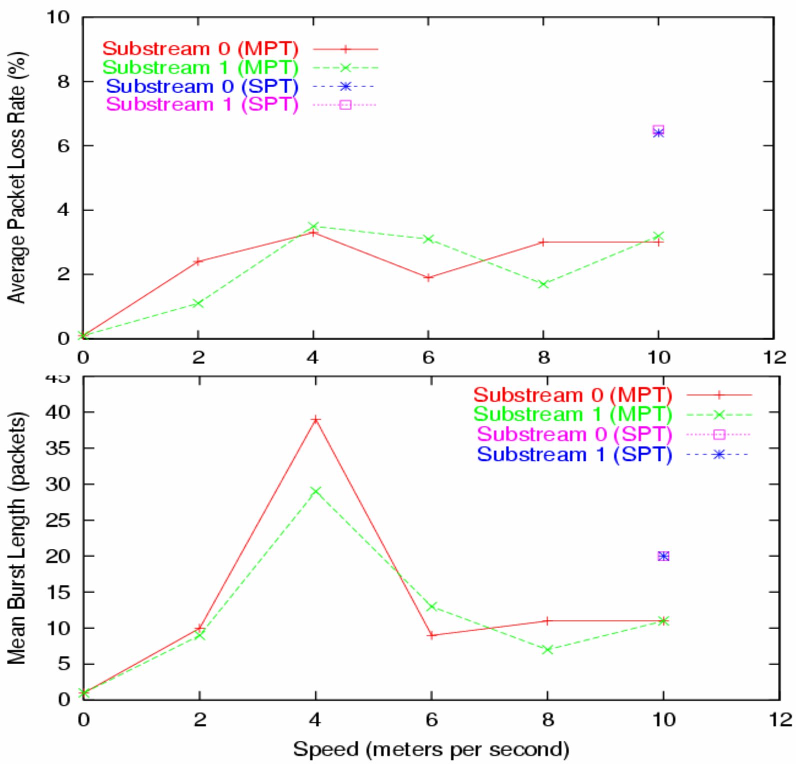


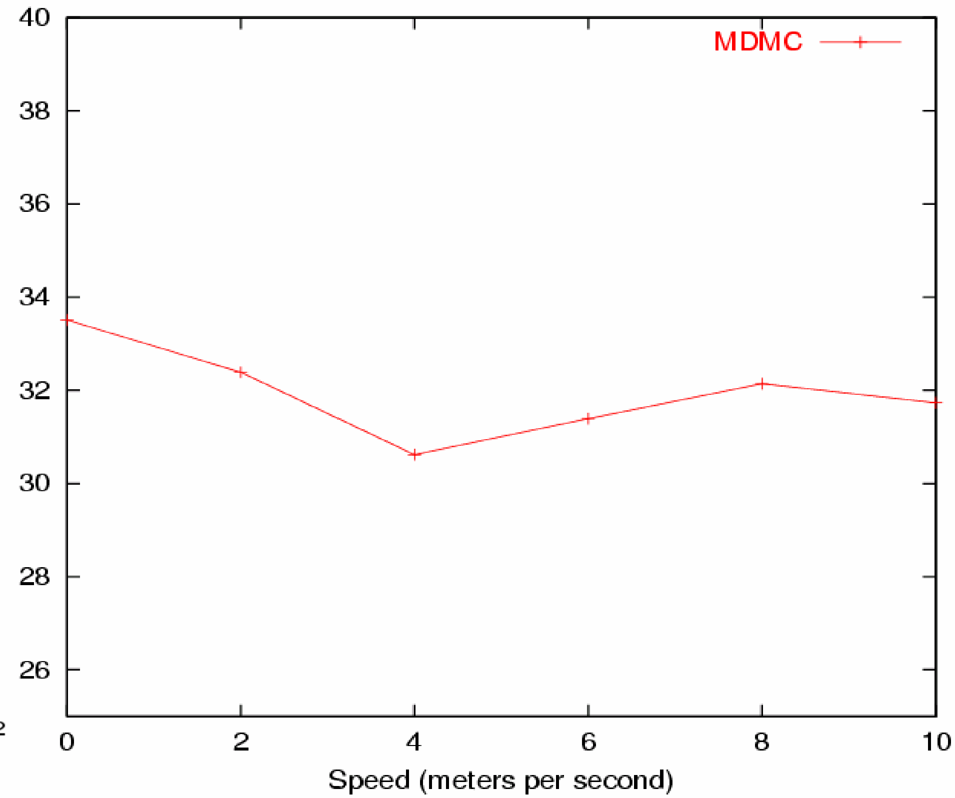

- When no feedback or retransmission is allowed, MDC is significantly better than LC
- Otherwise, LC+ARQ and RPS are better
  - RPS is best at low loss rates
  - LC+ARQ is better at higher loss rates (at the expense of extra delay)
- The performance of all three schemes improves gradually when the burst length increases
  - When the loss rate is the same, longer bursts means fewer frames are effected (bursty error is better than random error for compressed video!)
  - Longer burst lengths increase the diversity gain from using two paths
  - This trend reverses when the burst length exceeds more than 1 frame time.
- When the average loss rate is the same, the three schemes perform similar in the symmetric vs. asymmetric cases



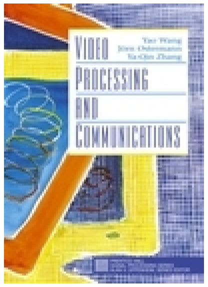

# OPNET Simulation Results

- 16 nodes, 600m by 600m, 10m/s, transmission range=250m, multipath routing DSR (dynamic source routing):

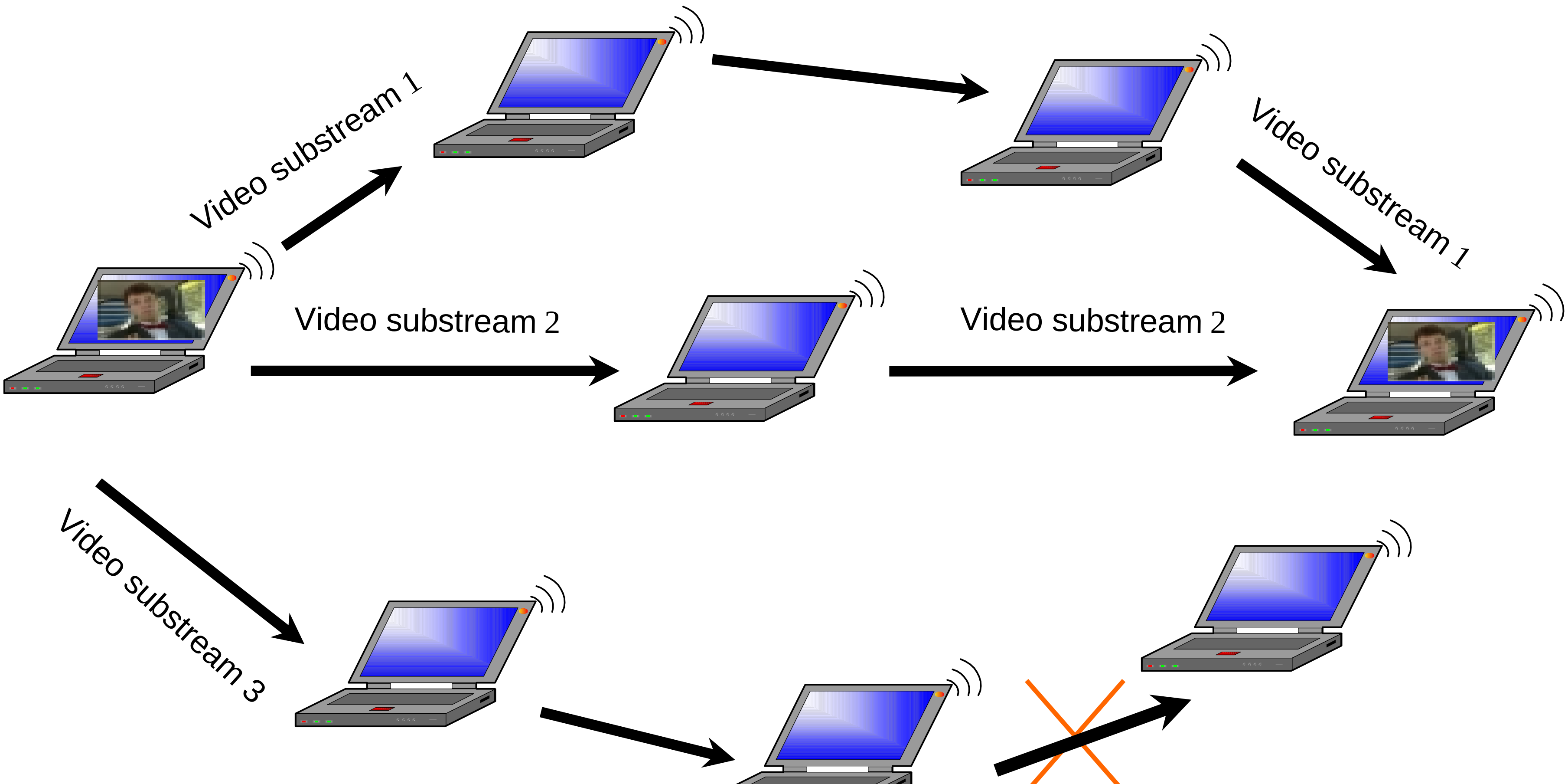


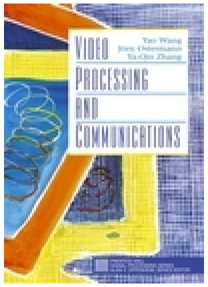

**Mobility can actually help!**




# A Wireless Video Streaming Testbed

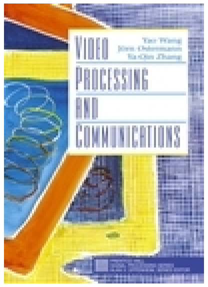



The current testbed consists of 4 mobile nodes: 1 source, 1 destination, 2 relays



# More on the Testbed

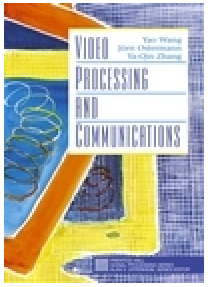
- Testbed Setup
  - Each mobile node is a Thinkpad notebook computer equipped with IEEE 802.11b cards working in the ad hoc mode
  - The source node sends a pre-encoded compressed bit stream (MDMC or LC) to the destination node via two separate relays
  - The destination node assembles the received the packets from two paths, and decodes and displays the video in real time (10 fps)
  - With the LC scheme, the destination node detects and requests retransmission of lost BL packets, the source node resend appropriate BL packets and drop corresponding EL packets
- Results
  - Tested when the nodes are placed in a Polytechnic building
  - With slow moving (walking) of the nodes and limited interference from other traffic, we can get good video quality within both 2s and 300 ms playout delay.
  - Results are fairly consistent with the simulations results using the Markov model, especially with the MDMC scheme




# Summary of the Video over Adhoc Project

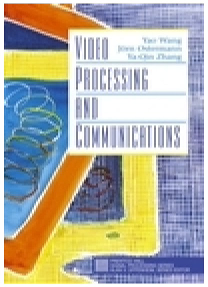
- Multiple Path Transport is feasible and effective for increasing both throughput and reliability in adhoc networks
- The right source coding strategy depends on application requirement and path conditions
  - MDC is effective when
    - one cannot have differential treatment in the network layer
    - delay constraint is very stringent and round trip delay is long
  - Layered coding is effective when the base layer can be transported reliably, through
    - a reliable physical channel, if available
    - retransmission, if delay is acceptable
  - RPS is effective when a feedback channel is available and channel loss rates are not too high
    - But not suitable for video streaming applications where video are pre-encoded

- Reference: S.Mao, S.Lin, Y.Wang, and S.Panwar, ``Video Transport over Ad Hoc Networks: Multistream Coding with Multipath Transport," IEEE J. Select. Areas Commun., Dec. 2003.




# Summary

- Transport level error control
  - Guarantee a basic level of quality
  - Retransmission is effective within the delay constraint
  - Unequal error protection: have practical difficulty, against separation of source coding and transport layer)
- Error resilient encoding
  - Trade off coding efficiency for error resilience
  - Some techniques are only useful for bit-error dominated channels
- Error concealment
  - Does not involve extra redundancy, motion-compensated temporal concealment is simple and yet offer visible improvements
- Encoder-decoder-network interactive error control
  - Require feedback info, may not be available
- Choice of techniques depends on underlying application and network




# 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.
- Y. Wang, J. Ostermann, Y.-Q. Zhang, Video processing and communications, Prentice Hall, 2002. Chap. 14.

# Homework

- Reading assignment
  - Y. Wang, J. Ostermann, Y.-Q. Zhang, Video processing and communications, Prentice Hall, 2002. Chap. 14, 15.