Video Transmission

Aggelos K. Katsaggelos

Professor
Director Motorola Center for Seamless Communications
Northwestern University
Department of EECS
Evanston, IL 60208
www.ece.northwestern.edu/~aggk

Cali, October 28, 2008

Plethora of exciting applications!

- Nomadic people who like to be seamlessly connected
  - Interact and communicate with other people (UMA/UME)
  - Watch the news "on the run"
- Sales of phone digital cameras surpassed sales of digital cameras in 2001
- We are not only users but also creators of content
- 70% of the internet traffic is video traffic (ie UTube)

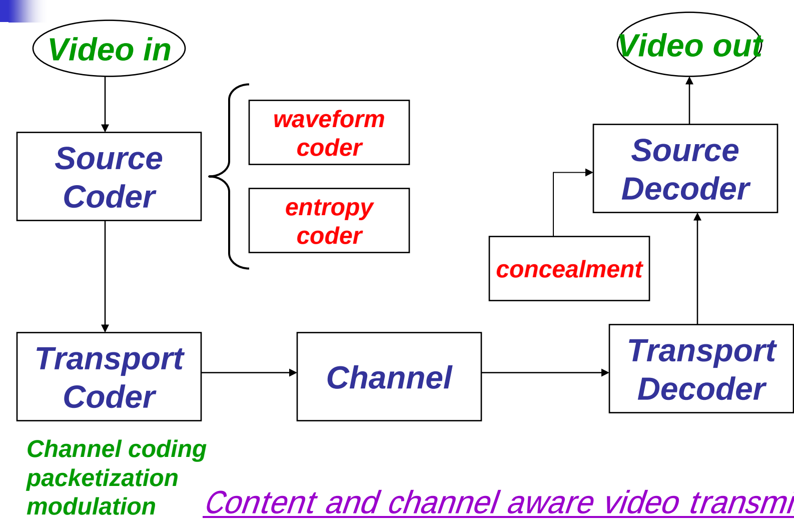
Video is the protagonist!

The Five+ Sense Communication Interface	Contribution to Human Perception
Visual	60%
Auditory	20%
Tactile	15%
Taste	3%
Smell	2%
Existence/Emotion	?%

Talk Objectives

- Describe the building blocks of a video transmission system
- Offer a framework and specific examples
- Discuss open problems, challenges, and opportunities

Video Communication System



Content and channel aware video transmission

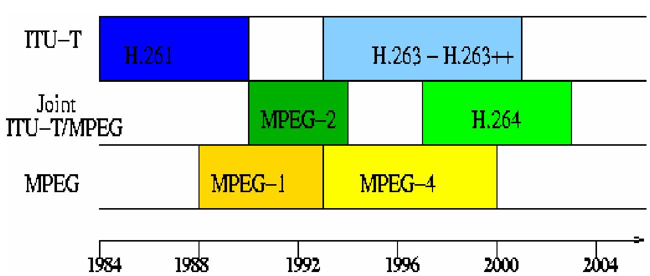
The 3 components

- Source Compression
- Channel Models
- Post-processing (concealment)

The 3 components

- Source Compression
- The Channel
- Post-processing (concealment)

Video Compression Standards



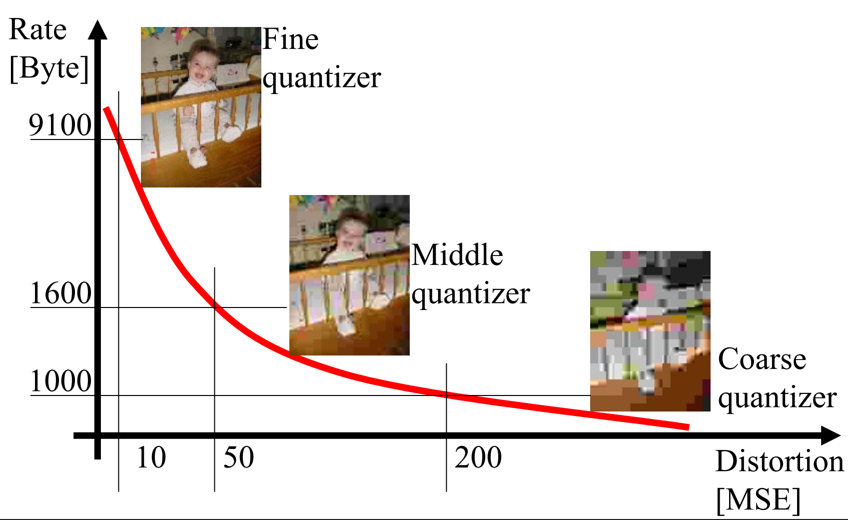
Compression Standards

Entertainment ⟶ ⟵ Communication

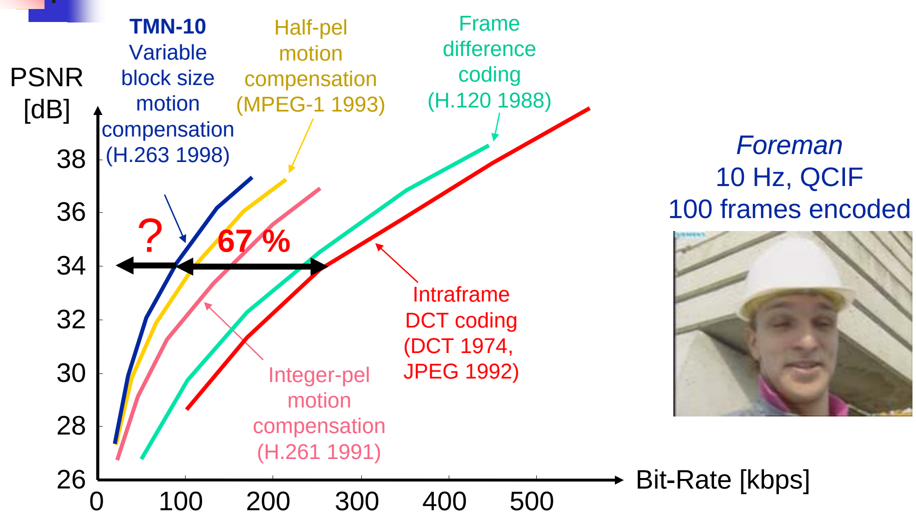
	MPEG-1	MPEG-2	MPEG-4	H.263	H.261
Dates of Standardization	11/92	11/94	1/99 Version 1 1/00 Version 2	5/96 Version 1 1/98 Version 2	12/90 Version 1 5/94 Revised
Primary Applications	Digital Storage Media	Broadcast/ DVD/ HDTV	Web Authoring, Multimedia Compression, Wireless Videophone	Desktop/ Wireless Video-conferencing	Wireline Video-conferencing
Typical Video Bitrates	1.5 Mbps	4-6 Mbps	20 Kbps - 6 Mbps	20-384 Kbps	128-384 Kbps
Typical Video Frame Size	352x240 (SIF)	720x480 (Rec. 601)	176x144 (QCIF) 352x288 (CIF) 720x480 (601)	176x144 (QCIF) 352x288 (CIF)	176x144 (QCIF) 352x288 (CIF)
Typical Associated Audio Quality	Stereo CD Quality	Surround Sound	Speech/Music/ Stereo CD/Surround Sound	Speech	Speech

Operational Rate-Distortion Function

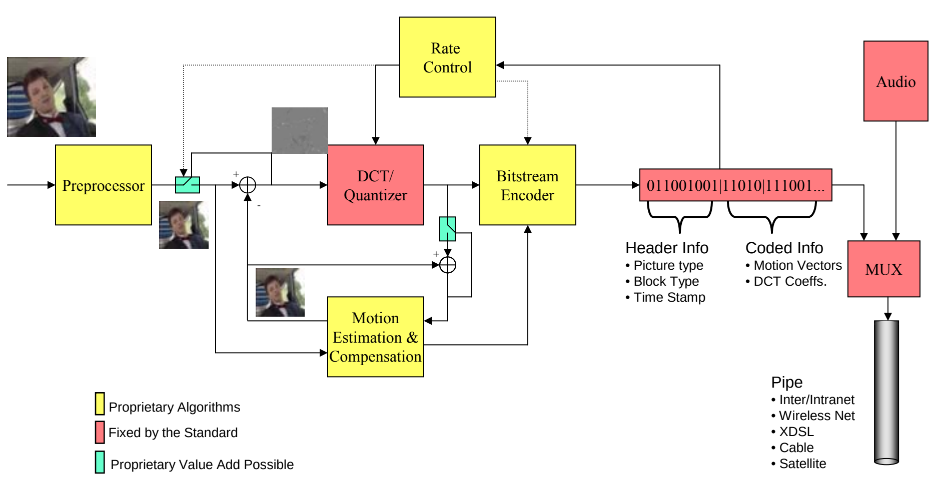
Find optimal quantizers with a given bit budget



Video Coding Efficiency



Hybrid Motion Compensated Video Encoding



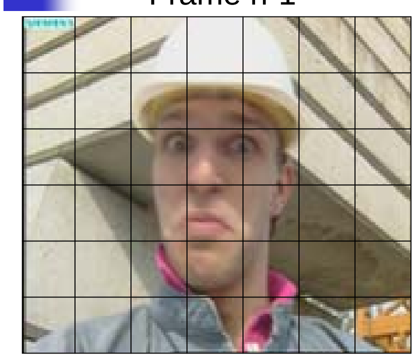
Video Coding

Frame n-1

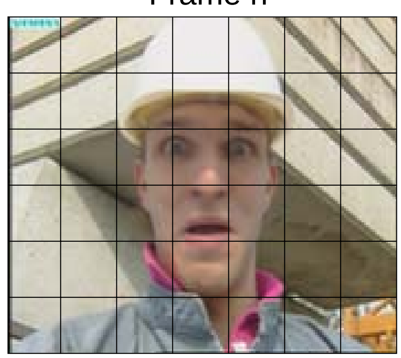
Frame n

Video Coding

Frame n-1

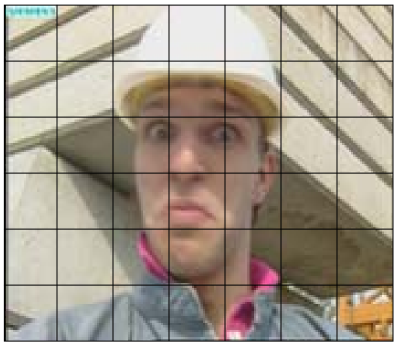


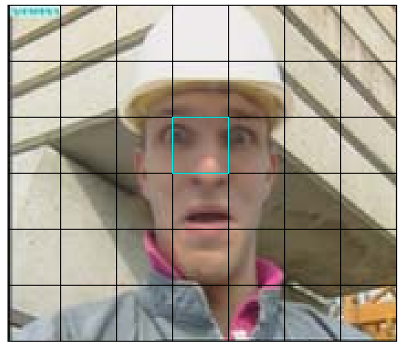
Frame n



Video Coding

Frame n-1 | Frame n



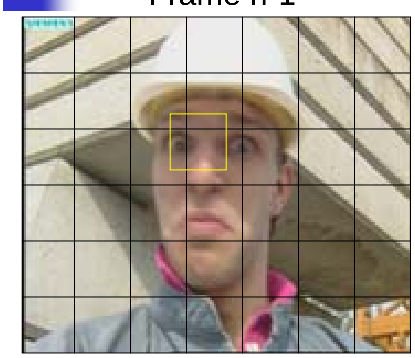


Intra Mode:
Independently encode the block in frame n

Inter Mode:
Use frame n-1 to predict the block in frame n

Video Coding

Frame n-1 | Frame n

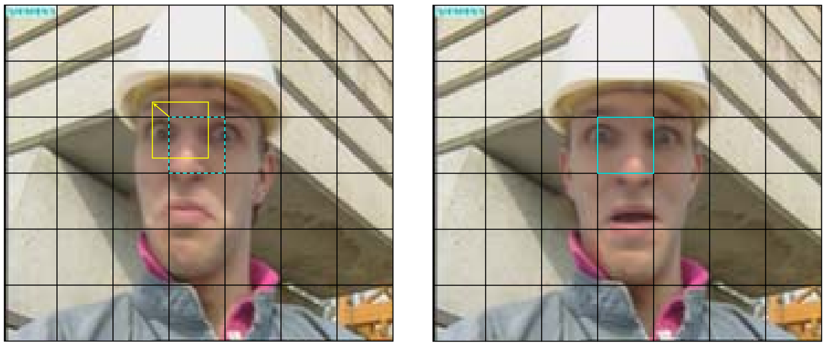


Intra Mode:
Independently encode the block in frame n

Inter Mode:
Use frame n-1 to predict the block in frame n

Video Coding

Frame n-1 Frame n



Intra Mode:
Independently encode the block in frame n

Inter Mode:
Use frame n-1 to predict the block in frame n

DFD Comparison

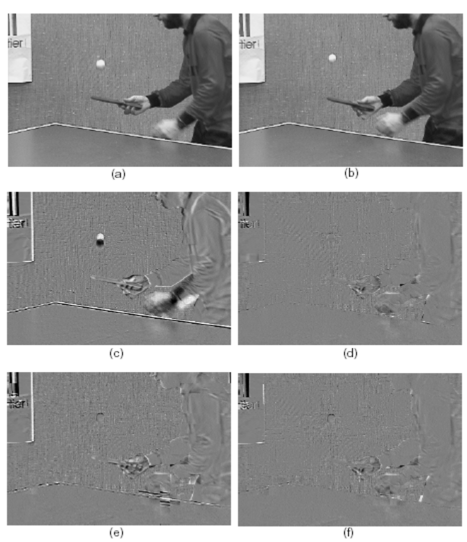
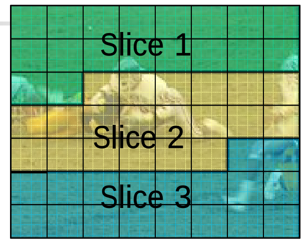


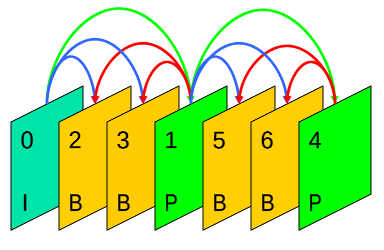
Figure 1: (a) reference picture; (b) current picture; (c) frame difference; (d) full-search; (e) logarithmic search; (f) three level hierarchical search

Video Coding Basics

- Slices
  - Contain groups of macroblocks
  - Macroblocks within a slice depend on each other (e.g., predictive coding of motion vectors and DC coefs)
    - Improves coding efficiency
  - All predictors are reset at slice boundary
    - Improves robustness to error
- Frames
  - I, P, B

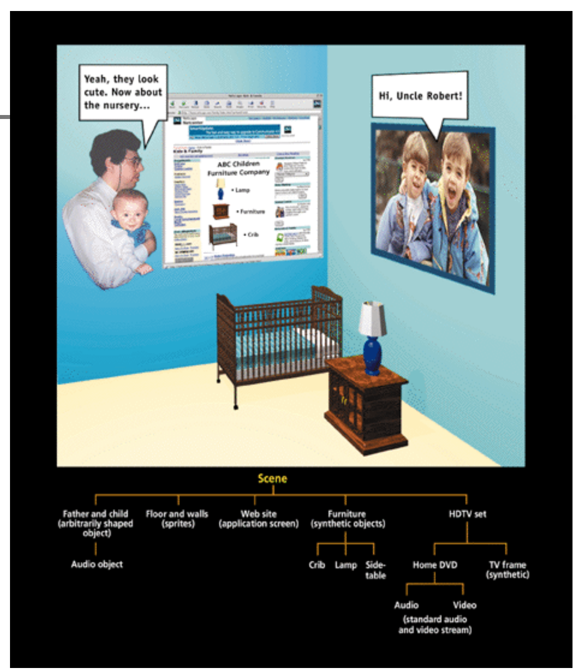


Video frame



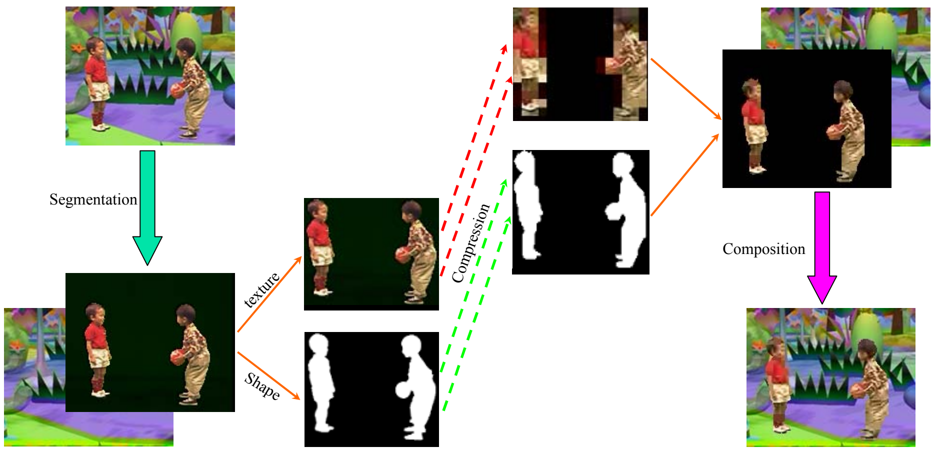
Video sequence

MPEG-4 Scene: Object-oriented composition

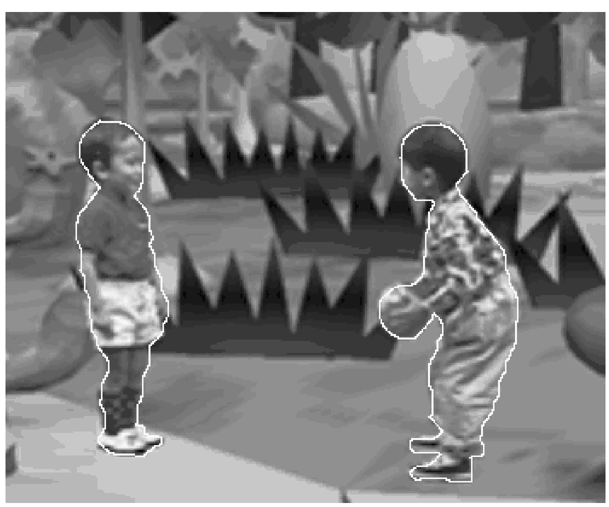


Rob Koenen, KPN Research

Object-based Video Encoding



Experimental Results



Source Compression

- The enabling technology
- Successful Standards
  - MPEG-4, H.264/AVC, SVC, H.265
- "Rule of Thumb" at a price
- Error Resiliency Challenges
  - Predictive coding
  - VLCs

Error Resiliency Challenges

- Non-Robust Nature of (VLBR) Video Coding
  - Highly Predictive
  - Variable Length Codes (requires resynchronization)
- A lot of care has been taken in source encoding to increase error resiliency

The 3 components

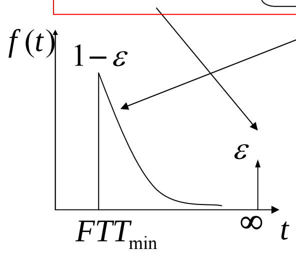
- Source Compression
- Channel Models
- Post-processing (concealment)

Channel Models

- Wireless vs Wired vs Hybrid
- Network model :

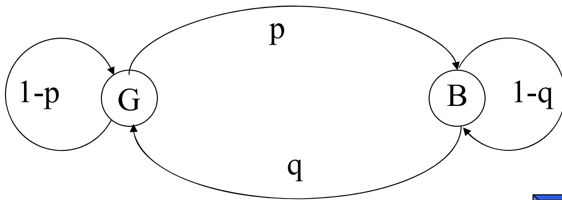
  an independent time-invariant packet erasure channel + random delays

$$\rho^k = \underbrace{\varepsilon} + (1-\varepsilon)\underbrace{P\{\Delta T_n(k) > \tau\}}$$



- Packet delay
  - Exponential distribution: fast decaying tail
  - Gamma distribution
  - Pareto distribution: slowly decaying (heavy) tail
- Packet Loss
  - Bernoulli
  - 2-state Markov (Gilbert)
  - High-order Markov
  - Nakagami-m fading model

Gilbert Model



G: Good State, B: Bad State
While in B, error prob=1-h
burst length=1/q*1000/bitrate msec
Overall BER=(1-h)*p/(p+q)

Gilbert Model
BER = $5x10^{-3}$
Burst Length = 2.5ms
Bad BER h = 0.5

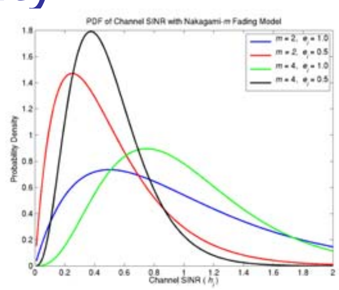
Outage Probability

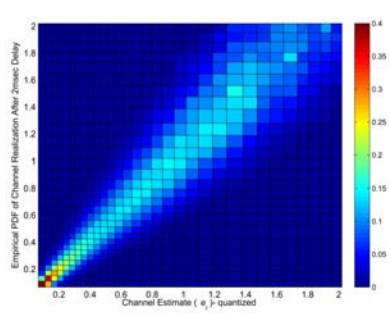
- Nakagami-*m* Fading Model
  - Mean at channel estimate (e_i)
  - Cumulative cdf at h_i

$$\phi_{e_i}(h_i) = \frac{\gamma\left(m, \frac{mh_i}{e_i}\right)}{\Gamma(m)}$$

- Use Empirical Distribution
- Packet Loss Rate

$$\varepsilon_i = \phi_{e_i}\left(\frac{n_i}{p_i} 2^{\frac{r_i}{n_i B}-1}\right)$$





Error Resiliency Challenges

- Broad Error Conditions
  - Random Bit Errors
  - Burst Errors
  - Packet Loss Errors
- Low Delay (and additional constraints on resources)
  - Interleaving can be a problem

The 3 components

- Source Compression
- Channel Models
- Post-processing (concealment)

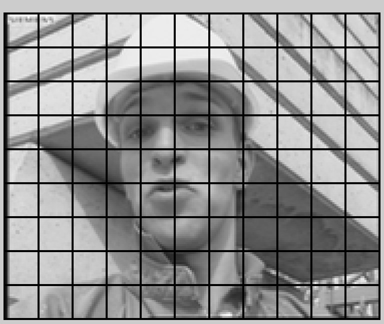
Error Concealment Approaches

- Simple
  - Packets from the same group are not used for concealment
- Complex
  - Incremental gain in quality due to the addition of a packet no longer additive

Error Concealment Scheme

Case 1: Packet k-1 RECEIVED

Encoder Reconstruction



Frame 2

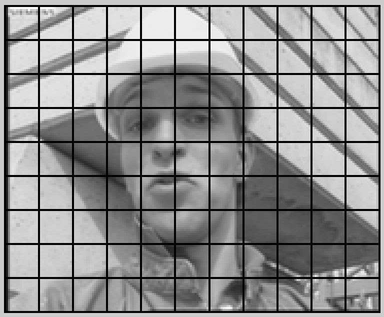
Decoder Reconstruction

Frame 2

Error Concealment Scheme

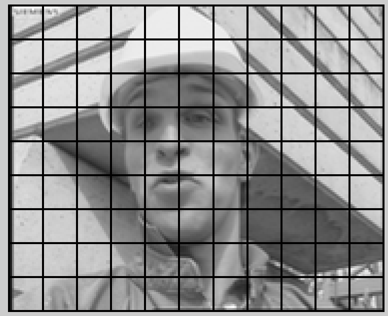
Case 1: Packet k-1 RECEIVED

Encoder Reconstruction



Frame 2

Decoder Reconstruction

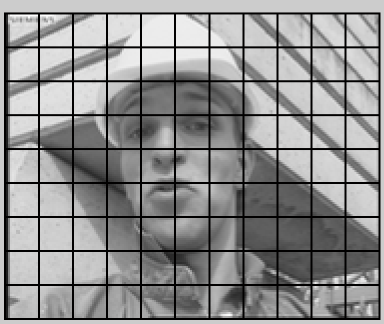


Frame 2

Error Concealment Scheme

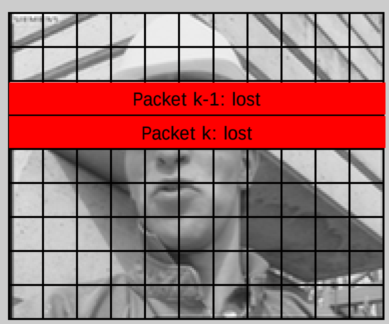
Case 2: Packet k-1 LOST

Encoder Reconstruction



Frame 2

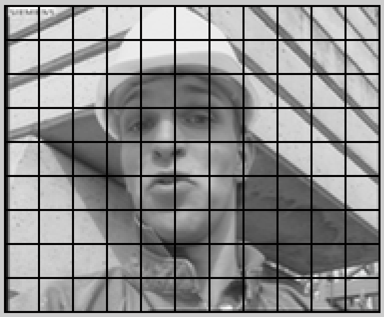
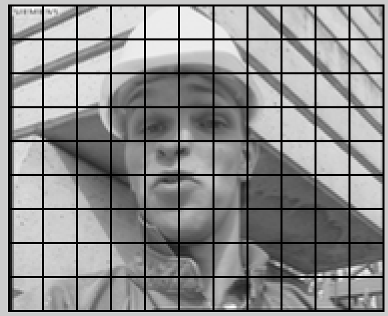
Decoder Reconstruction



Frame 2

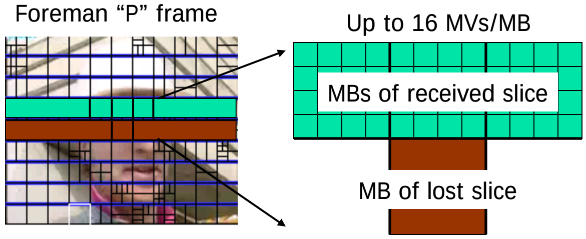
Error Concealment Scheme

Case 2: Packet k-1 LOST

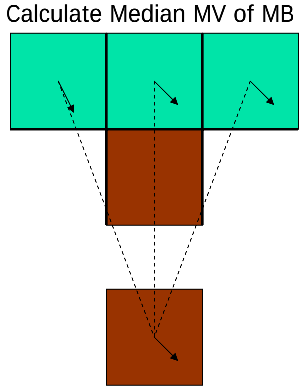
Encoder Reconstruction	Decoder Reconstruction
	
Frame 2	Frame 2

Error Concealment Scheme

Foreman "P" frame



Calculate Median MV of MB

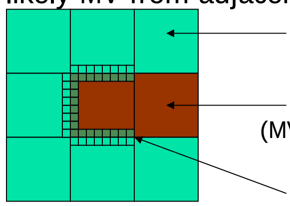


- Use most likely MV from top MBs
  - If top MBs are available, use median MV of top MBs, else use zero MV

Copy median of top MVs

Complex Concealment

- Use most likely MV from adjacent MBs

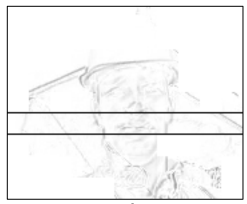


- Slice Distortions Not Additive

Pkt #4 Received

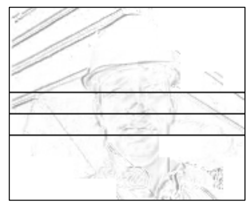
MSE Gain = 262

Pkt #5 Received



MSE Gain = 137

Pkts 4 and 5 Received

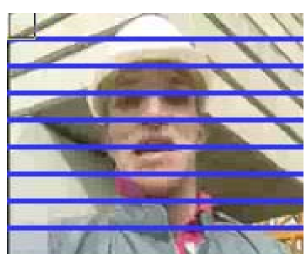


MSE Gain = 272

(Darker pixels indicate greater gain in quality compared to not receiving any packets at all)

Packet Importance

Frame 157



Frame 159

- Depends on
  - Concealability
  - Reliability of the Neighbors
  - Reliability of the Reference

The Connector: Expected Distortion

Depends on coding parameters for the current packet

Depends on concealment scheme (μ and ρ for other packets)

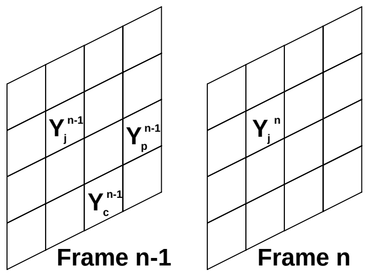
$$E\left[D^k\right] = \left(1-\rho^k\right) \underbrace{E\left[D_R^k(\mu^k)\right]} + \left(\rho^k\right) \underbrace{E\left[D_L^k\right]}$$

- What Affects the Expected Distortion?
  1. Source coding
  2. Probability of loss in the channel
  3. Error concealment

Computing the Expected Distortion

- Simulations
- Per-Pixel Accurate Techniques
  - Recursive calculations
  - ROPE and extensions (Zhang et al 2000)
- Model-Based Techniques
  - Reduce computational complexity
  - Reduce memory requirements
- End-to-end Frame Distortion (ave, max)
- Difference between expected and actual

Recursive Distortion Estimate



$$E[D(k)] = (1\text{-}p(k))E[D_r(k)] + p(k)E[D_l(k)]$$

where

$$E[D_r(k)] = x^2 - 2xE[Y^n_{r,j}] + E[(Y^n_{r,j})^2]$$

$$E[D_l(k)] = x^2 - 2xE[Y^n_{l,j}] + E[(Y^n_{l,j})^2]$$

x= original pixel value

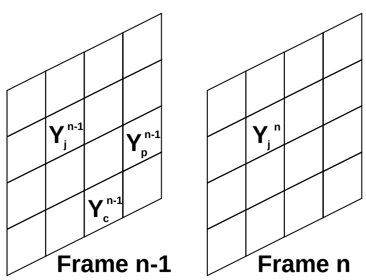
Y_j^n = jth reconstructed pixel in the nth frame

Y_p^{n-1} = pixel in previous frame used as the prediction for current pixel (Inter)

Y_c^{n-1} = pixel in previous frame used to conceal the current pixel if the previous packet is received

- Distortion metric = Mean Squared Error
- Concealment strategy presented earlier
- Can recursively calculate the distortion for the pixels in the current frame based on statistics of the previous frame

Recursive Distortion Estimate



$$E[Y^n_{r,j}] = \begin{cases} r_j + E[Y^{n-1}_p] & \text{Inter} \\ r_j & \text{Intra} \end{cases}$$

$$E[(Y^n_{r,j})^2] = \begin{cases} (r_j)^2 + 2(r_j)E[Y^{n-1}_p] + E[(Y^{n-1}_p)^2] & \text{Inter} \\ (r_j)^2 & \text{Intra} \end{cases}$$

$$E[Y^n_{l,j}] = (1\text{-}p(k\text{-}1))E[Y^{n-1}_c] + p(k\text{-}1)E[Y^{n-1}_j]$$

$$E[(Y^n_{l,j})^2] = (1\text{-}p(k\text{-}1))E[(Y^{n-1}_c)^2] + p(k\text{-}1)E[(Y^{n-1}_j)^2]$$

x= original pixel value

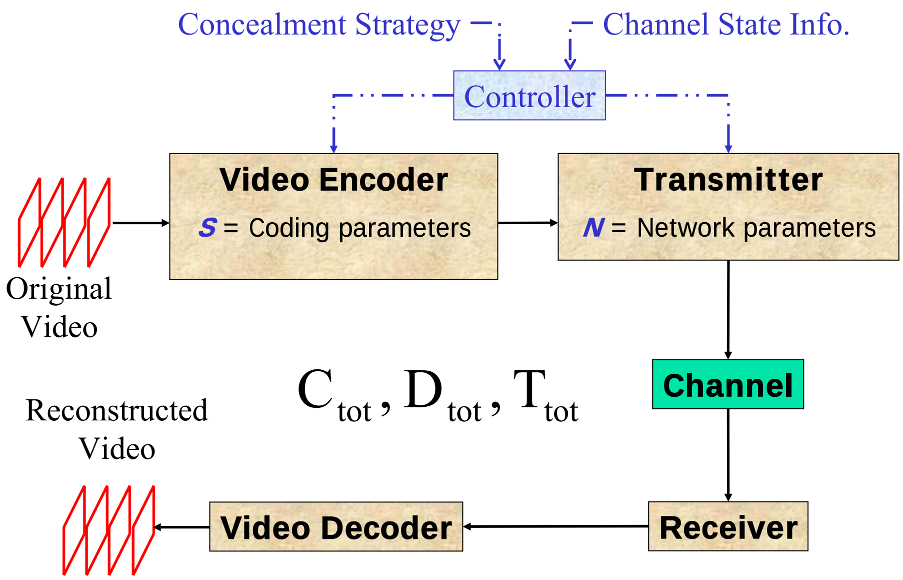
Y_j^n = jth reconstructed pixel in the nth frame

Y_p^{n-1} = pixel in previous frame used as the prediction for current pixel (Inter)

Y_c^{n-1} = pixel in previous frame used to conceal the current pixel if the previous packet is received

- 1st and 2nd moment of each pixel are required to accurately calculate the distortion per pixel
- Can recursively calculate these values from frame to frame.
- ROPE method: Zhang, Regunathan, Rose, “Video coding with optimal inter/intra...,” *IEEE JSAC*, June ‘02

System Model



$C_{tot}, D_{tot}, T_{tot}$

Examples of cost/constraints

- Cost
  - Transmission (computation) power
  - $$ for DiffServ
- Network Parameters
  - Scheduling
  - Transmission rate
    - Adaptive modulation
    - Variable rate spreading
  - Probability of packet loss
  - Adaptive FEC
  - ARQ (fast, hybrid)

F. Zhai and A. K. Katsaggelos, *Joint Source-Channel Video Transmission*, Morgan and Claypool, 2007.
A. K. Katsaggelos et al, "Advances in Efficient Resource Allocation for Packet-Based Real-Time Video Transmission", *IEEE Proc.,* Jan. 2005.

Transmission Cost

- Wireless = Transmission energy

$$E_{tot} = \sum_{k=1}^{K} E^k = \sum_{k=1}^{K} B^k(\mu^k)\frac{P^k}{R^k}$$

Total Energy
Packet index
Number of Bits
Source Coding Parameters
Transmission Power
Transmission Rate
Cost per bit

- DiffServ = Pricing

$$C_{tot} = \sum_{k=1}^{K} C^k = \sum_{k=1}^{K} B^k(\mu^k)c^k$$

Total Cost

Considerations

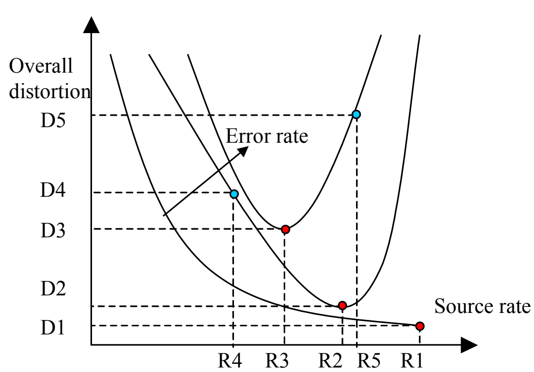
- Channel-Aware Resource Allocation
  - Joint-Source Channel Coding
  - Network Adaptive Video Transmission
  - Dynamic Resource Allocation
- Content-Adaptive Resource Allocation

Cross Layer Optimization

Performance Metrics

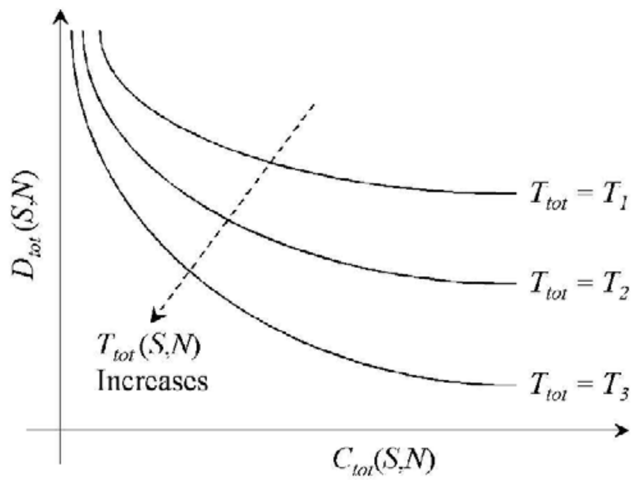
2.4	Max throughput	Throughput	Bandwidth	delay	jitter	Data Overhead	Loss-rate
app	**Max-goodput** can be measured by sending as large files as possible.	**goodput**: the number of useful bits per unit of time forwarded by the network from a certain source address to a certain destination, excluding protocol overhead, and retransmitted data packets. Such as (file size)/(file transfer time).	N/A	Refer to the definition of delay. * 1 and 3 are available for app.	Stream consistence Audio jitter & video jitter: fluctuation/variation of end-to-end delay from one packet to the next packet within the same packet stream/connection/flow. Or jitter for any stream: caption/subtitle etc.	Raw application data/packet size	Frame loss rate =(broken frame number)/(sequence number)
TRNS	Max throughput can be measured in any point of networking process by (number of bits)/(time) by sending as large file as possible.	throughput can be measured in any point of networking process by (number of bits)/(time)	N/A	1, 2 and 3 Can be measured at any point.	Can be measured at any point	Observe the data/packet size at any point of service process Can be per stream/application or per packet	Statistic of transmission protocol: TCP-like Any point with sequence support or be aware how many packet should expect.
NET	ditto	ditto	N/A	ditto	ditto	ditto	Ditto
LINK	ditto	ditto	Technicians sometimes use it as slang for baud rate, the rate at which symbols may be transmitted through the system. It is also used more colloquially to describe channel capacity, the rate at which bits may be transmitted through the system.	Ditto 3 become packet buffer queuing delay.	Variation of gaps between bits.	ditto	ATM cell lost rate, Buffer overflow, Networking Queuing problem, Data Unit loss rate.
PHY	Maximum throughput <= channel capacity. This is affected by modulation method and physical layer protocol overhead such as error correction coding, bit synchronization and equalizer training sequences. Spectral efficiency = (Max throughput)/(analog bandwidth)	N/A	The frequency range in which the signal's Fourier-transform is nonzero.	**Circuit delay**: process time by integrated circuit is 2. Transmission time in copper wire or air is 1.	Variation between signal peeks.	Physical layer protocol overhead such as error correction coding, bit synchronization and equalizer training sequences.	N/A

Joint Source-Channel Coding



The total bandwidth for source and channel rates is the same for the three curves

TRADE-OFFS



End-to-End Frame Distortion

Expected Distortion for the kth packet

Average Expected Distortion → $$D_{tot} = \frac{1}{K}\sum_{k=1}^{K} E\left[D^k\right]$$

Maximum Expected Distortion → $$D_{tot} = \max_{k=[1,\ldots,K]} \left\{ E\left[D^k\right] \right\}$$

Vector of Expected Distortions → $$D_{tot} = \left[E\left[D^1\right], E\left[D^2\right], \ldots, E\left[D^K\right] \right]^T$$

Resource-Distortion Optimization Framework

- Goal: Minimize transmission cost while limiting the end-to-end distortion and delay.

$$\min \quad \mathrm{C}_{\mathrm{tot}}(S, N) \qquad \text{Transmission Cost}$$

$$s.t.:$$

$$\mathrm{D}_{\mathrm{tot}}(S, N) \le \mathrm{D}_0 \qquad \text{End-to-End Distortion Constraint}$$

$$\mathrm{T}_{\mathrm{tot}}(S, N) \le \mathrm{T}_0 \qquad \text{Transmission Delay Constraint}$$

Resource-Distortion Optimization Framework

- Goal: Minimize the end-to-end distortion while limiting the transmission cost and delay.

$$\min \quad \mathrm{D}_{\mathrm{tot}}(S, N) \qquad \text{End-to-End Distortion}$$

$$s.t.:$$

$$\mathrm{C}_{\mathrm{tot}}(S, N) \le \mathrm{C}_0 \qquad \text{Transmission Cost Constraint}$$

$$\mathrm{T}_{\mathrm{tot}}(S, N) \le \mathrm{T}_0 \qquad \text{Transmission Delay Constraint}$$

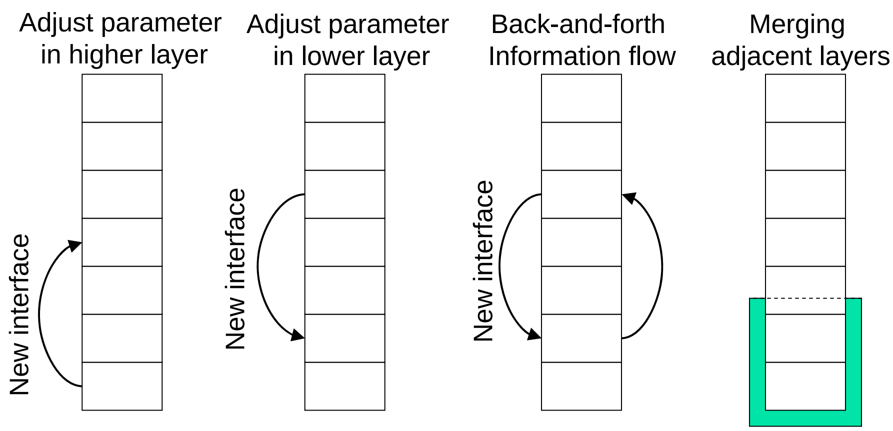
On Cross-Layer Design

- Problem
  - Wireless communication protocols generally operate independently and separately in each layer - this limits adjustability/adaptability in system parameters
- Approach
  - Create interfaces among layers
  - Parameters in one layer could be adjusted based on parameters in other layers
  - For example, adjust physical transmission based on application needs

Introduction

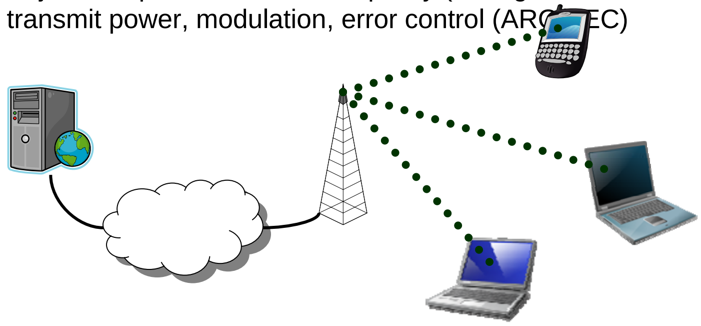
- Objectives/Requirements
  - Optimize system performance
  - Identify parameters in each layer of the protocol stack to be adjusted
  - Maintain integrity of the parameters and reliability/stability of all layers
  - Maintain the structure of the protocol stack
  - Achieve the desired complexity and performance tradeoff

Different Cross-Layer Design Approaches



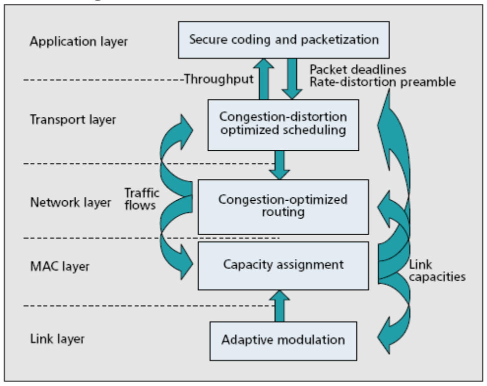
Cross-Layer Design Example

- Video Over Wireless
  - Objective: maximize the quality of video transmission over wireless channel
  - Constraint: limited maximum transmit power, provide rate and delay guarantee
  - Adjustable parameters: video quality (through video encoder), transmit power, modulation, error control (ARQ, FEC)



Cross-Layer Design Example

Application of cross-layer design in ad hoc networks for real-time video streaming



E. Setton et al. ,"Cross-layer design of ad hoc networks for real-time video streaming," IEEE Wireless Communications, vol. 12, no. 4, pp. 59-65, Aug. 2005.

Cross-Layer Design Example

Link Layer:

Transmission rate is adaptively adjusted according to the time-varying channel quality

MAC Layer:

Adjust time slot/channel allocation

Network Layer:

Congestion control is dynamically tuned with MAC protocol to determine the set of network flows

Application Layer:

Determine the video coding scheme according to the lower layer throughput and delay

To CL design or not?

- Good architectural design leads to proliferation and longevity
- Conventional layered architecture is a reasonable way to operate wireless networks and is in fact optimal up to an order
- Unintented cross-layer interactions can have undesirable consequences on overall system performance (e.g., Rate-adaptive MAC and minimum-hop routing)
- Unbridled cross-layer design can lead to spaghetti design, which can stifle further innovation and be difficult to upkeep

V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design", IEEE Wireless Comm., Feb. 2005.

To CL design or not?

- CL design can create loops; "law of unintended consequences" might take over
- Design in the presence of interacting dynamics needs care
  - Stability
  - Robustness
  - Timescale separation might need to be employed
- Architecture in system design pertains to breaking down a system into modular components, and systematically specifying the interactions between the components
  - The von Neumann architecture
  - The OSI architecture for networking
  - Source-channel separation and digital comm system architecture
- Performance vs Architecture
  - Performance optimization leads to short-term gain
  - Architecture is usually based on longer-term considerations
  - Short-term vs long-term gains

Architectural considerations for wireless networks

- Successful architectural designs by their pervasiveness influence how designers think
- It is not obvious that the layered architecture for wired networks is a priori appropriate for wireless networks
- The wireless medium allows modalities of communication that are nonexistent for wired networks
  - Multihop architecture with decoding and forwarding at each relay node creates special problems for wireless networks that need to be solved
    - A MAC protocol is needed to control the number of interferers
    - Routing problem (no problem for one hop)
    - Power control (interferer treated as noise)

Fundamental Properties of the Wireless Medium

- There is no intrinsic concept of a link
- No notion of a switch
- Infinitude of possibilities for operating wireless networks (cooperating nodes)
- The multihop decode and forward, treating interference as noise, is order optimal with respect to the transport capacity when load across nodes can be balanced by multipath routing