Ph.D Defense

Optimal Cross-Layer Resource Allocation for Real-Time Video Transmission over Packet Lossy Networks

Fan Zhai

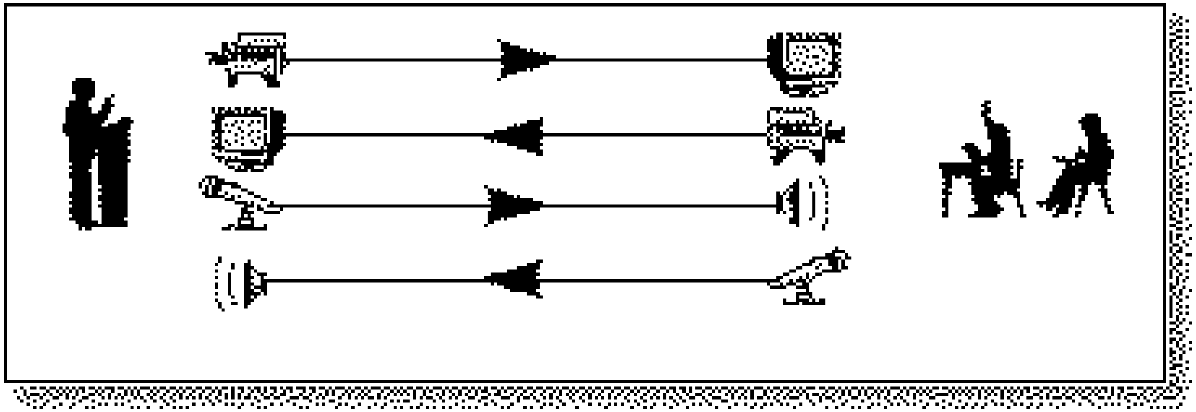
Dept. Electrical & Computer Engineering

Northwestern University

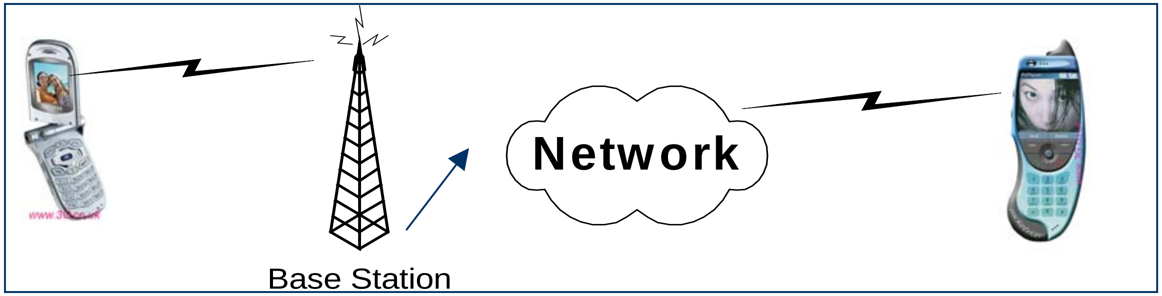
Feb. 9, 2004

Current Streaming Video Applications

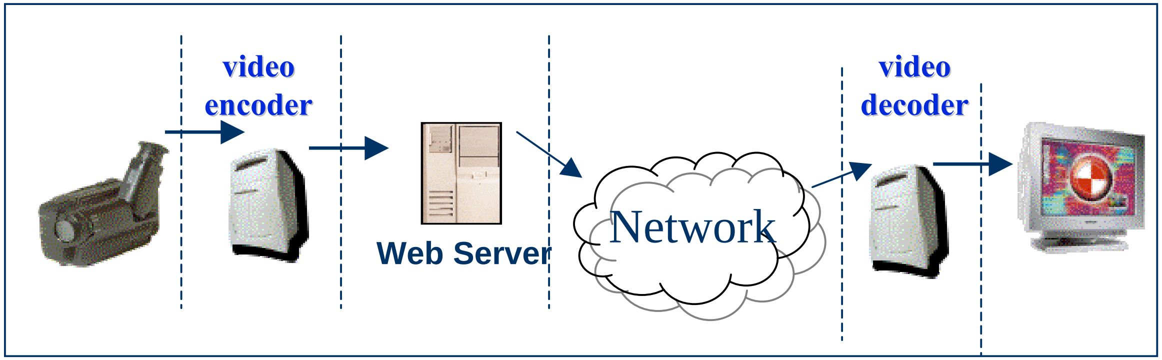
- ➢ Videoconferencing
- ➢ Distance learning



- ➢ Videophone

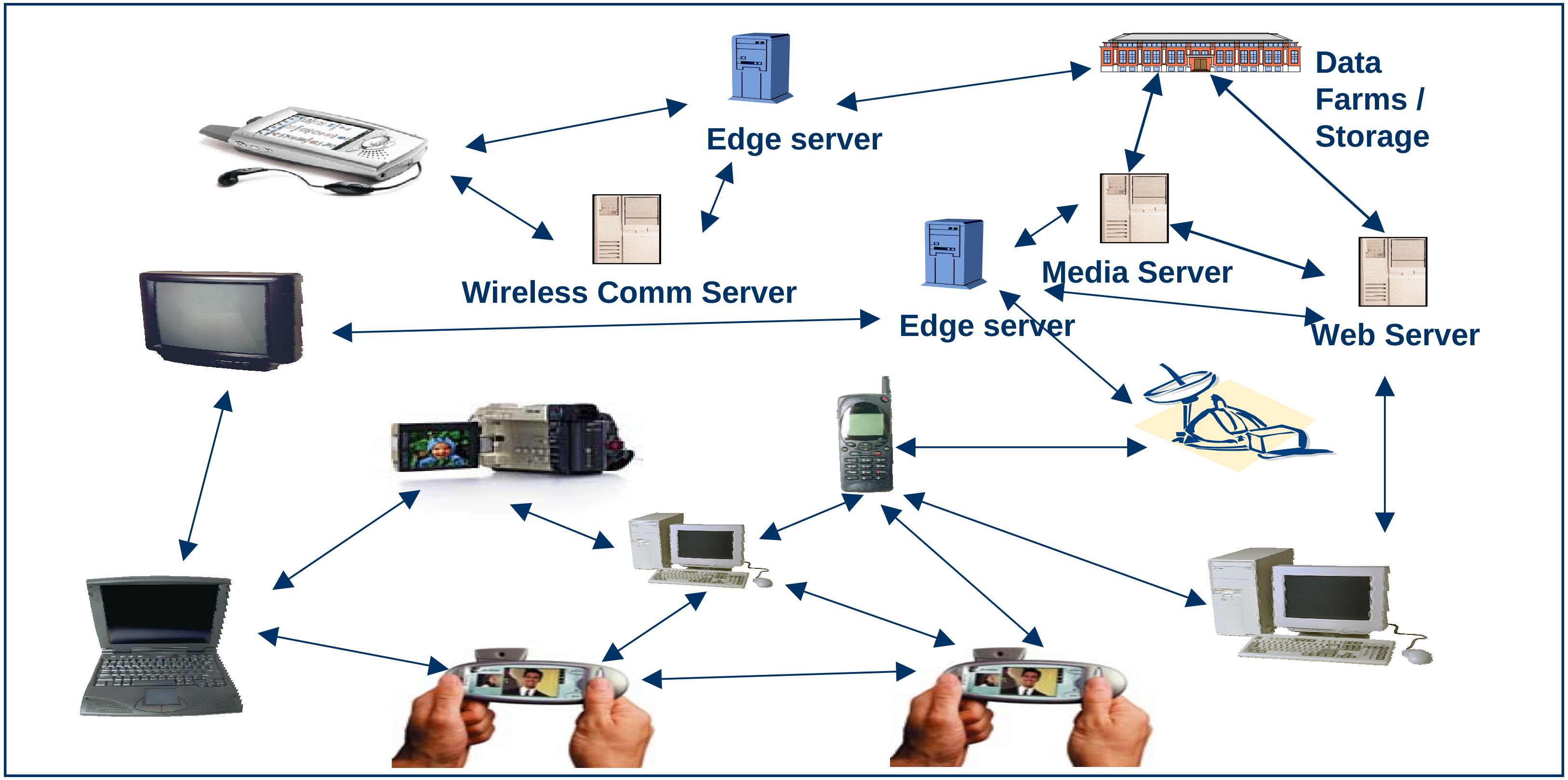


- ➢ On-demand streaming



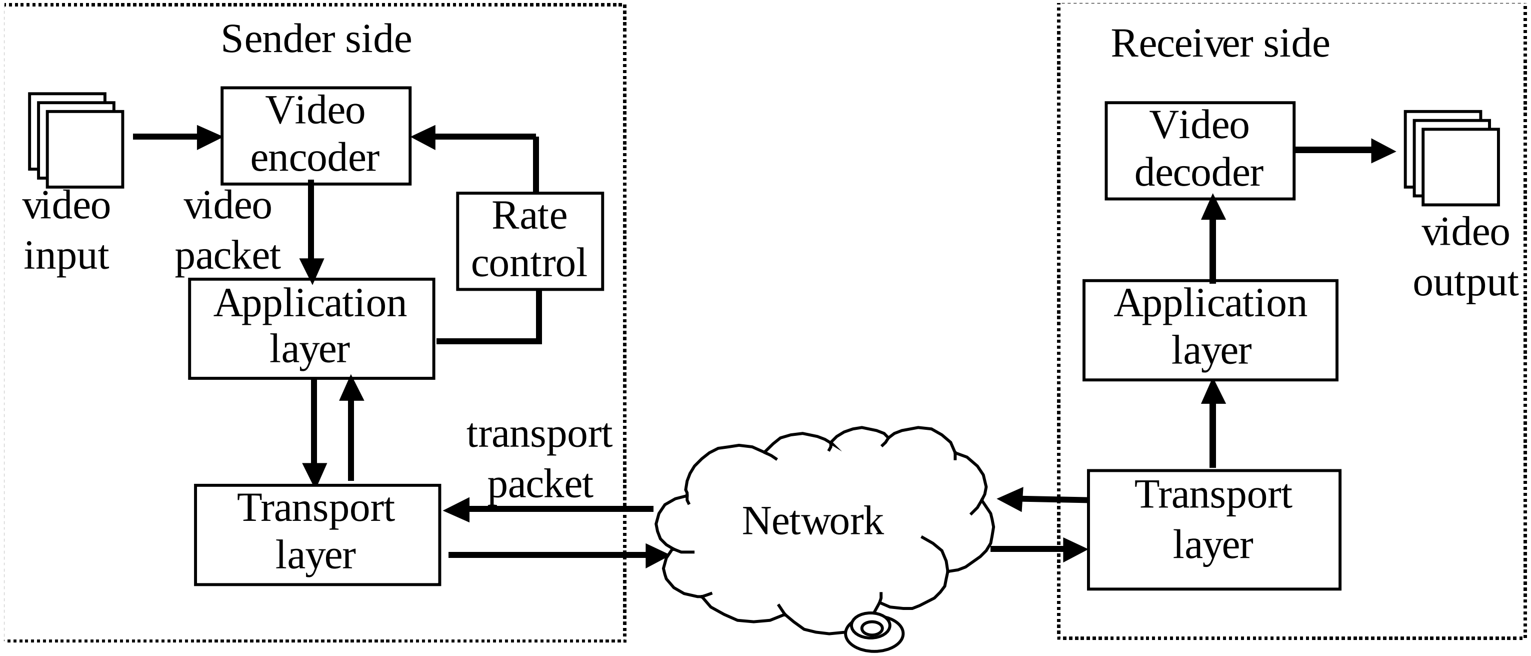
Emerging Streaming Video Applications

Anywhere, Anytime and Anyone



- Cellular mobile networks: 2.5G, 3G, 4G systems
- Wireless LANs: IEEE 802.11, Hiper LAN 2, Bluetooth

Video Transmission System



- Encoder: compression
- Rate control: to constraint bit rate based on estimated CSI
- Protocol stack: RTP/UDP/IP
- Network: packet loss and delay
- Decoder: display video in real-time; error concealment
- Max end-to-end delay (setup time): application-dependent

Challenges and Approaches

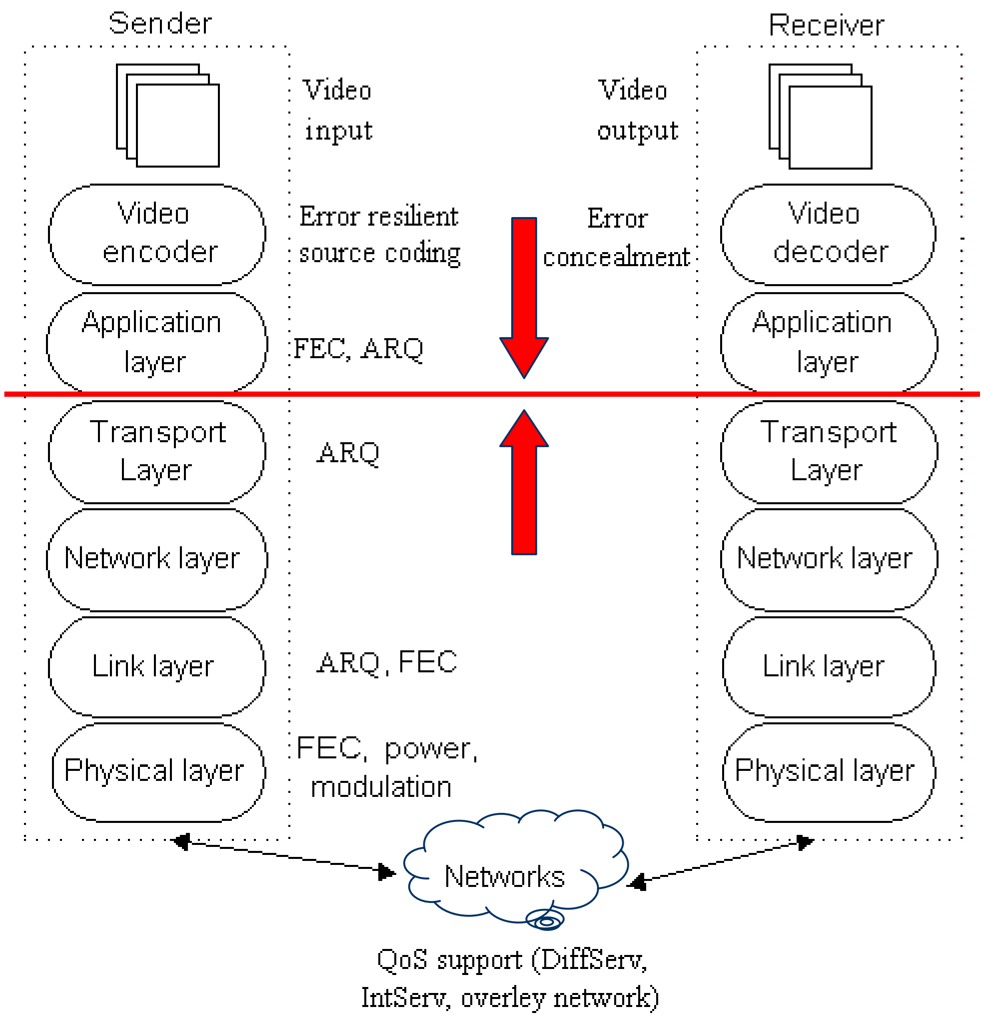
Challenges

- Bandwidth (limited, time-varying)
  - Internet: congestion
  - Wireless: fading, shadowing
- QoS (packet loss, delay)
  - Distortion
  - Strict end-to end delay constraint
- Limited resources
  - Internet: bandwidth, buffer
  - Wireless: battery,bandwidth

Approaches

- Network-adaptive
- Error control
  - Unequal error protection
- Cross-layer design

Cross-Layer Design



Error control Components

- Error resilient source coding
  - adapt source coding to channel
- FEC (forward error correction)
  - modify channel characteristics
- ARQ (delay constraint)
- Power control
  - modify channel characteristics
- Modulation, Rate adaptation
- QoS support
  - reserve resource to support QoS
- Error concealment
  - recover/conceal the error at decoder

Related Work

- Optimization on each error control component
  - Error resilient source coding: R. Zhang, et al, ’00
  - FEC: B. Hong, et al ’02
  - ARQ: P. A. Chou, et al, ’01
  - Error concealment: Y. Wang et al, ’00
- Cross-layer design
  - Joint source-channel coding: M. Gallant, et al, *IEEE CSVT’01*
  - Joint source coding and power control: Y. Eisenberg, et al, *CSVT*’02
  - Joint source-channel coding and power control:Y.S. Chan, et al, *JSAC* ’02, *sub*
  - Joint source-channel coding and power control:S. Appadwedula, et al, *ICIP’98*
- Difference in our work
  - The available error control components are considered in an integrated manner
  - We consider error resilient source coding
  - We consider error concealment

Presentation Outline

- Resource-distortion optimization framework
- Internet: JSCC -- Joint source-channel coding
- Wireless: JSCCPA -- Joint source-channel coding and power allocation
- DiffServ: JSCPC -- Joint source coding and packet classification
- Extensions to scalable video transmission
- Conclusions and future work

Resource-Distortion Optimization

- General formulation

$$\min_{\{\boldsymbol{\mu},\boldsymbol{\nu}\}} E[D(\boldsymbol{\mu},\boldsymbol{\nu})]$$ End-to-End Distortion

$$\text{s.t.} \quad C(\boldsymbol{\mu},\boldsymbol{\nu}) \le C_0$$ Cost Constraint

$$T(\boldsymbol{\mu},\boldsymbol{\nu}) \le T_0$$ Transmission Delay Constraint

- μ : Source coding parameter vector
- ν : Resource allocation parameter vector
- C_0 is explicitly determined by specific application
- T_0 is more implicitly determined by the application
  - obtained from higher-level rate controller, which is not incorporated into this work.

End-to-End Distortion

- Consider end-to-end distortion

Original

Encoded

Decoded

- Distortion of the *i*-th pixel in the *n*-th frame

$$E[d_i^{(n)}] = E[(f_i^{(n)} - \tilde{f}_i^{(n)})^2] = (f_i^{(n)})^2 - 2 f_i^{(n)} E[\tilde{f}_i^{(n)}] + E[(\tilde{f}_i^{(n)})^2]$$

- Distortion for the *k*-th packet that has *M* pixels

$$E[D_k] = \sum_{i=1}^{M} E[d_i^{(n)}] / M$$

ROPE (Recursive Optimal Per-pixel Estimate)

- 1^{st} order expected value (same fashion for 2^{nd} order)
  - Intra: $\mathrm{E}[\tilde{f}_i^{(n)}] = (1-\rho_k)\mathrm{E}[\tilde{f}_i^{(n)}] + \rho_k \mathrm{E}[\tilde{f}_i^{(n-1)}]$
  - Inter: $\mathrm{E}[\tilde{f}_i^{(n)}] = (1-\rho_k)(\hat{e}_i^{(n)} + \mathrm{E}[\tilde{f}_j^{(n-1)}]) + \rho_k \mathrm{E}[\tilde{f}_i^{(n-1)}]$
- In order to calculate the distortion for one frame, only need the 1^{st} order and 2^{nd} order expected reconstructed pixel value in the previous frame
- This algorithm is per-pixel accurate and can fully capture error propagation.

- R. Zhang, S.L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, June 2000.

Expected Distortion

- Expected distortion for the *k*-th packet

$$E[D_k] = (1-\rho_k)\,E[D_{R,k}] + \rho_k\,E[D_{L,k}]$$

Probability of Reception

Expected distortion if the packet is received

Probability of Loss

Expected distortion if the packet is lost

- $E[D_{R,k}]$, $E[D_{L,k}]$: depends on source coding parameter μ_k
- $\rho_k(\varepsilon, \nu)$: depends on CSI and resource allocation parameter

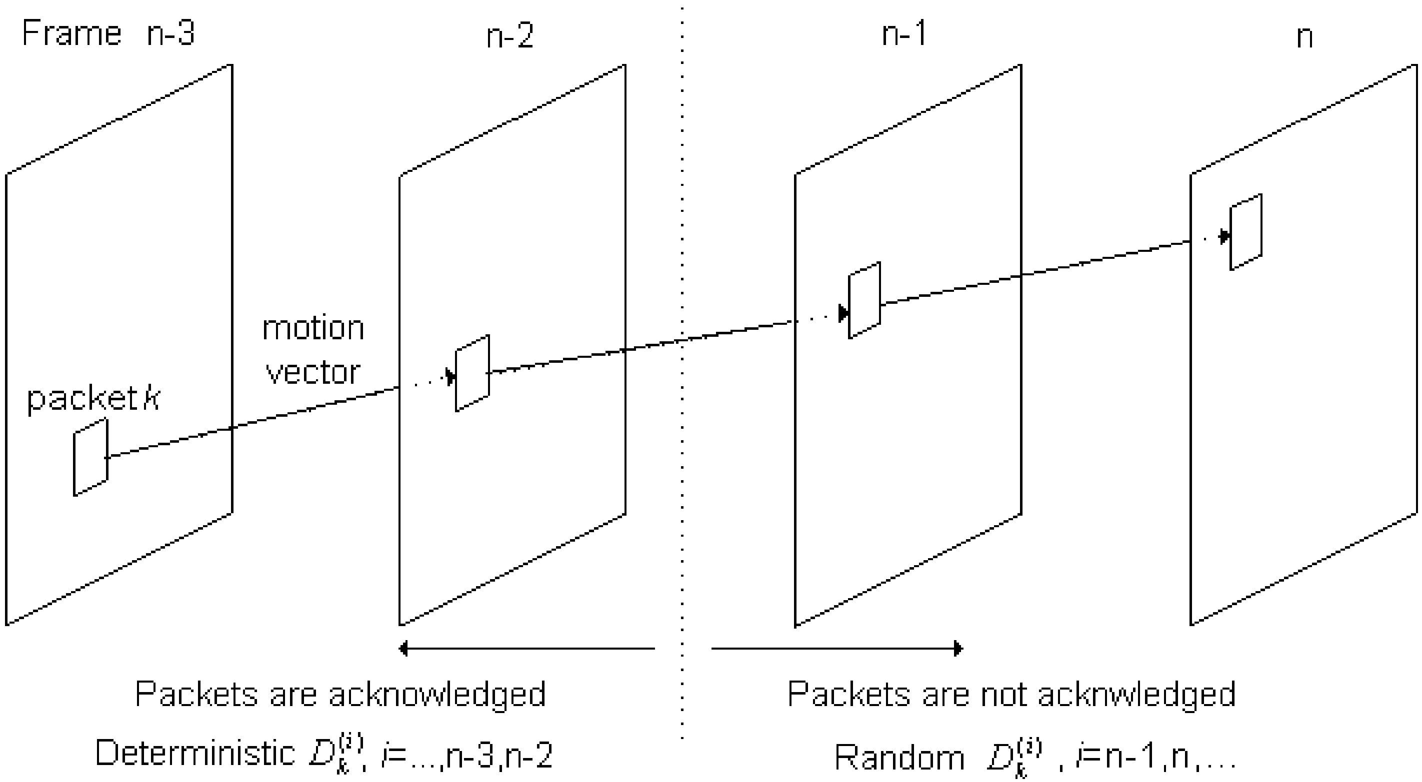
FEC parameter

$\varepsilon \xrightarrow{\nu_k} \rho_k$

Prob of loss for transport packet

Prob of loss for source/video packet

Distortion Estimation Based on Feedbacks



- Expectations are taken with respect to the updated probability distribution of channel losses given the available feedback

Transmission Delay and Cost

- Transmission Delay:

Number of packets in a frame

$$T = \sum_{k=1}^{M} \frac{B_k(\mu_k)}{R_T}$$

Source coding parameter

Packet index

Transmission rate

Number of bits

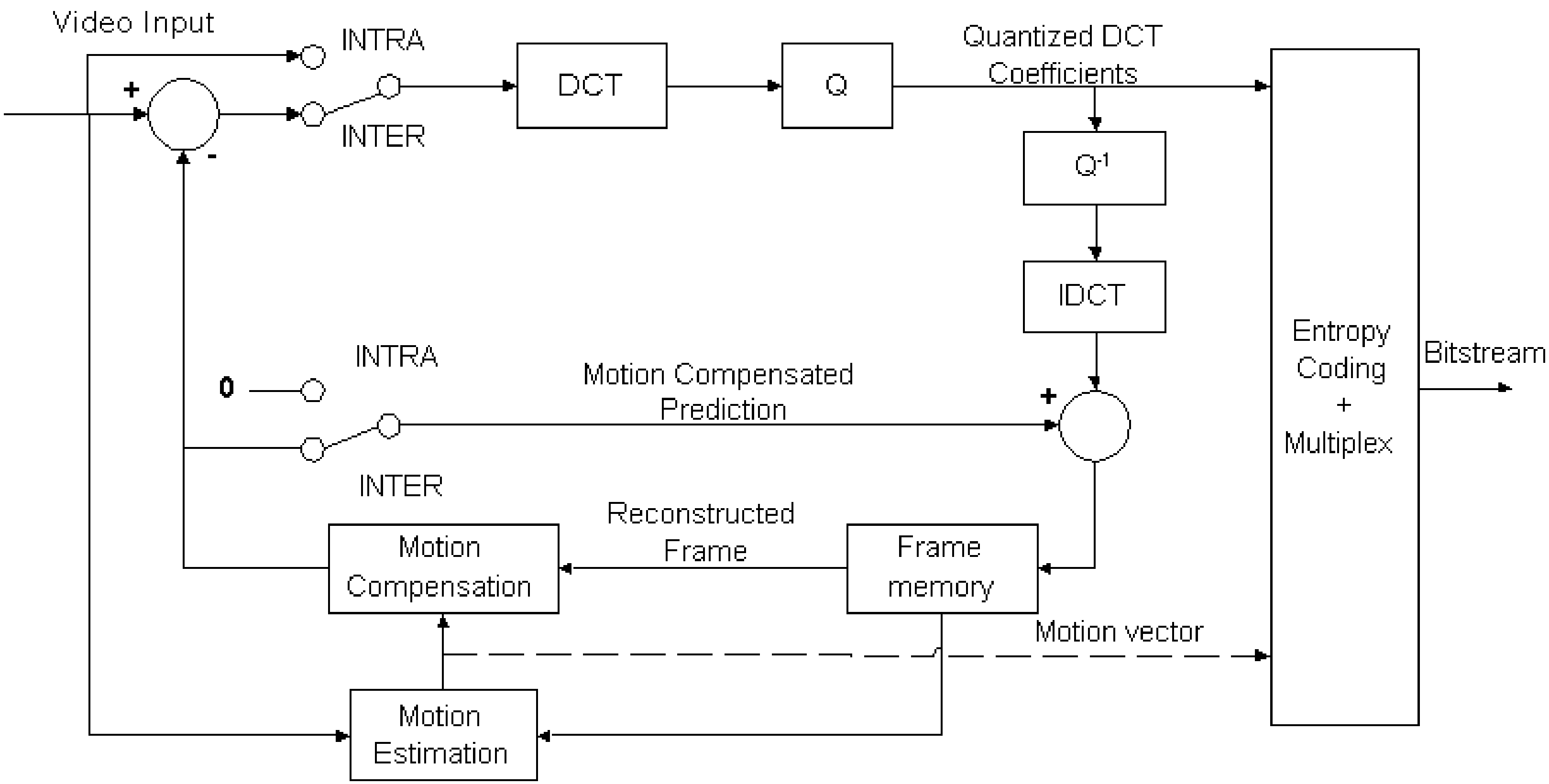
Power allocation parameter

- Energy:

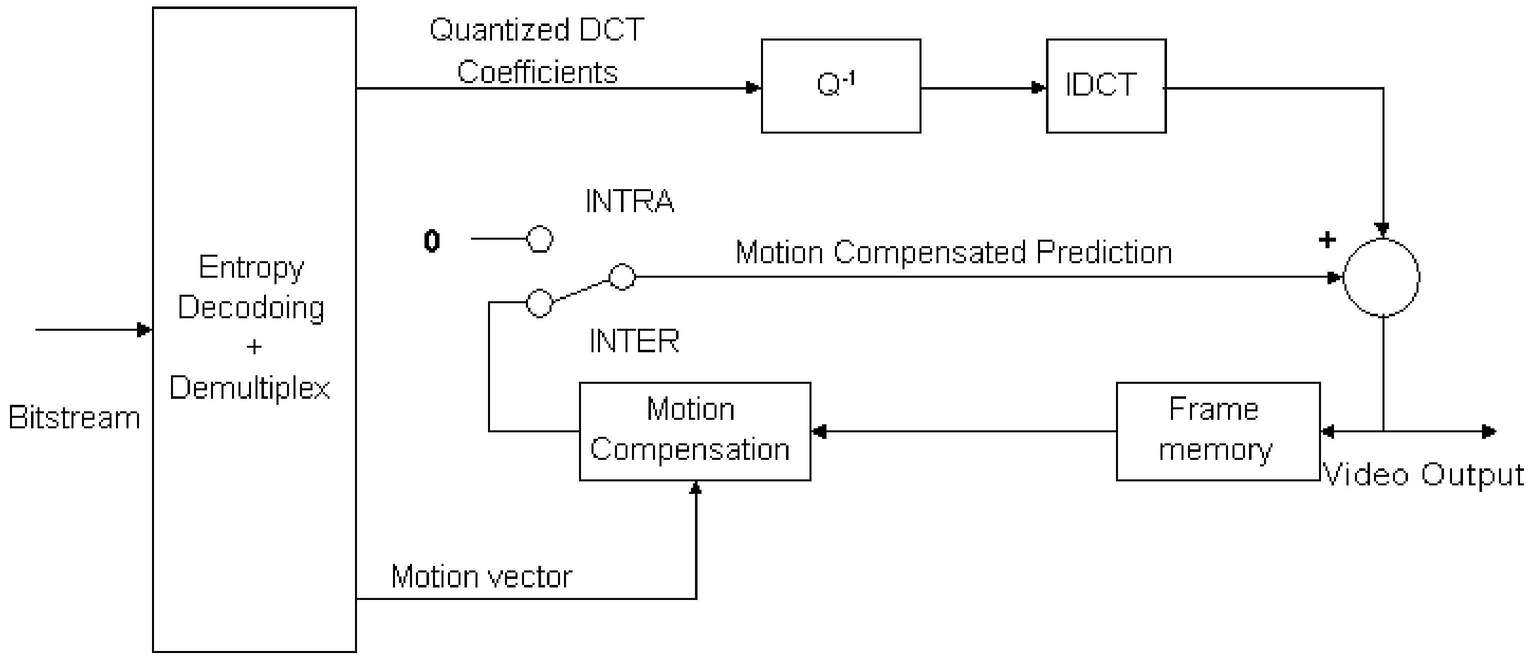
$$C = \sum_{k=1}^{M} \frac{B_k(\mu_k) P_k(\nu_k)}{R_T}$$

Transmission power

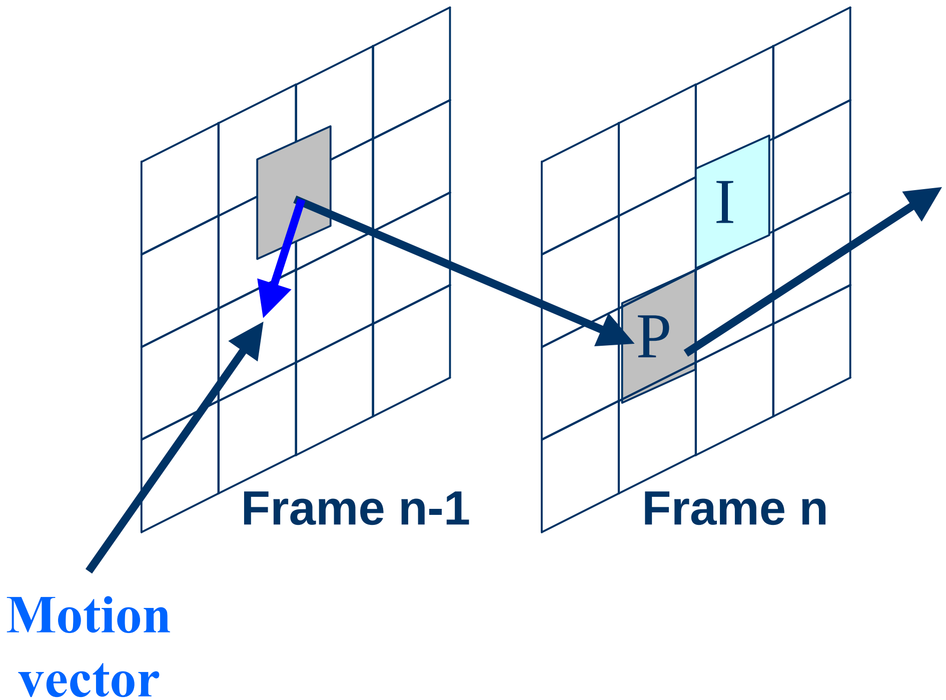
Motion-Compensated Video Encoder



Motion-compensated Video Decoder



Motion-compensated Video Codec



DCT coefficients

120	60	40	30	4	3	0	0
70	48	32	3	4	1	0	0
50	36	4	4	2	0	0	0
40	4	5	1	1	0	0	0
5	4	0	0	0	0	0	0
3	2	0	0	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantizer

Quantized coefficients

12	6	3	2	0	0	0
7	3	2	0	0	0	0
3	2	0	0	0	0	0
2	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

10	10	15	20	25	30	35	40
10	15	20	25	30	35	40	50
15	.20	25	30	35	40	50	60
20	25	30	35	40	50	60	70
25	30	35	40	50	60	70	80
30	35	40	50	60	70	80	90
35	40	50	60	70	80	90	100
40	50	60	70	80	90	100	110

Quantization table

- Source coding parameter μ_k
  - Prediction mode (Intra, Inter, Skip)
  - Quantization stepsize
- Different encoding modes result in different levels of coding efficiency and error robustness

Channel Model

- Network model
  - Independent time-invariant packet erasure channel
- Packet loss
  - Internet:
    - Before FEC : ε -- Bernoulli process
    - After FEC : $\rho_k = f(\varepsilon, v_k)$
  - Wireless:

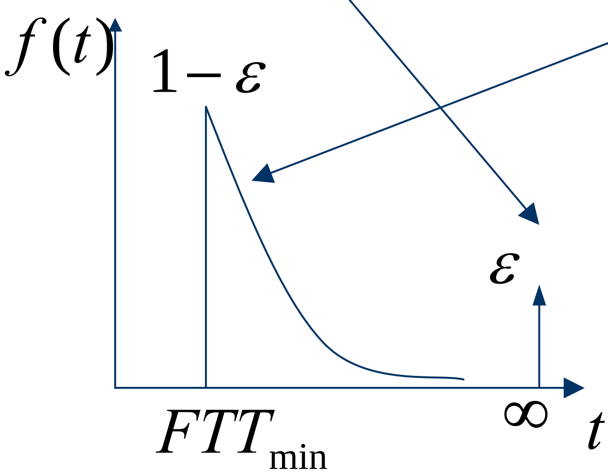
$$\rho_k = 1 - (1 - p_{e,k})^{B_k}$$

    - i.i.d fading
    - B_k : Packet size
    - $p_{e,k}$: BER after channel coding (depending FEC, channel SNR, and fading model)

Channel Model—Internet

- Network model :
  - Independent time-invariant packet erasure + random delays
  - Receiver responds to a received/lost packet with a positive/negative acknowledgement.
  - Perfect feedback channel:constant delay and no error

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



- 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

Channel Model—Wireless

- Packet-erasure channel, from the point of view of the applications
- Channel BER (Bit Error Rate)

$$\text{BER} = \frac{1}{2}\left(1 - \sqrt{\frac{\alpha E_b}{N_0 + \alpha E_b}}\right)$$ Rayleigh fading

$$\text{BER} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$ AWGN

- Probability of Packet Loss

$$\rho_k = 1 - (1 - p_{e,k})^{B_k}$$

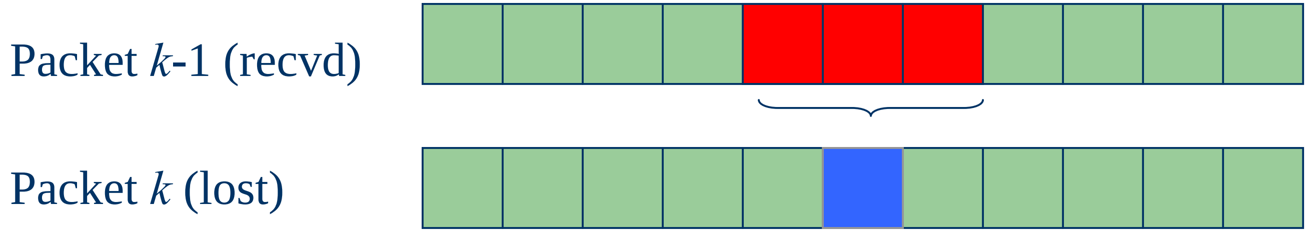
$p_{e,k}$: BER after channel coding

B_k : Packet size

- Wireless channel: probability of packet loss depends on source coding parameter, channel coding parameter, and power level

Implementation Issues

- Packetization: each row of blocks (GOB) is packetized into a packet
- Error concealment: temporal concealment, median motion vectors

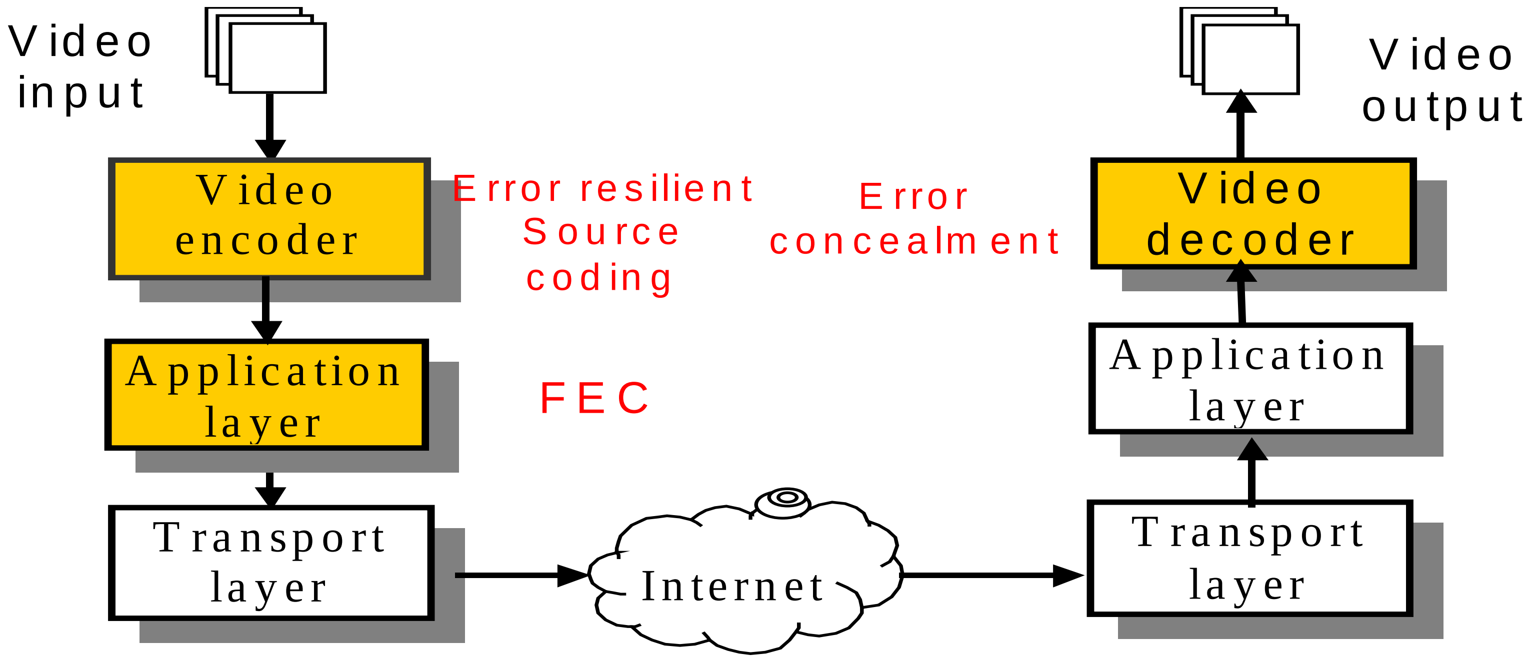


- Expected distortion

$$E[D_k] = (1-\rho_k)\,E[D_{R,k}] + \rho_k\,E[D_{L,k}]$$

$$E[D_k] = (1-\rho_k)\,E[D_{R,k}] + \rho_k\,(1-\rho_{k-1})\,E[D_{C,k}] + \rho_k\,\rho_{k-1}\,E[D_{Z,k}]$$

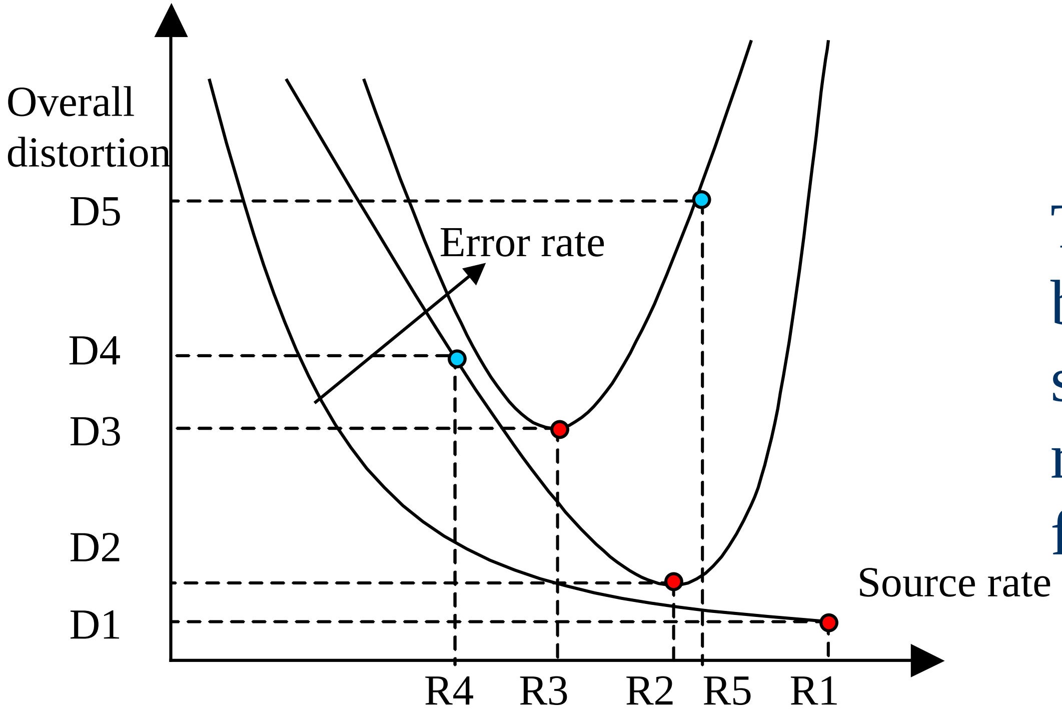
Internet Video Transmission



- We jointly consider error resilient source coding, FEC, and error concealment

Joint Source-Channel Coding

- Shannon's separation theorem does not strictly hold due to the delay and complexity constraints



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

- Overall distortion = source distortion + channel distortion
- Optimal bit allocation depends on the channel characteristics

Sequential Joint Source-Channel Coding

- Step 1: Bit allocation between source and channel

$$\min_{\{v \in R\}} E[D(\boldsymbol{v})]$$

$$s.t. \quad T(\boldsymbol{v}) = B_s(\boldsymbol{\mu}(\boldsymbol{v}))/R_T + B_c(\boldsymbol{v})/R_T \le T_0$$

- Step 2: Source coding based on given bit budget

$$\min_{\{\mu \in Q\}} E[D(\boldsymbol{\mu})]$$

$$s.t. \quad T_s(\boldsymbol{\mu}) = B_s(\boldsymbol{\mu})/R_T \le T_{s,0}$$

μ : source coding parameter

ν : channel coding parameter

T_0: transmission delay constraint

B_s : source bits

B_c : channel bits

$T_{s,0}$: transmission delay constraint for the source

Integrated Joint Source-Channel Coding

- Solve (1) Bit allocation between source and channel and (2) source coding and channel coding, in one step

$$\min_{\{\boldsymbol{\mu}\in Q, \boldsymbol{\nu}\in R\}} E[D(\boldsymbol{\mu},\boldsymbol{\nu})]$$

$$s.t. \quad T(\boldsymbol{\mu},\boldsymbol{\nu}) = B(\boldsymbol{\mu},\boldsymbol{\nu}) / R_T \le T_0$$

B: total bits used for both source and channel coding

R_T : transmission rate

T_0 : transmission delay constraint for this frame

F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, A. K. Katsaggelos, "An integrated joint source-channel coding framework for video transmission over packet lossy networks," *IEEE ICIP'04*, submitted.

Experimental Results

- System 1: Integrated JSCC

$$\min_{\{\boldsymbol{\mu}\in Q,\boldsymbol{\nu}\in R\}} E[D(\boldsymbol{\mu},\boldsymbol{\nu})]$$

- System 2: Fixed channel coding rate

$$\min_{\{\boldsymbol{\mu}\in Q\}} E[D(\boldsymbol{\mu},\nu_0)]$$

- System 3: Sequential JSCC

$$\min_{\{\nu\in R\}} E[D(\boldsymbol{\nu})]$$

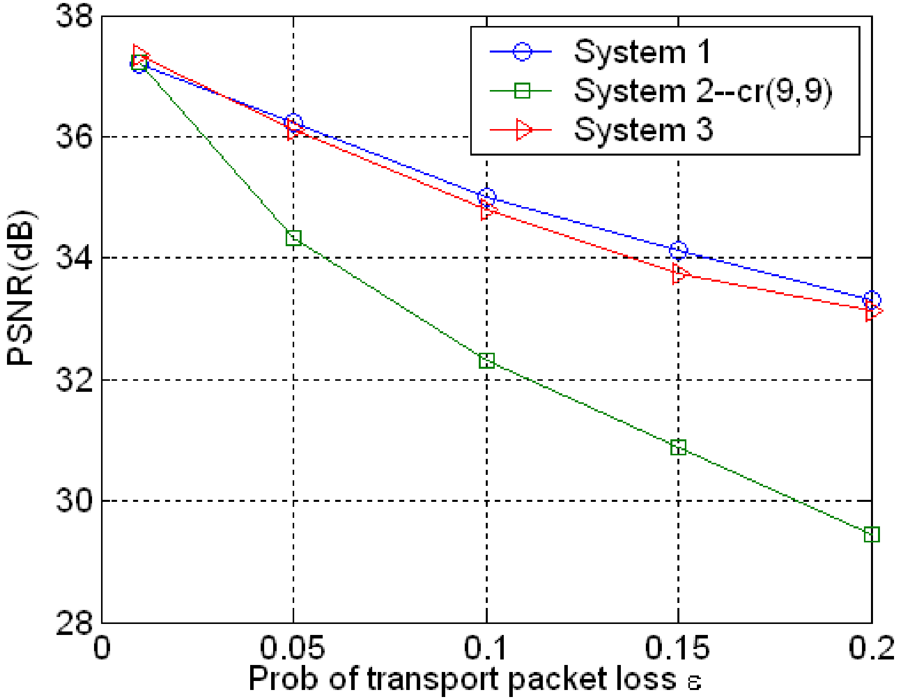
$$s.t.\quad T(\boldsymbol{\nu}) = B_s(\boldsymbol{\mu}(\boldsymbol{\nu}))/R_T + B_c(\boldsymbol{\nu})/R_T \le T_0$$

$$\min_{\{\mu\in Q\}} E[D(\boldsymbol{\mu})]$$

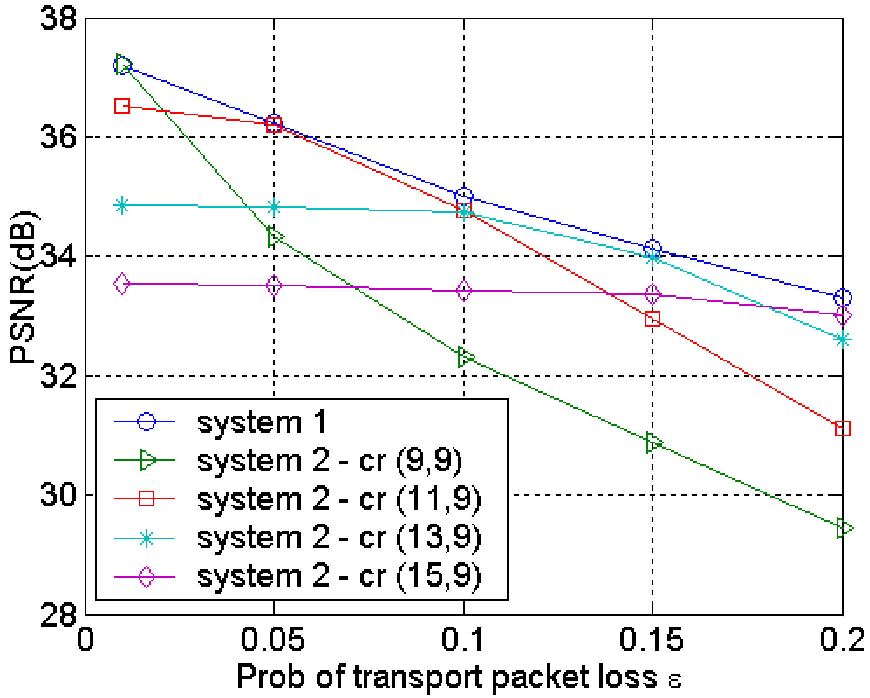
$$s.t.\quad T_s(\boldsymbol{\mu}) = B_s(\boldsymbol{\mu})/R_T \le T_{s,0}$$

Experimental Results

Average PSNR vs. transport packet loss probability



System 1 vs. System 2 and 3

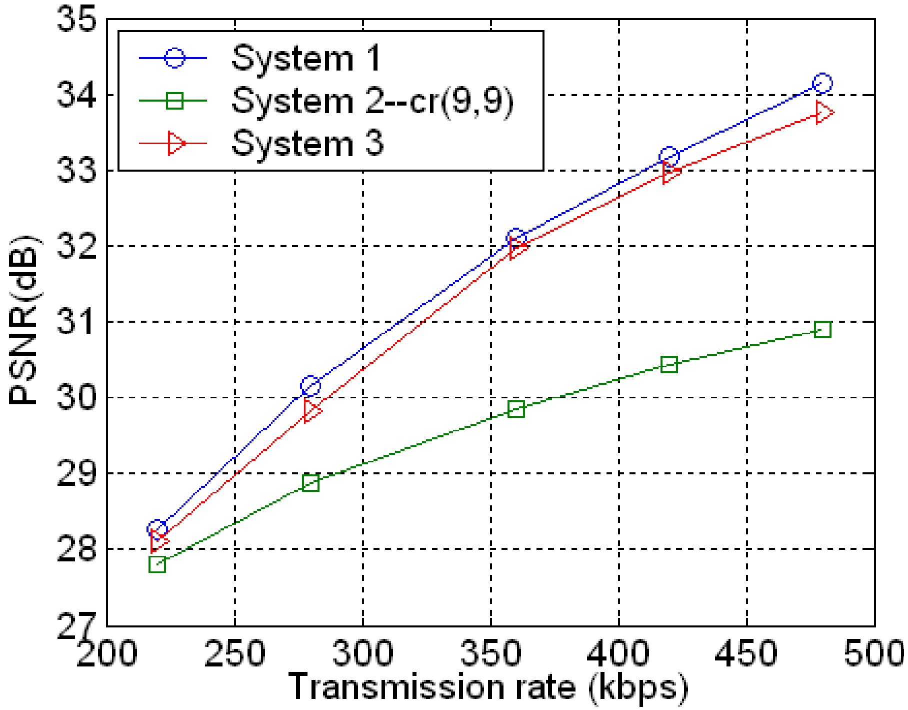


System 1 vs. System 2 with indicated channel rates

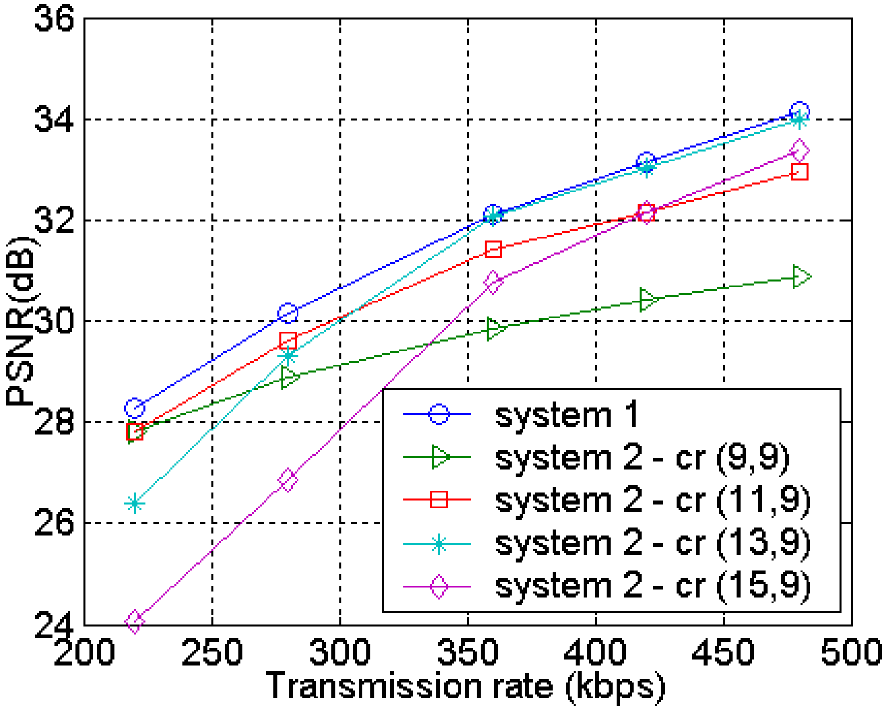
(R_T=480 kbps, F=15 fps, cr in the legend denotes channel rates).

Experimental Results

Average PSNR vs. transmission rate



System 1 vs. System 2 and 3

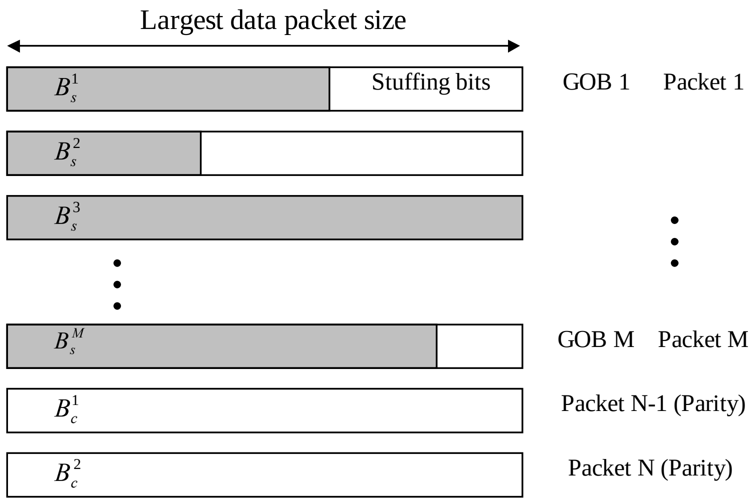


System 1 vs. System 2 with indicated channel rates

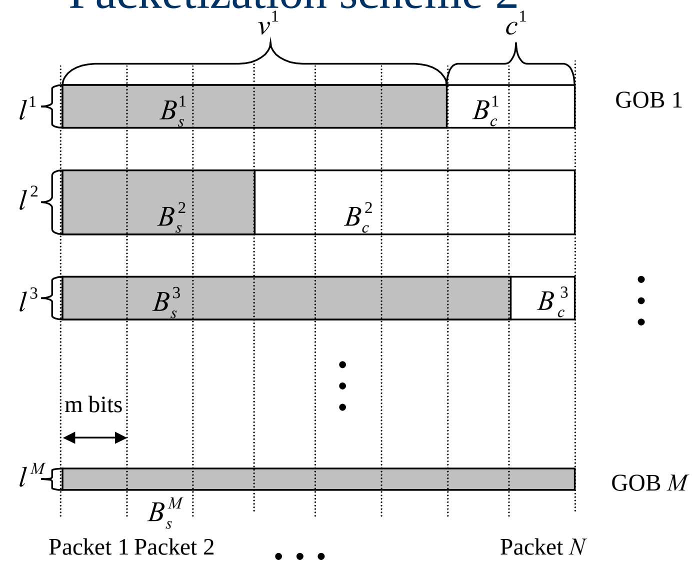
(ε=0.15, F=15 fps, cr in the legend denotes channel rates).

JSCC—Packetization

Packetization scheme 1



Packetization scheme 2



RS(*N*,*M*)

$$\rho = \varepsilon P_b(N-1,k) = \varepsilon\left(1-\sum_{i=0}^{N-1-k}\binom{N-1}{i}\varepsilon^i(1-\varepsilon)^{N-1-i}\right)$$

RS(*N*,v^k)

$$\rho^k = P_b(N,v^k) = 1-\sum_{i=0}^{N-v^k}\binom{N}{i}\varepsilon^i(1-\varepsilon)^{N-i}$$

F. Zhai, Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, A. K. Katsaggelos, "Packetization schemes for forward error correction in Internet video streaming," *Proc. 41st Allerton Conf. Communication, Control, and Computing,* Oct. 2003.

JSCC — Solution Algorithm

- Lagrangian relaxation and DP

$$\min_{\{\boldsymbol{\mu},\boldsymbol{\nu}\}} \sum_{k=1}^{M} J_k = \sum_{k=1}^{M} E[D_k(\mu,\nu)] + \lambda B_k(\mu_k,\nu)/R_T$$

- Packetization 1

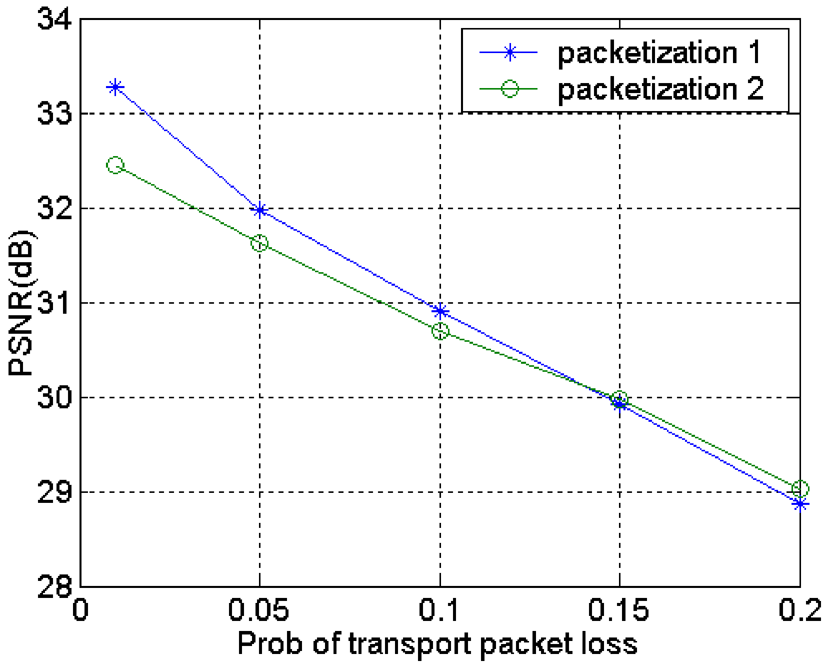
$$\min_{\{\nu\}} \sum_{k=1}^{M} J_k(\boldsymbol{\mu}^* \mid \nu) = \min_{\{\nu\}} \left\{ \min_{\{\boldsymbol{\mu}\}} \sum_{k=1}^{M} J_k(\boldsymbol{\mu},\nu) \right\}$$

- Packetization 2

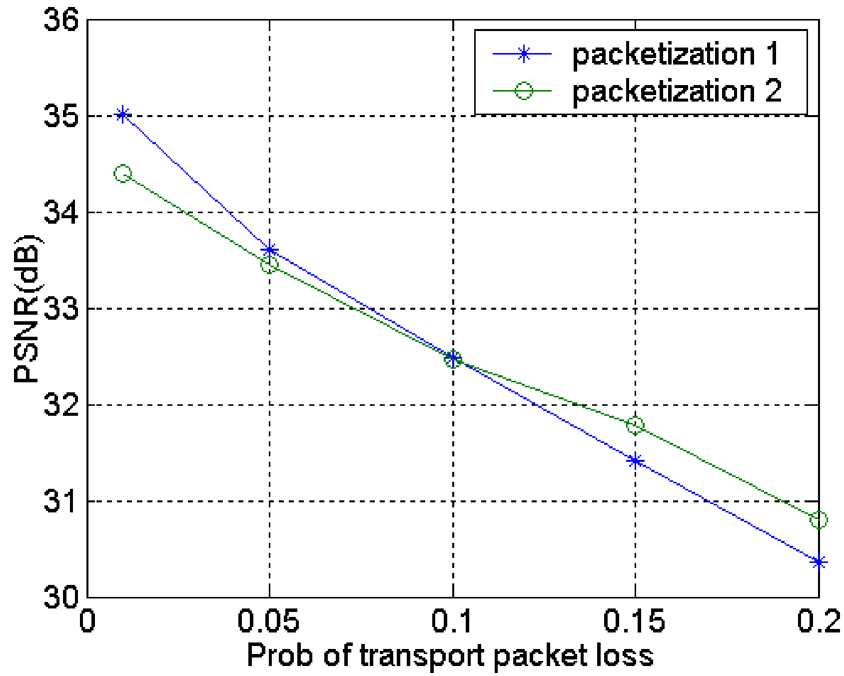
$$\min_{\{\boldsymbol{\mu},\boldsymbol{\nu}\}} \sum_{k=1}^{M} J_k = \min_{\{\boldsymbol{\mu},\boldsymbol{\nu}\}} \sum_{k=1}^{M} J_k(\mu_{k-1},\mu_k,\nu_{k-1},\nu_k)$$

Experimental Results

Packetization scheme 1 vs 2



Transmission rate 360Kbps



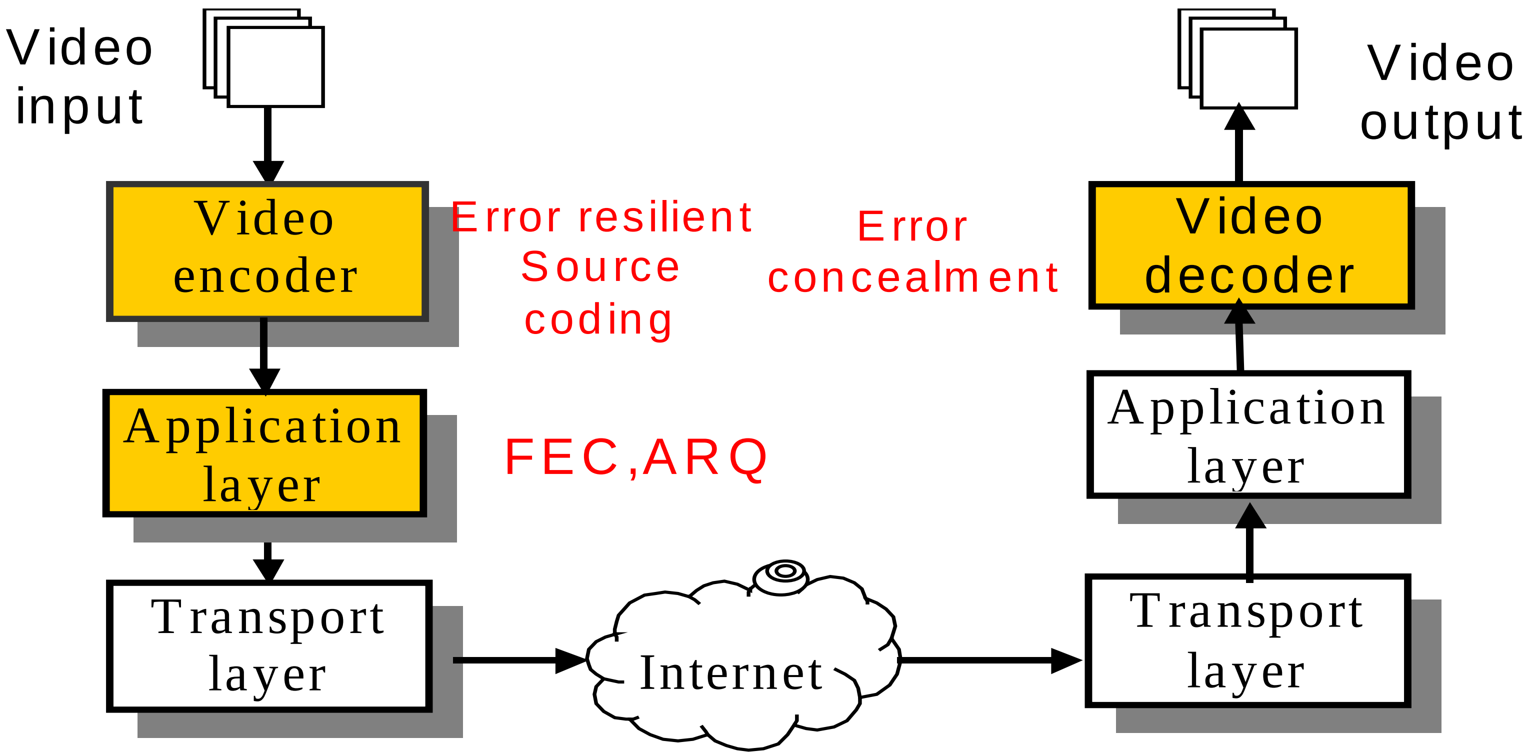
Transmission rate 480Kbps

RTP/UDP/IP header = 40bytes

FEC vs. ARQ

- FEC:
  - Usually preferred for real-time video applications
  - Cannot completely avoid packet loss
  - Incur constant overhead even when the channel has no loss
  - Depend on the accurate CSI estimation
- ARQ:
  - Can automatically adapt to the channel loss characteristics
  - Longer delay
  - Useful for long end-to-end delay applications, e.g., on-demand video streaming
  - Useful for short RTT situations, e.g., LAN

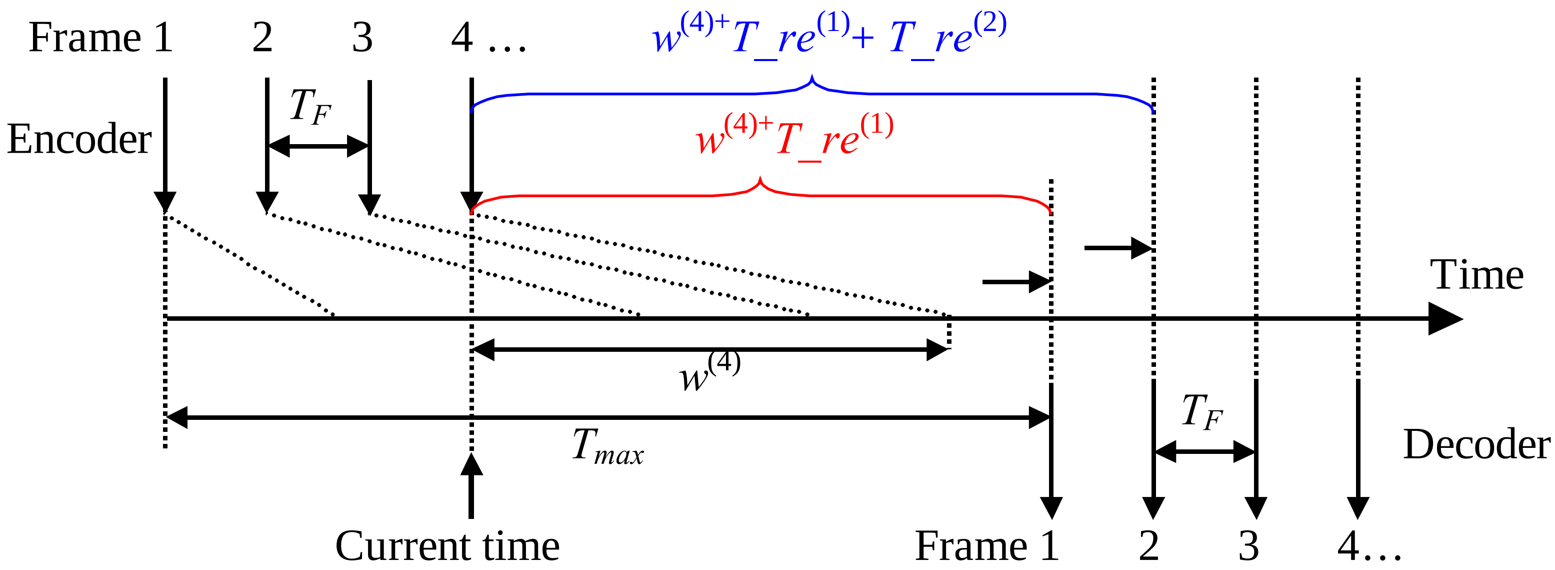
JSCC—Hybrid FEC/Retransmission



- We jointly consider error resilient source coding, FEC, application-layer ARQ, and error concealment

F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Rate-distortion optimized hybrid error control for real-time packetized video transmission," *IEEE ICC'04*, accepted.

Delay-Constraint Retransmission



- T_{max}: Maximum end-to-end delay (imposed by application)
- $w^{(n)}$: waiting time for the n-th frame in the encoder buffer
- $T\_re^{(n)}$: transmission delay for retransmitting packets in the n-th frame
- T_F : One frame's time (=1/T_F)

Problem Formulation

$$\begin{cases} \min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{\sigma}\}} \sum_{i=0}^{A} E[D^{(n-i)}] = \sum_{i=1}^{A} E[D^{(n-i)}(\boldsymbol{\sigma}^{(n-i)})] + \sum_{k=1}^{M} E[D_k^{(n)}(\boldsymbol{\mu},\gamma)] \\ \text{s.t. } w^{(n)} + \sum_{i=1}^{j} \sum_{k=1}^{M} \sigma_k^{(n-i)} T_k^{(n-i)} \le T_{\max} - (A+1-j)T_F \quad j = 1 \ldots A \\ \quad w^{(n)} + \sum_{i=1}^{A} \sum_{k=1}^{M} \sigma_k^{(n-i)} T_k^{(n-i)} + \sum_{k=1}^{M} T_k^{(n)} \le T_{\max} \end{cases}$$

- A: # of frames that retransmission is eligible
- $\sigma^{(n)}=\{\sigma_1^{(n)},\ldots,\sigma_M^{(n)}\}\in\{0,1\}$: Retrans parameter vector for frame n
- γ : FEC parameter (using RS code)
- $T_k^{(n)}$: Transmission delay for packet k in frame n.
- T_{max} can be replaced by T_{max}-$\Delta T^{(n)}$, where $\Delta T^{(n)}$ is decided by a rate controller.

Problem Formulation

$$
\begin{cases}
\min\limits_{\{\boldsymbol{\mu},\gamma,\boldsymbol{\sigma}\}} \sum\limits_{i=0}^{A} E[D^{(n-i)}] = \sum\limits_{i=1}^{A} E[D^{(n-i)}(\boldsymbol{\sigma}^{(n-i)})] + \sum\limits_{k=1}^{M} E[D_k^{(n)}(\boldsymbol{\mu},\gamma)] \\
\text{s.t.} \quad \sum\limits_{i=1}^{A}\sum\limits_{k=1}^{M} \sigma_k^{(n-i)} T_k^{(n-i)} + \sum\limits_{k=1}^{M} T_k^{(n)} \le T_0^{(n)}
\end{cases}
$$

- We assume $T_0^{(n)}$ is given from higher-level rate controller
- Optimization with a sliding window of A+1 frames
- Optimization decisions: retransmission policy for the first A frames, and source coding and FEC for the current frame
- Based on the updated probabilities of packet loss, recalculate the expected distortion of all packets in the window
- Optimization window shifts at the frame level

Probability of Packet Loss for the Current Frame

- Probability of packet loss

$$\rho_k^{(n)} = \rho_{k,FEC}^{(n)} \rho_{k,RET}^{(n)}$$

Due to FEC

Due to retransmission

- The probability of loss in future retransmissions can only be estimated since the acknowledgement information and retransmission decisions are not available in the encoding of the current frame

- Approximation formula

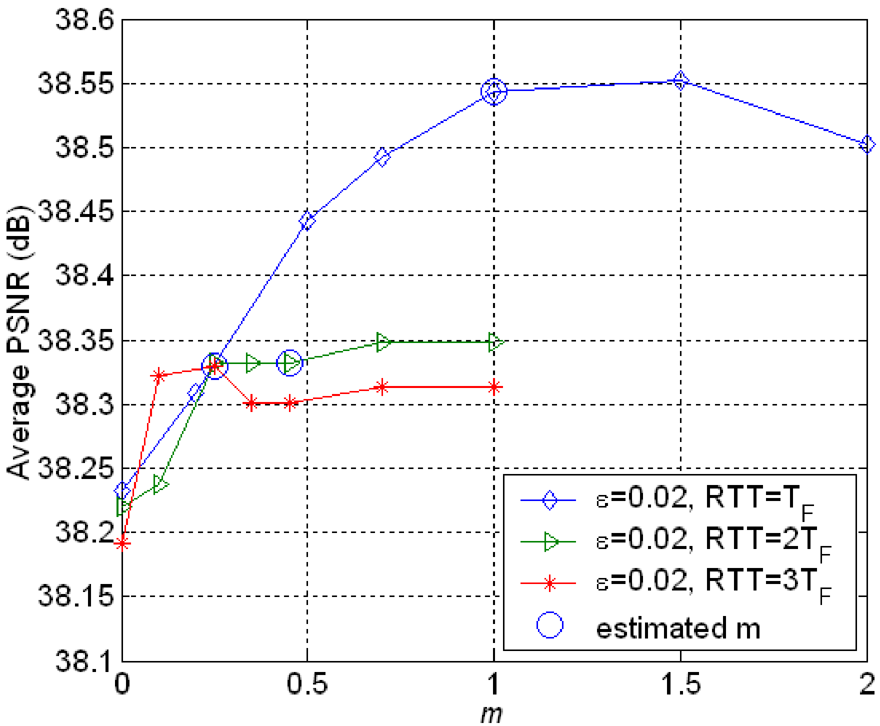
$$\rho_{k,RET}^{(n)} = \varepsilon^m$$

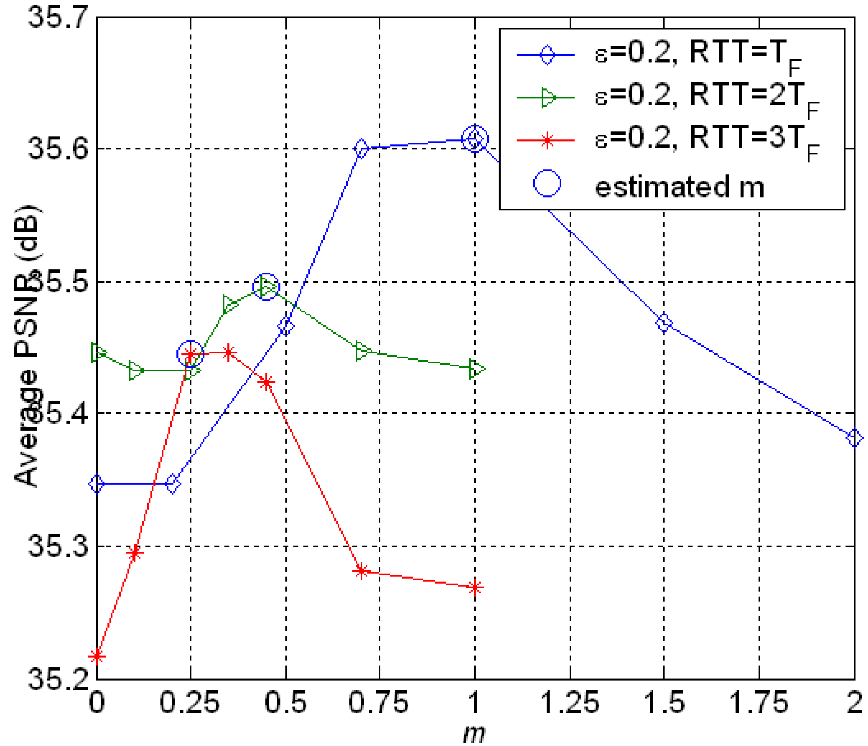
- The estimate of the total number of retransmissions

$$\tilde{m} = \frac{A}{(1+RTT)^2}$$

Estimating $\tilde{m}$

Average PSNR vs. $\tilde{m}$

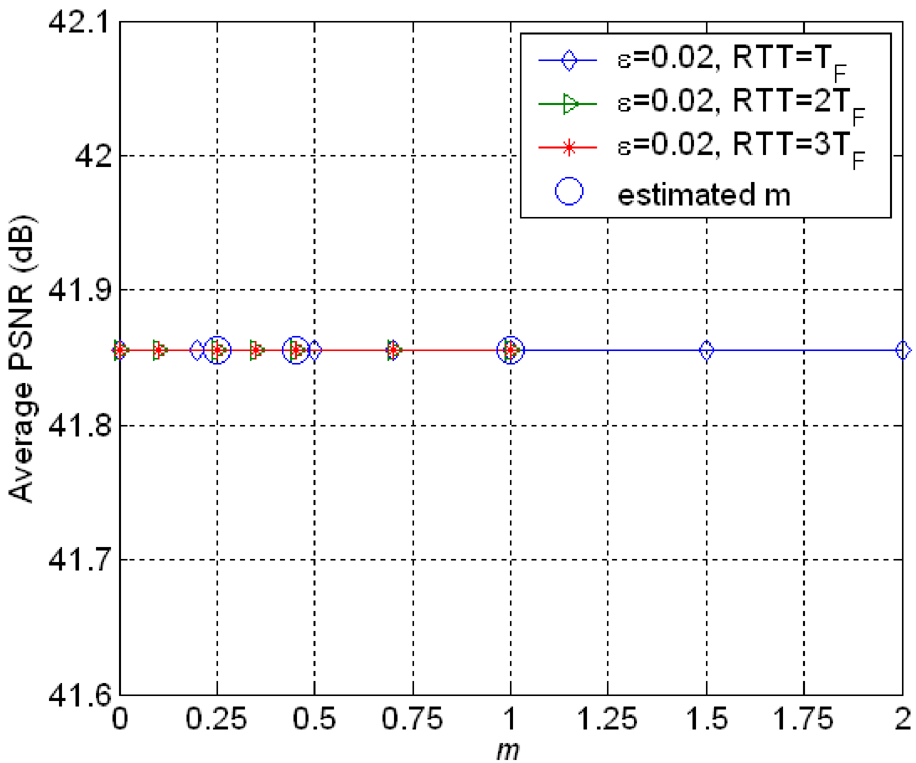


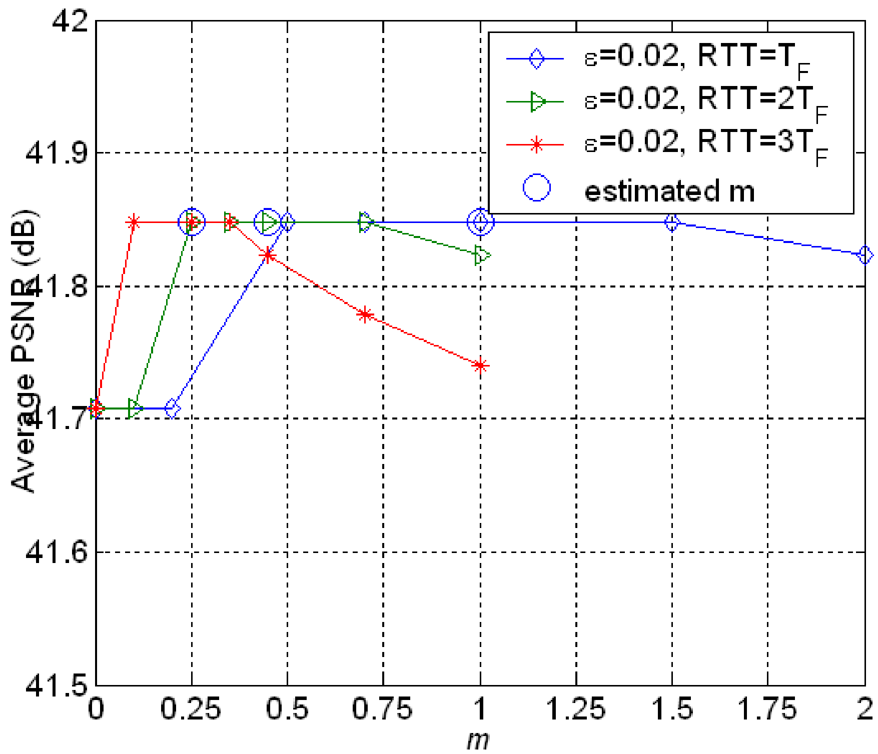


QCIF Foreman sequence at F=15 fps, R_T=480 kbps and A=4

Estimating $\tilde{m}$

Average PSNR vs $\tilde{m}$





QCIF Akiyo sequence at F=15 fps, R_T=360 kbps and A=4

Probability of Packet Loss for the Past Frame

- Probability of packet loss

$$\rho_k^{(n)} = \rho_{k,UPD}^{(n)} \rho_{k,RET}^{(n)}$$

Updated loss probability based on feedback

Due to retransmission

- $\rho^{(n)}_{k,UPD}$ can be accurately calculated
- $\rho^{(n)}_{k,RET}$ can be calculated using the similar way as before

Probability of Packet Loss for the Past Frame

- Assume protected by RS(N, M),
  - L: # packets lost
  - V: # retransmitted packets
  - $J=L+M-N$
- If $V<J$

$$\rho_{k,RET}^{(n-i)} = \varepsilon^{\sigma_k^{(n-i)}}$$

- If $V=J$

$$\rho_{k,RET}^{(n-i)} = \begin{cases} \varepsilon & \text{if } \sigma_k^{(n-i)} = 0 \\ 1-(1-\varepsilon)^J & \text{if } \sigma_k^{(n-i)} = 1 \end{cases}$$

- If $V>J$

$$\rho_{k,RET}^{(n-i)} = \begin{cases} \sum_{j=V-J+1}^{V} \frac{j}{V} \binom{V}{j} \varepsilon^j (1-\varepsilon)^{V-j} & \text{if } \sigma_k^{(n-i)} = 0 \\ \sum_{j=V-J+1}^{V} \binom{V}{j} \varepsilon^j (1-\varepsilon)^{V-j} & \text{if } \sigma_k^{(n-i)} = 1 \end{cases}$$

Solution Algorithm

- Lagrangian Relaxation

$$\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{\sigma}\}} \sum_{i=0}^{A} J^{(n-i)} = \sum_{i=1}^{A} E[D^{(n-i)}(\boldsymbol{\sigma}^{(n-i)})] + \sum_{k=1}^{M} E[D_k^{(n)}(\boldsymbol{\mu},\gamma)]$$

$$+\lambda\left\{\sum_{i=1}^{A}\sum_{k=1}^{M}\sigma_k^{(n-i)}T_k^{(n-i)} + \sum_{k=1}^{M} T_k^{(n)}\right\}$$

- Minimization of the Lagrangian

$$\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{\sigma}\}} \sum_{i=0}^{A} J^{(n-i)} = \min_{\{\boldsymbol{\sigma}\}} \sum_{i=1}^{A} J^{(n-i)}(\boldsymbol{\sigma}^{(n-i)})] + \min_{\{\gamma\}}\left\{\min_{\{\boldsymbol{\mu}\}} \sum_{k=1}^{M} J_k^{(n)}(\boldsymbol{\mu},\gamma)\right\}$$

Retransmission Exhaustive search ($\min_{\{\boldsymbol{\sigma}\}}$)

FEC Exhaustive search ($\min_{\{\gamma\}}$)

Source coding DP ($\min_{\{\boldsymbol{\mu}\}}$)

Simulation – Hybrid FEC/Retransmission

- NFNR– Neither FEC Nor Retransmission

$$\min_{\{\boldsymbol{\mu}\}} E[D(\boldsymbol{\mu}, \gamma_0, \boldsymbol{\sigma}_0)]$$

- Pure Retransmission

$$\min_{\{\boldsymbol{\mu}, \boldsymbol{\sigma}\}} E[D(\boldsymbol{\mu}, \gamma_0, \boldsymbol{\sigma})]$$

- Pure FEC

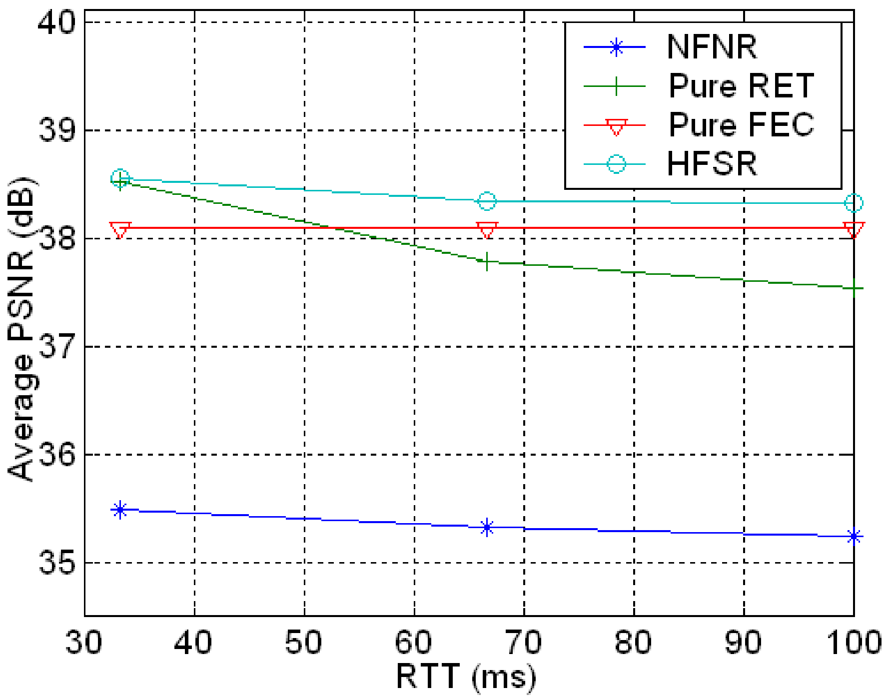
$$\min_{\{\boldsymbol{\mu}, \gamma\}} E[D(\boldsymbol{\mu}, \gamma, \boldsymbol{\sigma}_0)]$$

- HFSR (Hybrid FEC and Selective Retransmission)

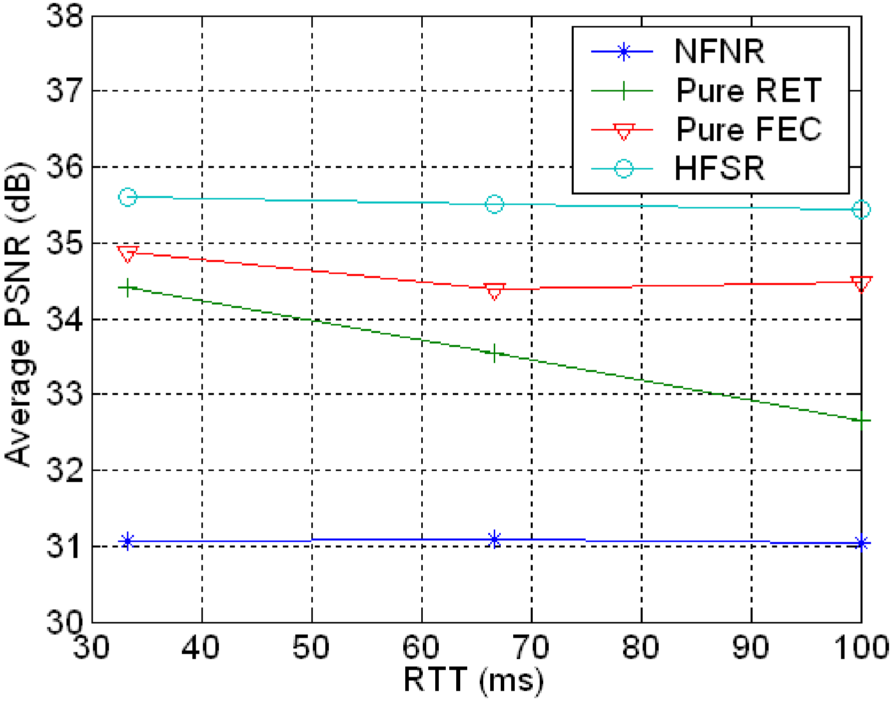
$$\min_{\{\boldsymbol{\mu}, \gamma, \boldsymbol{\sigma}\}} E[D(\boldsymbol{\mu}, \gamma, \boldsymbol{\sigma})]$$

Sensitivity to RTT

Average PSNR vs. RTT



ε=0.02

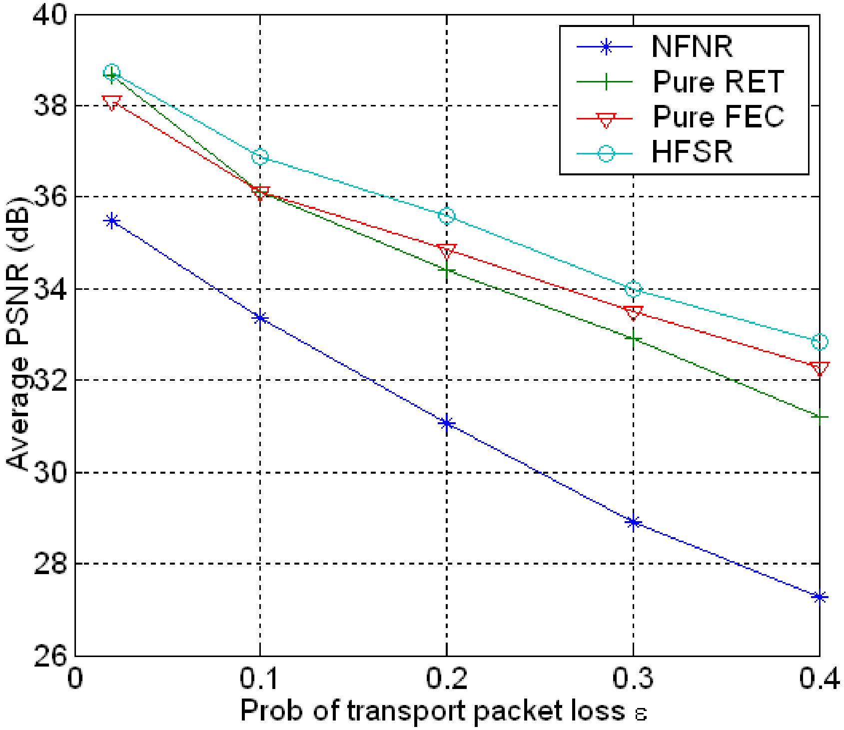


ε=0.2

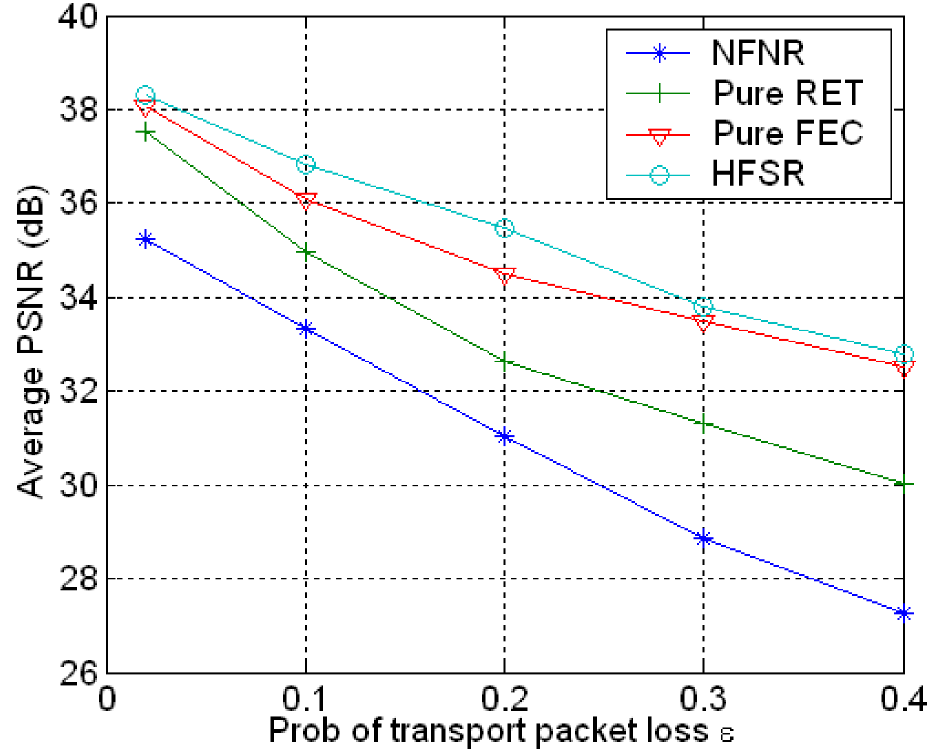
R_T=480 kbps, F=15 fps , QCIF Foreman

Sensitivity to Packet Loss Rate

Average PSNR vs. probability of transport packet loss ε



RTT=T_F

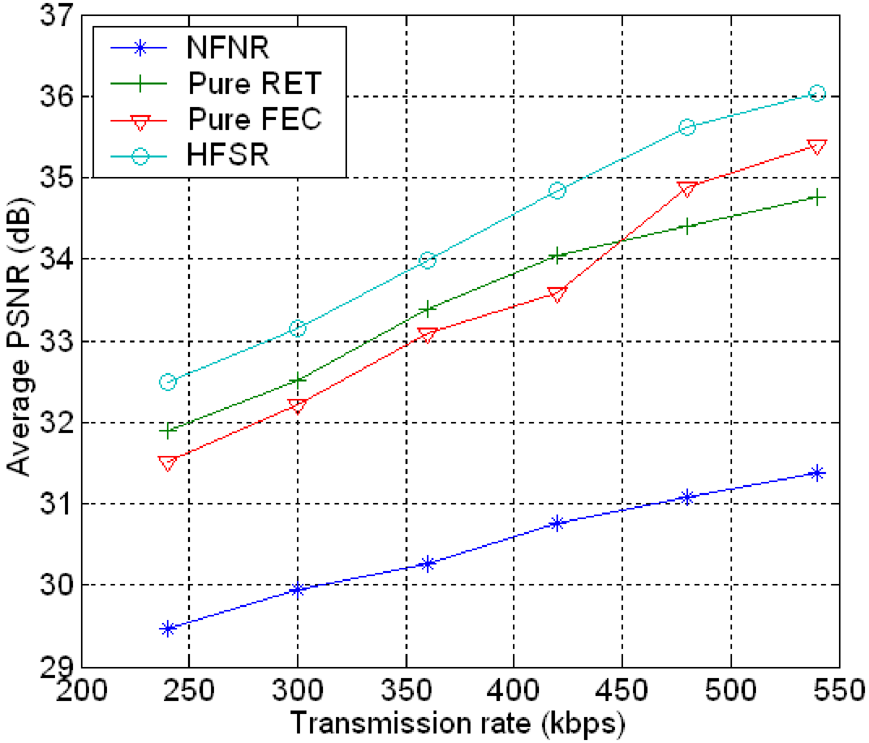


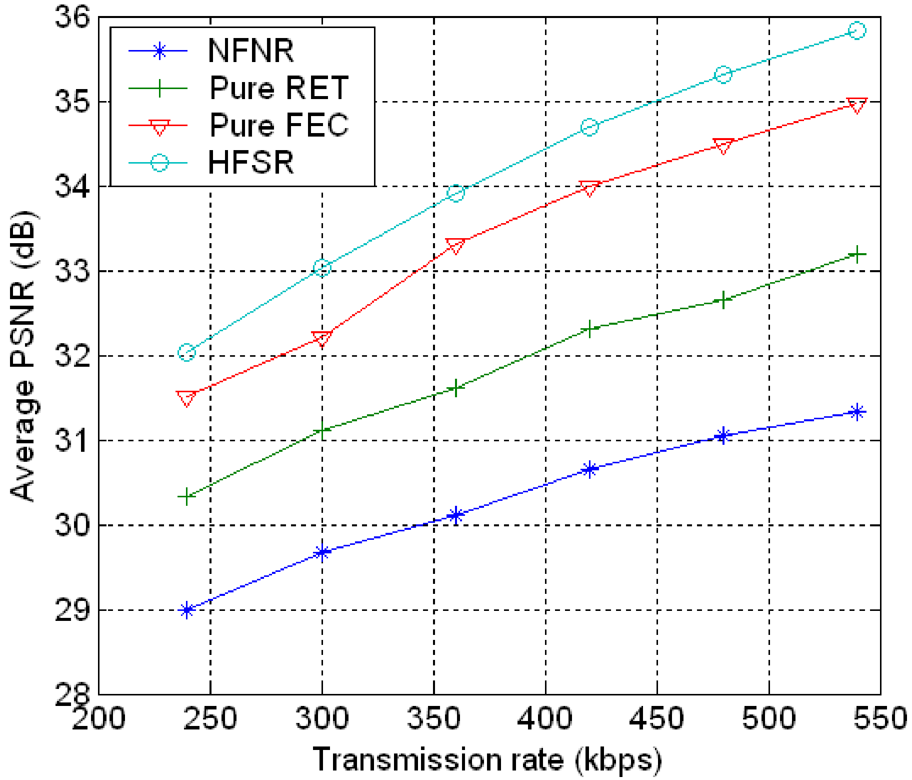
RTT=3T_F

R_T=480 kbps, F=15 fps , QCIF Foreman

Sensitivity to Transmission Rate

Average PSNR vs. channel transmission rate R_T





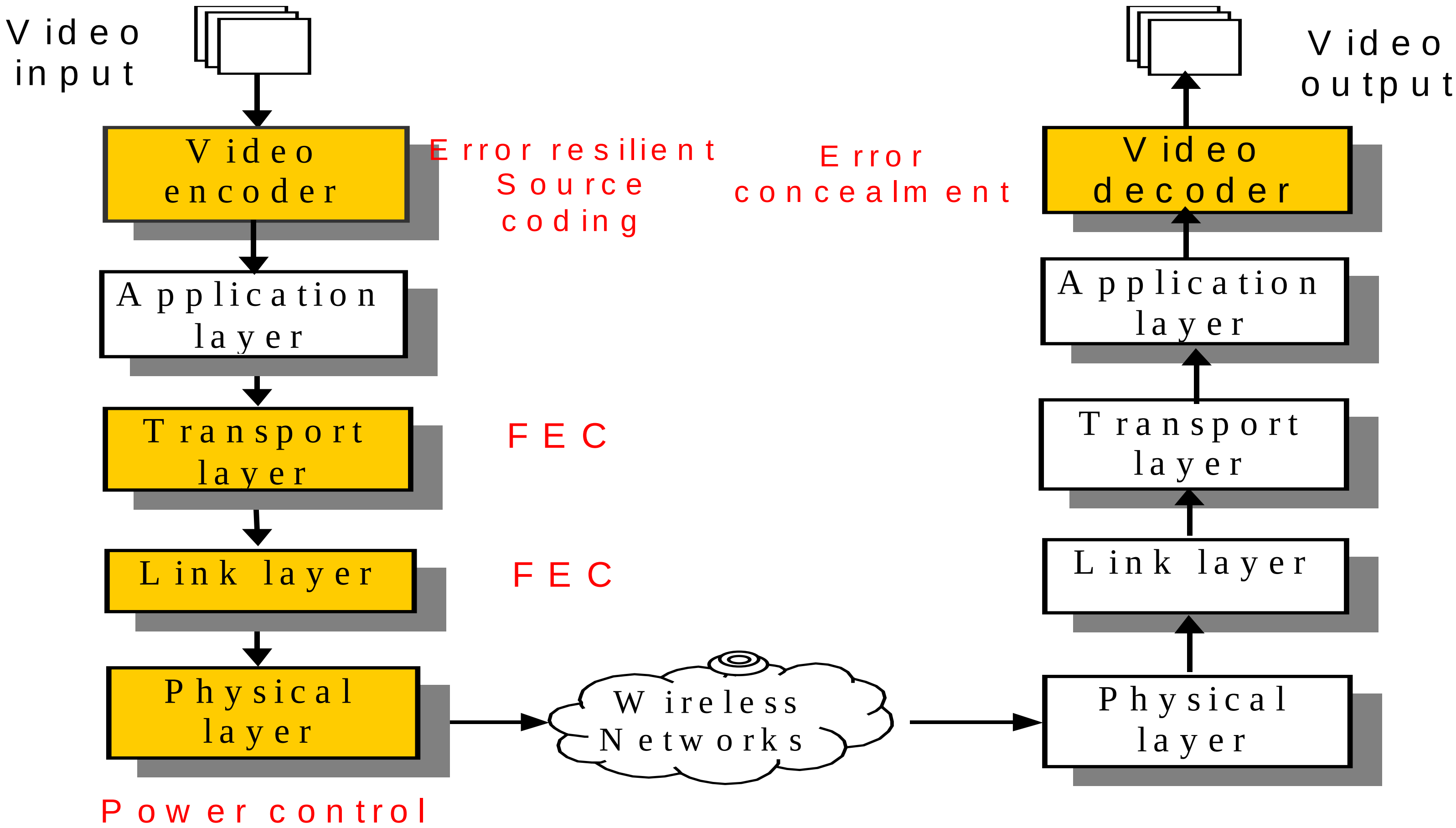
RTT=T_F RTT=$3T_F$

ε=0.2, F=15 fps , QCIF Foreman

JSCC– Conclusions

- Retransmission is suitable for short network RTT, low probability of packet loss, and low transmission rate.
- FEC is more suitable otherwise.
- The proposed hybrid FEC/selective retransmission scheme outperforms both (average gain: 0.7 dB in PSNR).
- In our simulations, we assume the CSI is accurately estimated, which favors FEC, because ARQ does not require accurate CSI.

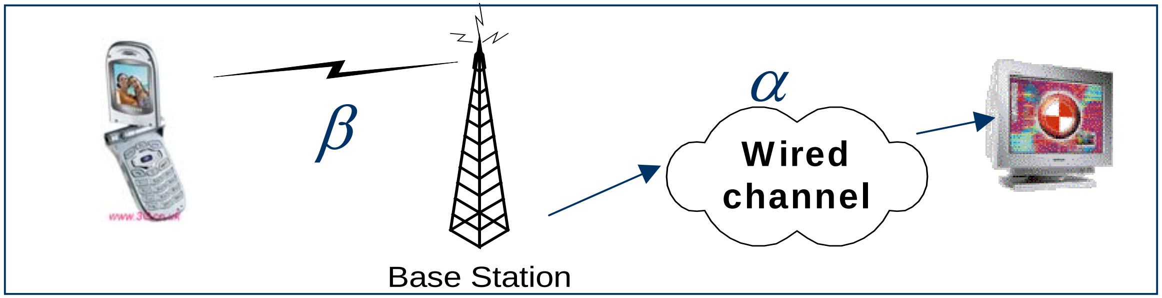
Wireless Video Transmission



➢ JSCCPA– Jointly Source-Channel Coding and Power Allocation

F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, A. K. Katsaggelos, "Joint source-channel coding and power allocation for energy efficient wireless video communications," *Allerton Conf. Communication, Control, and Computing,* Oct. 2003.

Hybrid Wireless Network



- At the IP level, the hybrid network can be modeled as the combination of two independent packet erasure channels: the wired part with loss rate α and the wireless part with loss rate β
- Overall probability of loss for a transport packet

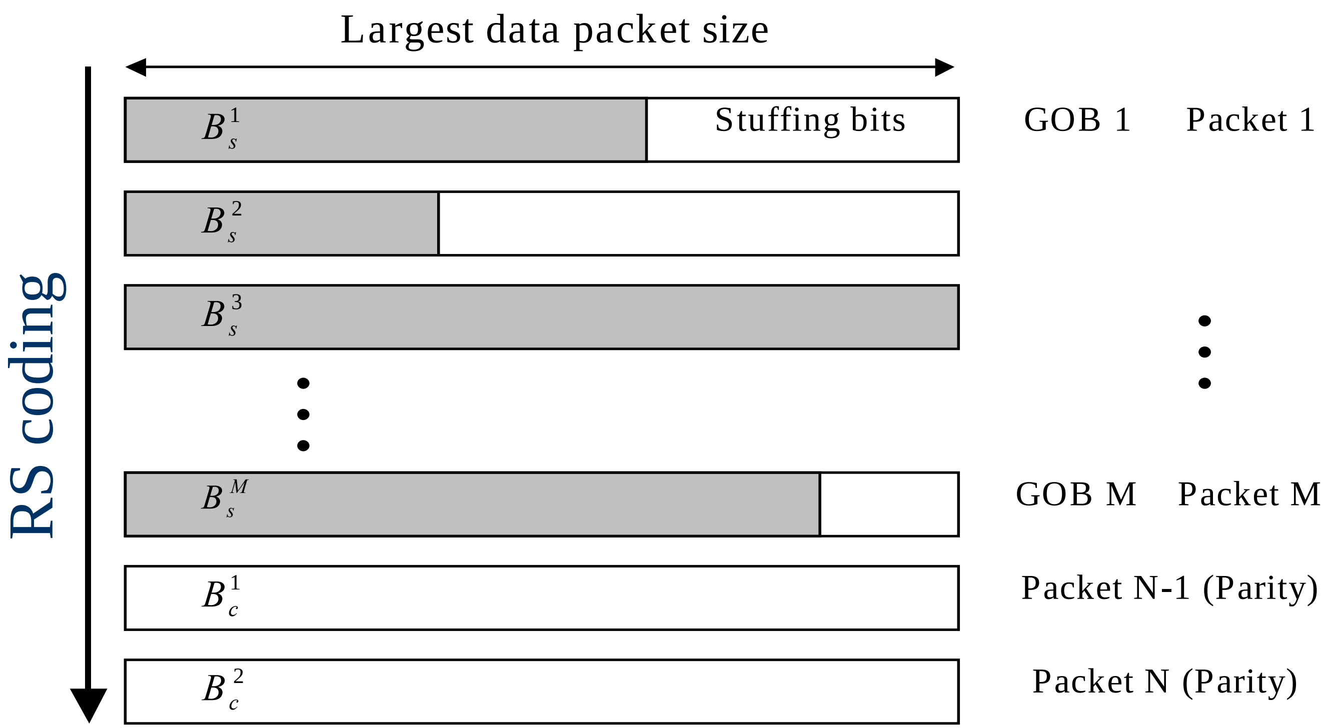
$$\varepsilon_k = \alpha + (1-\alpha)\beta_k$$

- Probability of loss for a transport packet in the wireless channel

$$\beta_k(\mu_k, \nu_k, \eta_k) = 1-(1-p_b)^{B_k}$$

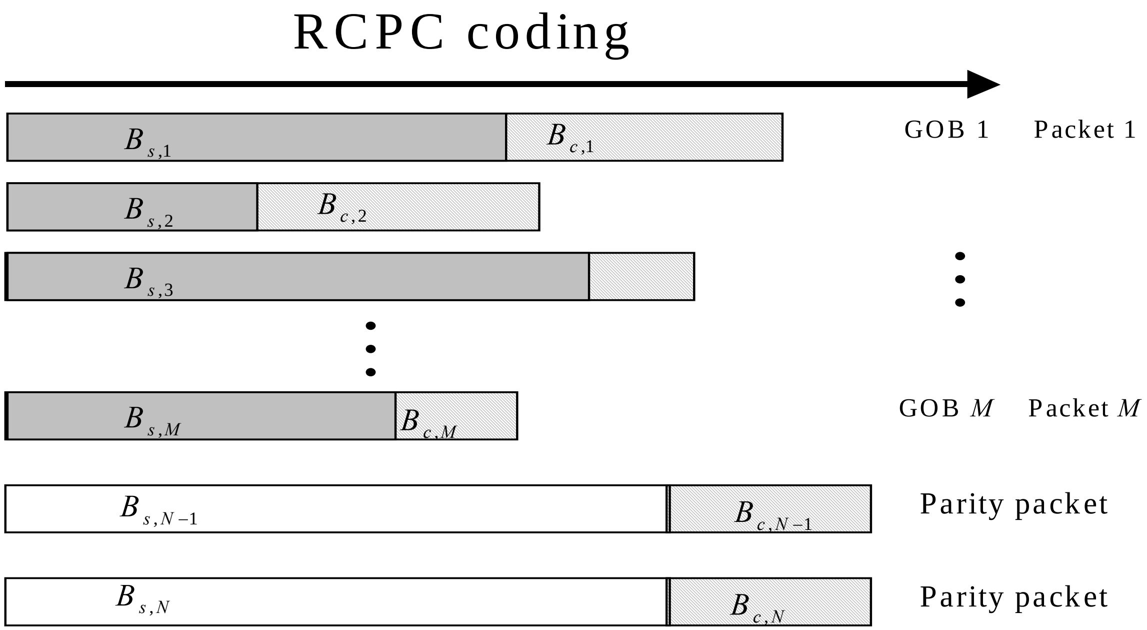
- We need Product FEC to combat different types of channel error

Transport-Layer Packetization



- One row corresponds to one source packet
- Inter-packet FEC
- RS coding RS(N,M)—transport-layer FEC
- Transport-layer FEC parameter: γ

Link-Layer Packetization



- One row corresponds to one source/transport packet
- Intra-packet FEC
- RCPC coding—link-layer FEC
- Link-layer FEC parameter vector $\boldsymbol{\nu} = \{\nu_1, \ldots, \nu_N\}$

Problem Formulation – Hybrid Network

$$\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}\}} E[D(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta})]$$

s.t.

$$C=\sum_{k=1}^{N(\gamma)} B_k P_k / R_T \le C_0$$

$$T=\sum_{k=1}^{N(\gamma)} B_k / R_T \le T_0$$

- $\mu=\{\mu_1,\ldots,\mu_M\}$: Source coding parameter vector
- $\gamma\in\{(N_1,M),\ldots,(N_q,M)\}$: Transport-layer FEC parameter
- $\nu=\{\nu_1,\ldots,\nu_N\}$: Link-layer FEC parameter vector
- $\eta=\{\eta_1,\ldots,\eta_N\}$: Power level parameter vector

Solution Algorithm

- Lagrangian relaxation

$$L(\boldsymbol{\mu}, \gamma, \boldsymbol{v}, \boldsymbol{\eta}, \lambda_1, \lambda_2) = E[D] + \lambda_1 (C - C_0) + \lambda_2 (T - T_0)$$

  - Have proposed an iterative search algorithm to solve the above problem (Allerton'03)

- Minimization of the Lagrangian

$$g(\lambda_1, \lambda_2) = \min_{\{\boldsymbol{\mu}, \gamma, \boldsymbol{v}, \boldsymbol{\eta}\}} L(\boldsymbol{\mu}, \gamma, \boldsymbol{v}, \boldsymbol{\eta}, \lambda_1, \lambda_2)$$

Solution Algorithm

- Lagrangian relaxation

$$L(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta},\lambda_1,\lambda_2)=E[D]+\lambda_1(C-C_0)+\lambda_2(T-T_0)$$

- Minimization of Lagrangian

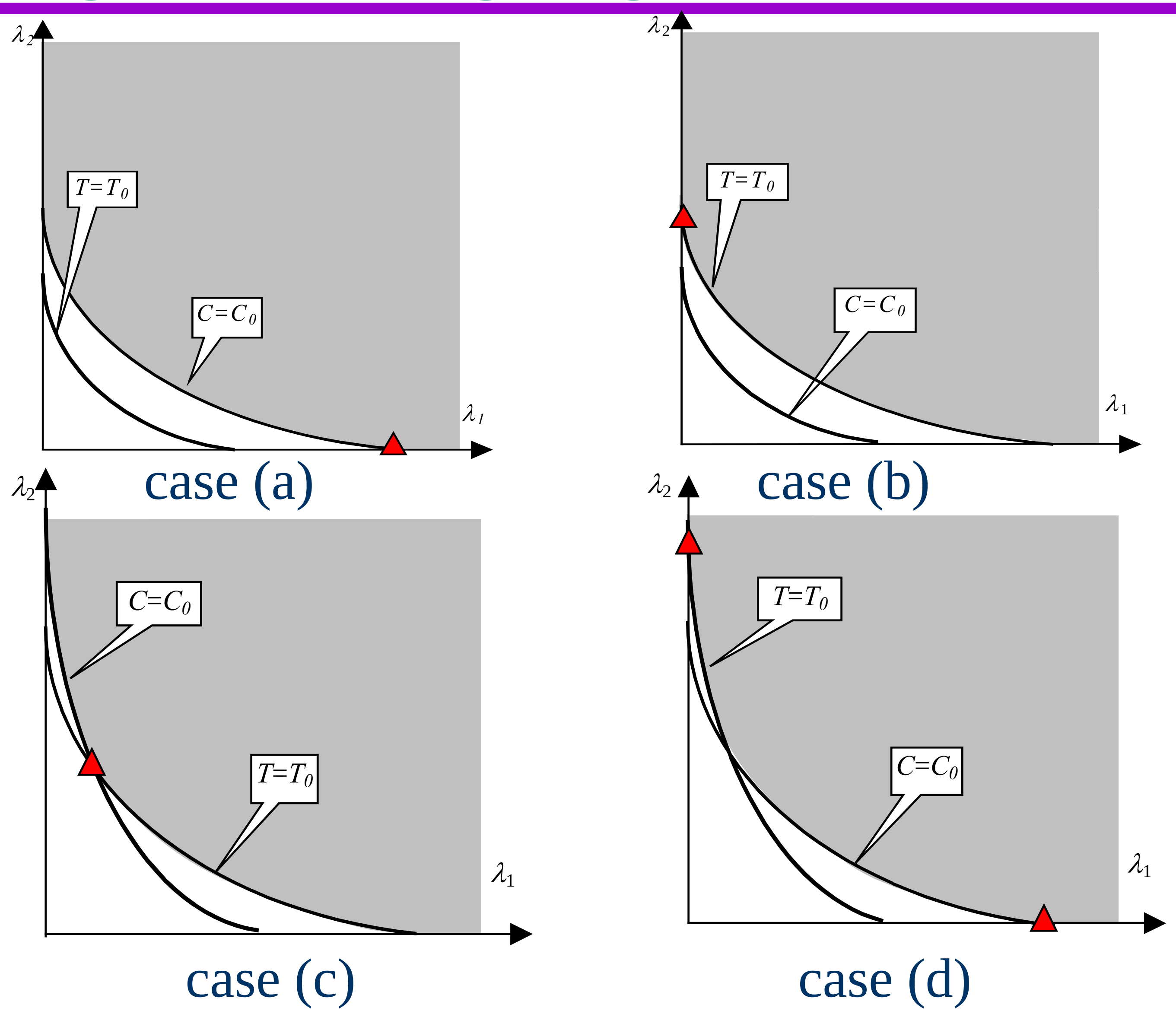
$$g(\lambda_1,\lambda_2)=\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}\}} L(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta},\lambda_1,\lambda_2)$$

- Dual Problem

$$\max_{\{\lambda_1\geq 0,\lambda_2\geq 0\}} g(\lambda_1,\lambda_2)$$

- Goal: find the convex-hull solution
  - The duel problem is concave
  - complementary slackness applies

Algorithm– Lagrangian relaxation



Algorithm– Lagrangian Relaxation

Step 1 (*case a, d*): Let λ_2 =0, find λ_1^* to satisfy $C(H(\lambda_1^*,0)) \le C_0$.

If $T(H(\lambda_1^*,0)) \le T_0$, $H(\lambda_1^*,0)$ corresponds to the optimal solution. Otherwise,

Step 2: (*case b, d*): Let λ_1 =0, find λ_2^* to satisfy $T(H(0,\lambda_2^*)) \le T_0$.

If $C(H(\lambda_1^*,0)) \le C_0$, $H(0,\lambda_1^*)$ corresponds to the optimal solution. Otherwise,

Step 3 (*case c*):

- **i.** Let $\lambda_1^l = 0, \lambda_1^r = \lambda_1^*$, $\lambda_2^b = 0, \lambda_2^t = \lambda_2^*$.
- **ii.** $\lambda_1^m = (\lambda_1^l + \lambda_1^r)/2$, find λ_2^* within [λ_2^b, λ_2^t] to satisfy $T(H(\lambda_1^m, \lambda_2^*)) \le T_0$.
- **iii.** If $C(H(\lambda_1^m, \lambda_2^*)) > C_0$, then let $\lambda_1^l = \lambda_1^m$, $\lambda_2^t = \lambda_2^*$, and go to step 3ii. Otherwise,
- **iv.** If $C(H(\lambda_1^m, \lambda_2^*)) < C_0 - \varepsilon$ (ε is a relatively small number), then let $\lambda_1^r = \lambda_1^m$, $\lambda_2^b = \lambda_2^*$, and go to step 3ii. Otherwise,
- **v.** The optimal solution corresponds to $H(\lambda_1^m, \lambda_2^*)$.

Calculation of Probability of Packet Loss

$$\rho_k(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}) = \sum_{j=N(\gamma)-M+1}^{N(\gamma)} P_b(N,j) = \sum_{j=N-M+1}^{N} \sum_{Q_j^t \in I_j(N,k)} \left(\prod_{i \in Q_j^t} \varepsilon_i \prod_{l \in \overline{Q}_j^t} (1-\varepsilon_l) \right)$$

$$Q_j^t(N),\ j = 1,\ldots,N,\ t = 1,\ldots,\binom{N}{j}$$

the t-th subset with j elements of $Q(N)$

$$I_j(N,k) = \{Q_j^t \in Q(N) \mid k \in Q_j^t(N), |Q_j^t| = j\}$$

- Loss probability of a source packet depends on the parameters chosen for all the other packets in the frame
    - ε_k differs from packet to packet
    - Inter-packet dependency introduced by inter-packet FEC

Calculation of Probability of Packet Loss

$$\rho_k(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}) = \sum_{j=N(\gamma)-M+1}^{N(\gamma)} P_b(N,j)$$

$$= \sum_{j=N-M+1}^{N} \sum_{Q_j^t \in I_j(N,k)} \left(\prod_{i \in Q_j^t} \varepsilon_i \prod_{l \in \overline{Q}_j^t} (1-\varepsilon_l) \right)$$

$P_b(N,j)$ Probability that the k-th packet is correctly decoded by the RCPC decoder and the total number transport packets that are not correctly received from the group of N packets is j

$$Q(3)=\{1,2,3\},\ \ Q_1^1(3)=\{1\}, Q_1^2(3)=\{2\}, Q_1^3(3)=\{3\}$$

$$Q_2^1(3)=\{1,2\}, Q_2^2(3)=\{1,3\}, Q_2^3(3)=\{2,3\}, Q_3^1(3)=\{1,2,3\},$$

Minimization of the Lagrangian– Hybrid

- Iterative descent algorithm (alternating variables)

Step 1: Set $\boldsymbol{x}^{(0)}$ corresponds to any initial state

Step 2: let t_n=(t mod N) (round-robin style)

if $i \neq t_n$, let $x_i^{(n)} = x_i^{(n-1)}$; otherwise, for $i=t_n$

$$x_i^{(t)} = \arg\min_{x^{(t)}} L(x_1^{(t)},\ldots,x_{i-1}^{(t)},x_i,x_{i+1}^{(t)},\ldots,x_N^{(t)})$$

Step 3: If not convergence, go back to Step 2

$x = \{x_1, x_2, \ldots, x_N\}$ Parameter selected for the N packets

$x^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \ldots, x_N^{(t)}\}$ for $t = 0,1,\ldots$ at step t

Experimental Setup

- **"Real Time" Applications**
  - Short-delay applications, where ARQ is not applicable
  - Delay constraint equals one frame time
- **Channel Model**
  - Flat Rayleigh fading channel with AWGN
  - i.i.d. channel fading
  - RTT = $2T_F$, which preclude retransmission

Experimental Setup

Performance of RCPC (in BER) over a
Rayleigh fading channel with interleaving

SNR(dB)	2	6	10	14	18
Cr = 1/2	$1.4*10^{-3}$	$2.2*10^{-5}$	$2.1*10^{-6}$	$2.4*10^{-7}$	$6.4*10^{-8}$
Cr = 4/7	$1.1*10^{-1}$	$5.3*10^{-4}$	$4.1*10^{-5}$	$1.2*10^{-5}$	$3.8*10^{-6}$
Cr = 2/3	$3.2*10^{-1}$	$7.4*10^{-3}$	$1.7*10^{-4}$	$3.5*10^{-5}$	$1.2*10^{-5}$
Cr = 4/5	$4.2*10^{-1}$	$4.0*10^{-2}$	$6.6*10^{-4}$	$1.1*10^{-4}$	$3.6*10^{-5}$

- Generator polynomial (133, 171), mother code rate ½, and puncturing rate 4
- Soft Viterbi decoding in conjunction with BPSK

PFEC vs. Link-Layer FEC

- Approach 1 – Link Layer FEC (LFEC)

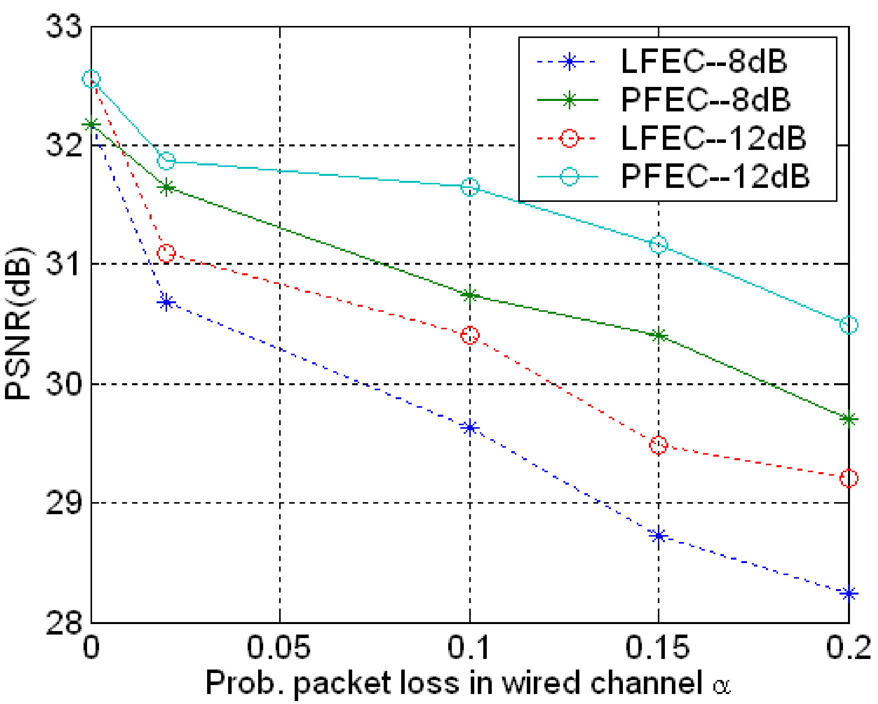
$$\min_{\{\boldsymbol{\mu},\boldsymbol{v}\}} E[D(\boldsymbol{\mu},\gamma_0,\boldsymbol{v},\boldsymbol{\eta}_0)]$$

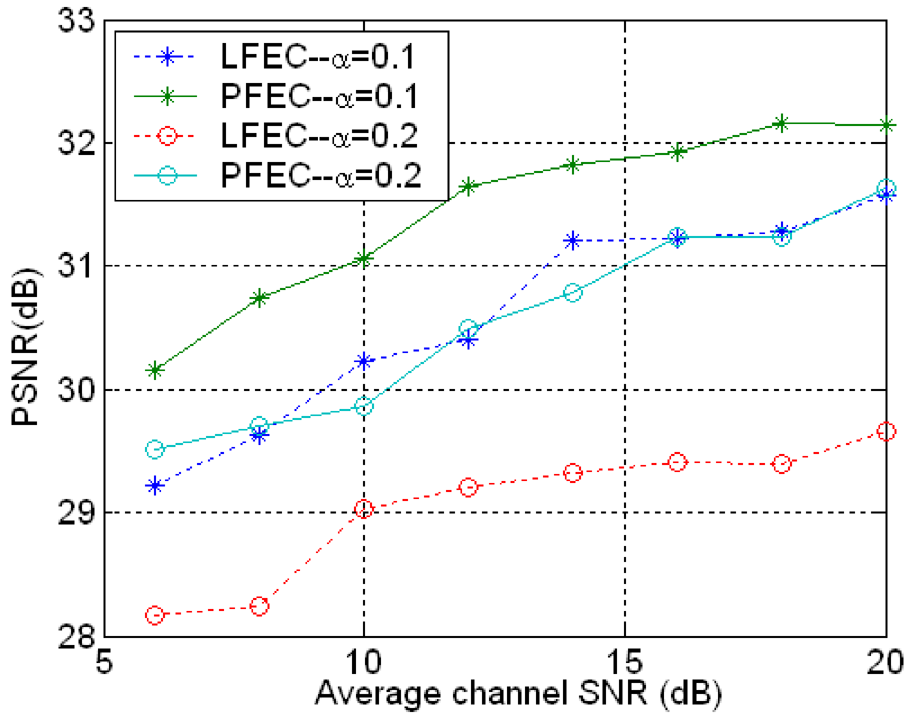
- Approach 2 – Product FEC (PFEC)

$$\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{v}\}} E[D(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}_0)]$$

F. Zhai, Y. Eisenberg, T.N. Pappas, R. Berry, A..K. Katsaggelos, "Rate-distortion optimized product code forward error correction for video transmission over IP-based wireless networks," *IEEE ICASSP'04*, accepted.

PFEC vs. LFEC





Average PSNR vs. α

Average PSNR vs. channel SNR

QCIF Foreman, 30 fps

UEP vs. EEP

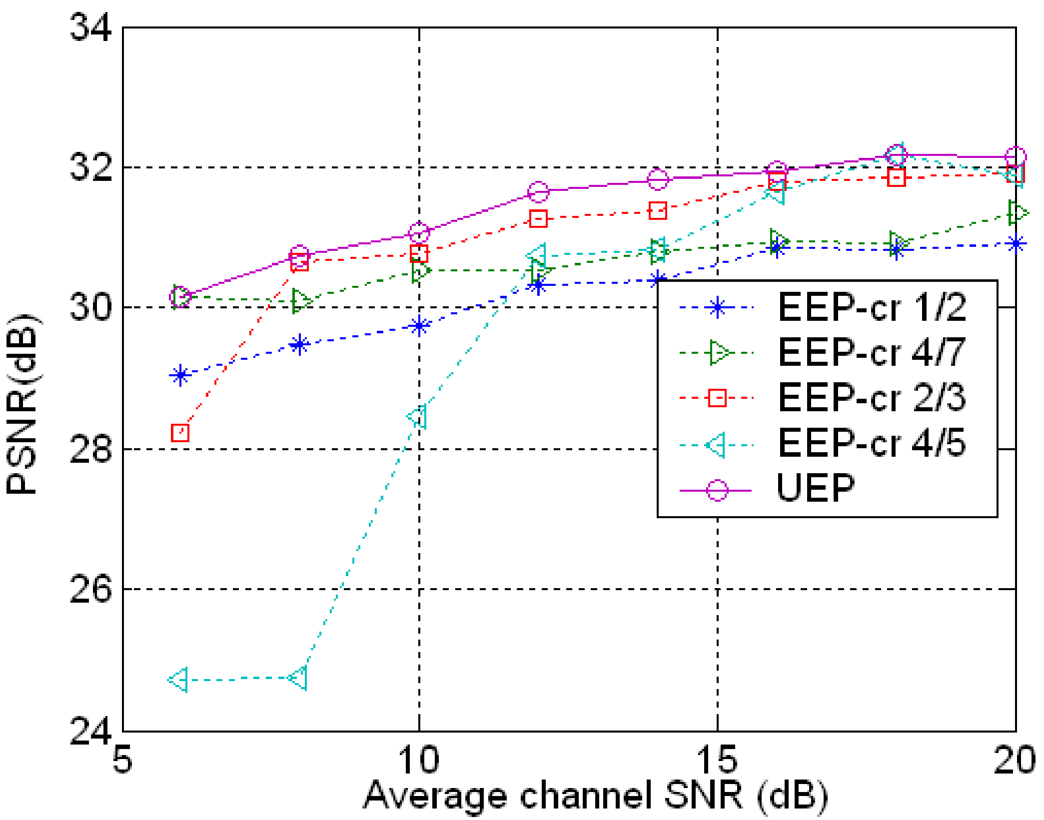
- Approach 1 – EEP-PFEC

$$\min_{\{\boldsymbol{\mu},\gamma\}} E[D(\boldsymbol{\mu},\gamma,\boldsymbol{v}_0,\boldsymbol{\eta}_0)]$$

- Approach 2 – UEP-PFEC

$$\min_{\{\boldsymbol{\mu},\gamma,\boldsymbol{v}\}} E[D(\boldsymbol{\mu},\gamma,\boldsymbol{v},\boldsymbol{\eta}_0)]$$

UEP vs. EEP



PSNR vs. channel SNR, α=0.1

QCIF Foreman, 30 fps

Problem Formulations—Wireless Channel

$$\min_{\{\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\eta}\}} E[D(\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\eta})]$$

$$\text{s.t.} \qquad C = \sum_{k=1}^{M} B_k P_k / R_T \le C_0$$

$$T = \sum_{k=1}^{M} B_k / R_T \le T_0$$

Solution Algorithm— Wireless

- Lagrangian relaxation

$$L(\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\eta}) = \sum_{k=1}^{M} J_k = \sum_{k=1}^{M} E[D_k] + \lambda_1 C_k + \lambda_2 T_k$$

- Dynamic Programming

$$J_k = J_k(\mu_{k-1}, v_{k-1}, \eta_{k-1}, \mu_k, v_k, \eta_k)$$

Simulation 1– Wireless Channel

- Goal: To show the benefit by adjusting power levels

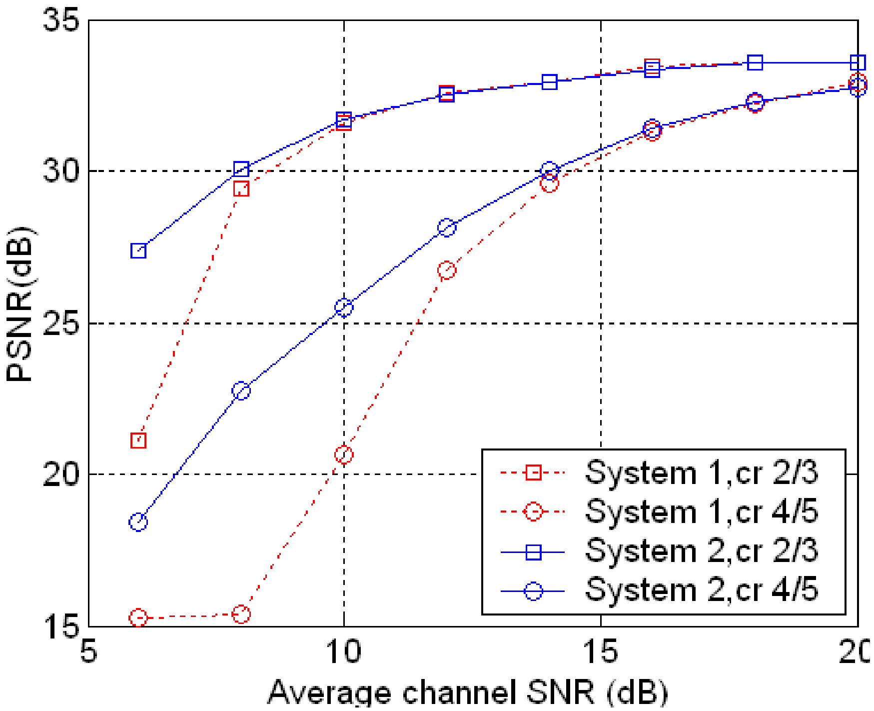
- System 1 -- Optimal error-resilient source coding

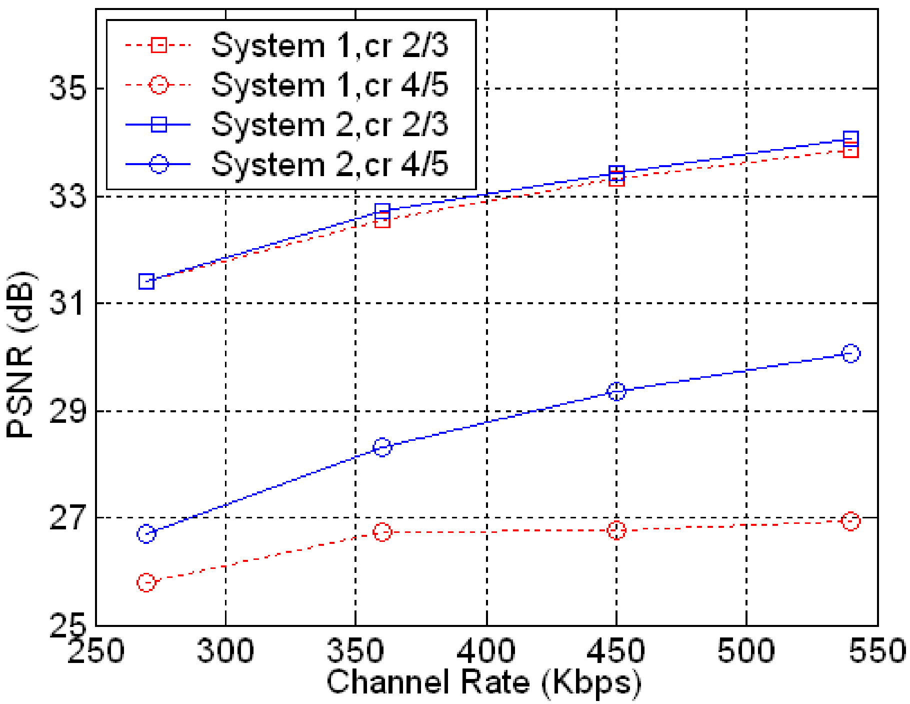
$$\min_{\{\boldsymbol{\mu}\}} E[D(\boldsymbol{\mu}, \boldsymbol{v}_0, \boldsymbol{\eta}_0)]$$

- System 2 -- Joint source coding and power allocation

$$\min_{\{\boldsymbol{\mu}, \boldsymbol{\eta}\}} E[D(\boldsymbol{\mu}, \boldsymbol{v}_0, \boldsymbol{\eta})]$$

Simulation 1





PSNR vs. channel SNR

(Transmission rate = 360 kbps)

PSNR vs. transmission rate

(reference SNR = 12 dB)

QCIF Foreman, 30 fps

Simulation 1

Allocation of power level (1,2,3,4,5) in percentage in JSCPA system

Ref SNR(dB)	6	12	20
Cr = 1/2	(2.4,18.5,73.9,5.1,0)	(12.6,32.4,34,19.6,1.4)	(62.3,0,12.9,0.24.8)
Cr = 4/7	(18,0,14.3,66.1,1.6)	(2.3,29.9,56.4,11,0.3)	(10,35,39.2,13.4,2.3)
Cr = 2/3	(40,0,0,13,47)	(0.7,14,66,18.7,0.6)	(11.6,10.8,69.1,6.9,1.6)
Cr = 4/5	(45.8,0,0,0,54.2)	(2,4,41.8,47.3,5)	(8.2,31.5,43.8,15.3,1.3)

- The reference power level 3
- Each value here denotes the percentage of packets using the corresponding transmission power level.
- The transmission power is proportional to the power level parameter

Simulation 2

- Goal: To show the benefit by adjusting channel rates

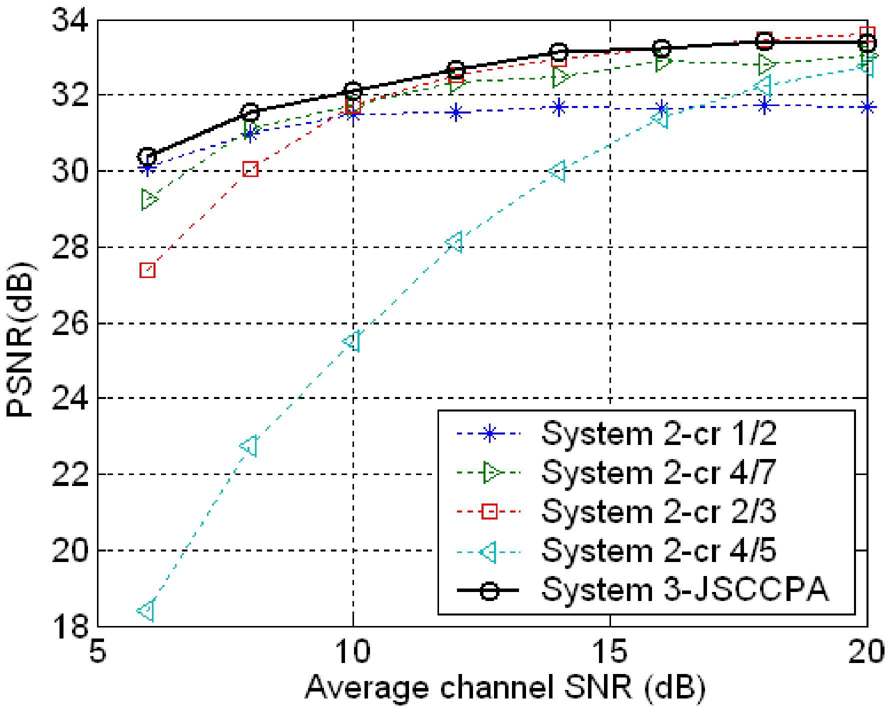
- System 2 -- Joint source coding and power allocation

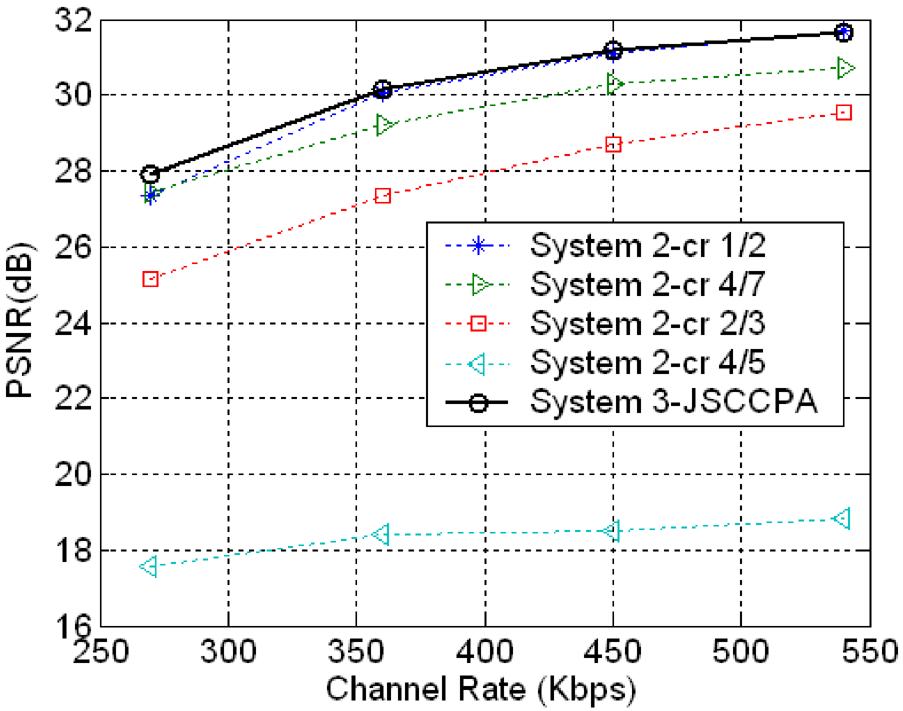
$$\min_{\{\boldsymbol{\mu},\boldsymbol{\eta}\}} E[D(\boldsymbol{\mu},\boldsymbol{v}_0,\boldsymbol{\eta})]$$

- System 3 -- JSCC and power allocation (JSCCPA)

$$\min_{\{\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\eta}\}} E[D(\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\eta})]$$

Simulation 2





PSNR vs. channel SNR

(transmission rate = 360 kbps)

PSNR vs. transmission rate

(reference SNR = 12 dB)

QCIF Foreman, 30 fps

Simulation 2

Channel Coding rates in percentages in JSCCPA system

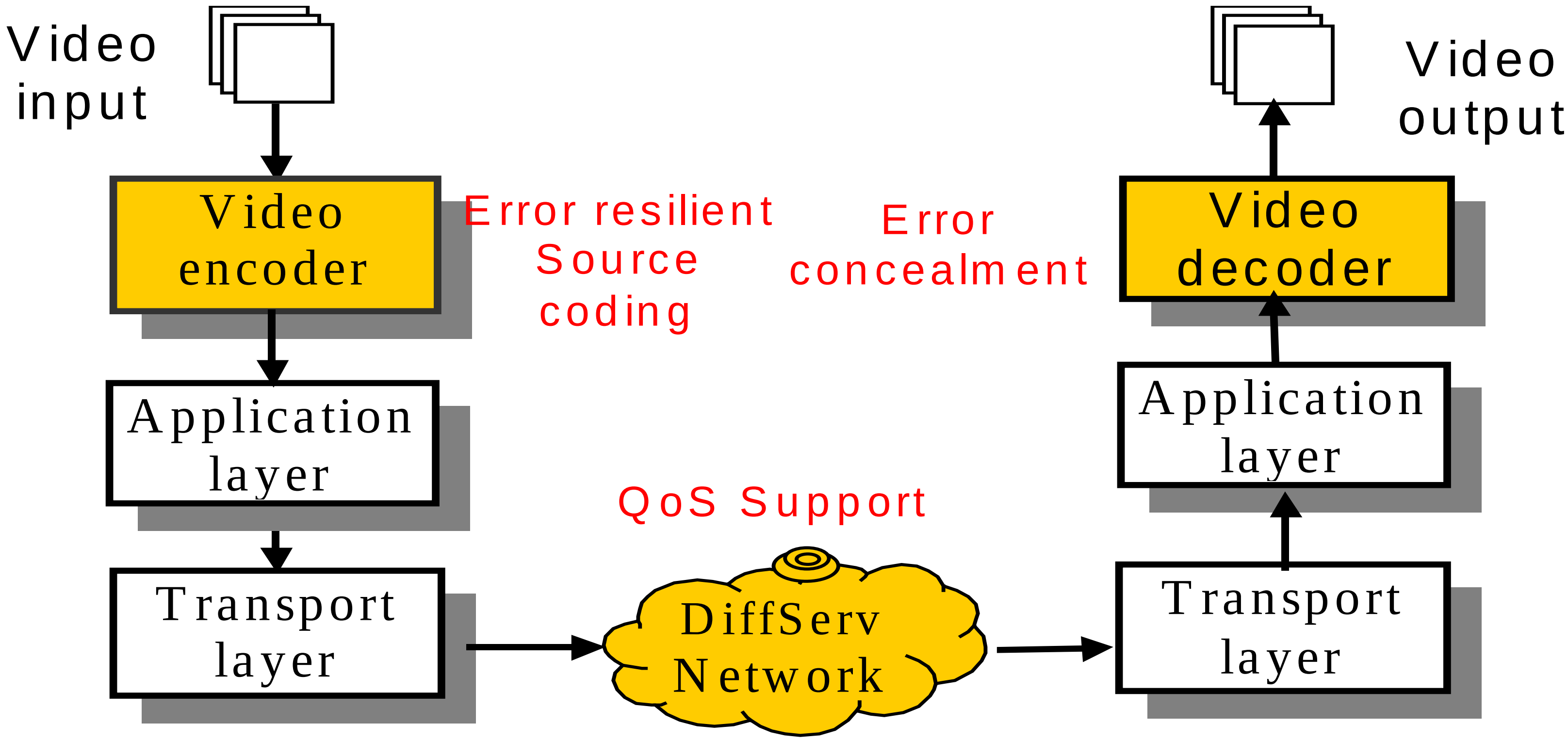
Ref SNR(dB)	6	8	10	12	14	16	18	20
Cr = 1/2	96.2	67.7	41.2	19.6	6.7	4.7	1.0	1.6
Cr = 4/7	3.8	31.9	57.3	69.6	61.3	35.0	17.8	5.6
Cr = 2/3	0	0.4	1.5	10.8	31.3	57.7	73.9	69.6
Cr = 4/5	0	0	0	0	0.7	2.6	7.3	23.2

- The better channel, the less channel coding protection

JSCCPA– Conclusions

- Optimal cross-layer resource allocation to provide UEP for wireless video transmission
  - Error resilient source coding
  - Transport-layer FEC
  - Link-layer FEC
  - Physical-layer power allocation
  - Error concealment
- Transport-layer FEC (inter-packet FEC) is not necessary if the wired link has no error.
- Adapting either power or link-layer FEC (instead both) may be adequate to achieve the near-optimal result in some situations.

DiffServ Video Transmission



- JSCPC– Jointly Source Coding and Packet Classification

F. Zhai, C. E. Luna, Y. Eisenberg, T. N. Pappas, R. Berry, A. K. Katsaggelos, "Joint source coding and packet classification for real-time video transmission over differentiated services networks," *IEEE Trans. Multimedia, to appear*, 2004

Differentiated Services Networks

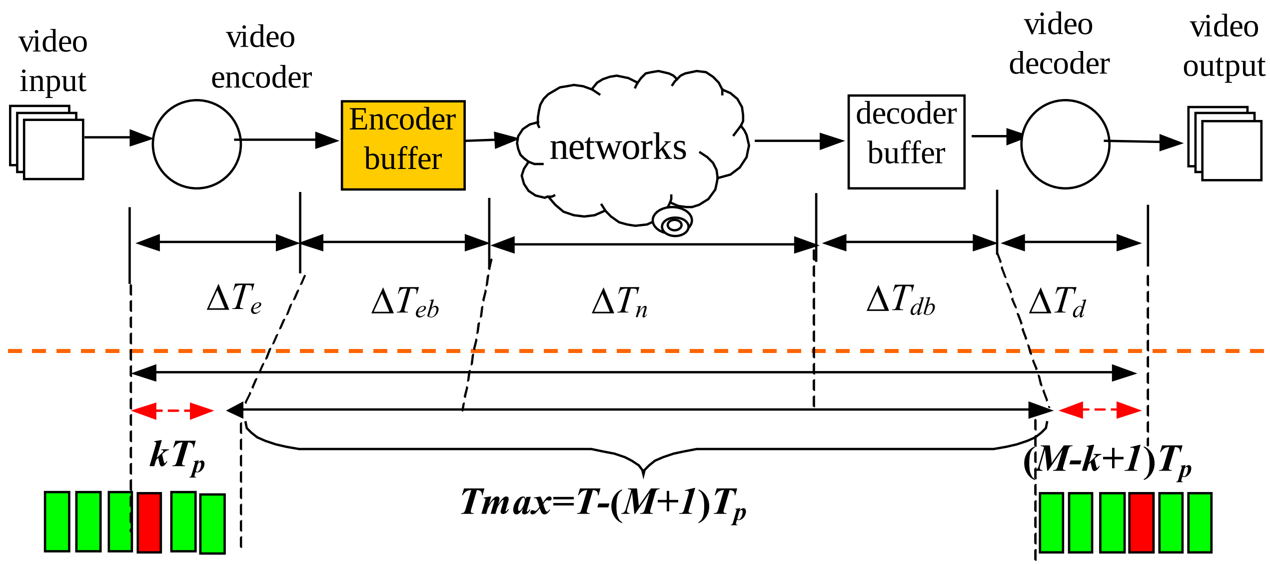
- Allocating resources discriminatorily to aggregated traffic flows
- Multiple service classes: different class has different end-to-end statistical behavior
- Sender is charged for each transmitted bit based on its service class

Delay Components

- Constant end-to-end delay: $T = \Delta T_e + \Delta T_{eb} + \Delta T_n + \Delta T_{db} + \Delta T_d$
- Frame level
- Packet level

  An example

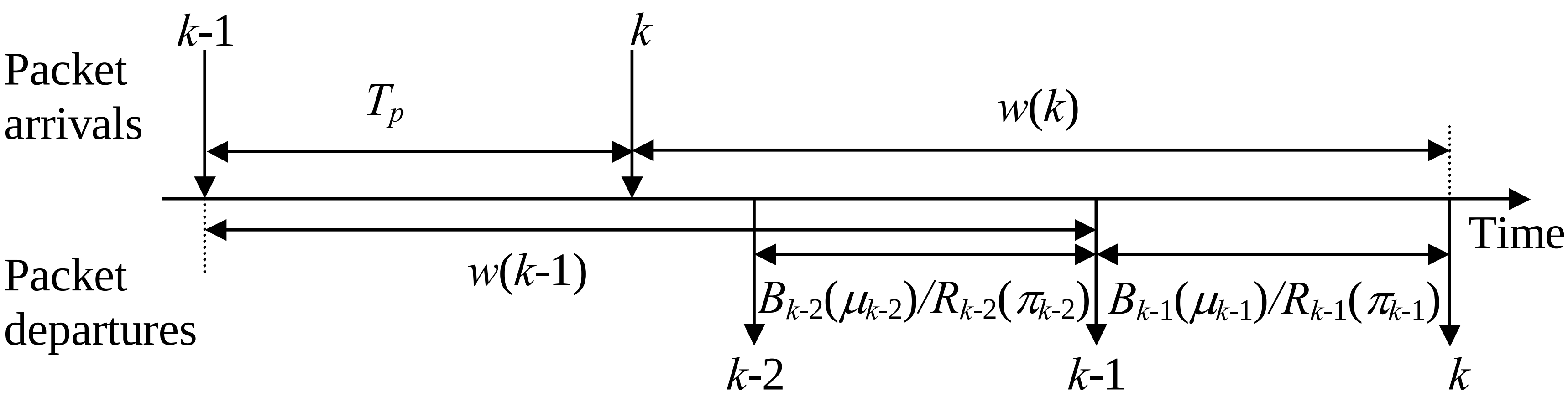
  (M=6, k=3)



$$\Delta T_{eb}(k) + \Delta T_n(k) + \Delta T_{db}(k) = T - (M+1)T_p = T_{\max}$$

to avoid excessive delay : $\Delta T_{db}(k) \geq 0 \Rightarrow \Delta T(k) = \Delta T_{eb}(k) + \Delta T_n(k) \leq T_{\max}$

Delay Components—cont.



encoder buffer delay: $\Delta T_{eb}(k) = w(k) + \dfrac{B_k(\mu_k)}{R_k}$

waiting time: $w(k) = w(k-1) + \dfrac{B_{k-1}(\mu_{k-1})}{R_{k-1}} - T_p$

max allowable network delay: $\tau(k) = T_{\max} - \Delta T_{eb}(k) = T_{\max} - w(k) - \dfrac{B_k(\mu_k)}{R_k}$

JSCPC– Solution Algorithm

- Lagrangian relaxation 1

$$\min_{\{\boldsymbol{\pi},\boldsymbol{\mu}\}} L(\boldsymbol{\pi},\boldsymbol{\mu},\lambda_1,\lambda_2) = \sum_{k=1}^{M} \{D_k + \lambda_1 C_k + \lambda_2 T_k\}$$

- Lagrangian relaxation 2

$$\min_{\{\mu,\pi\}} L(\mu,\pi,\lambda) = \min_{\{\mu,\pi\}} \sum_{k=1}^{M} D_k + \lambda C_k$$

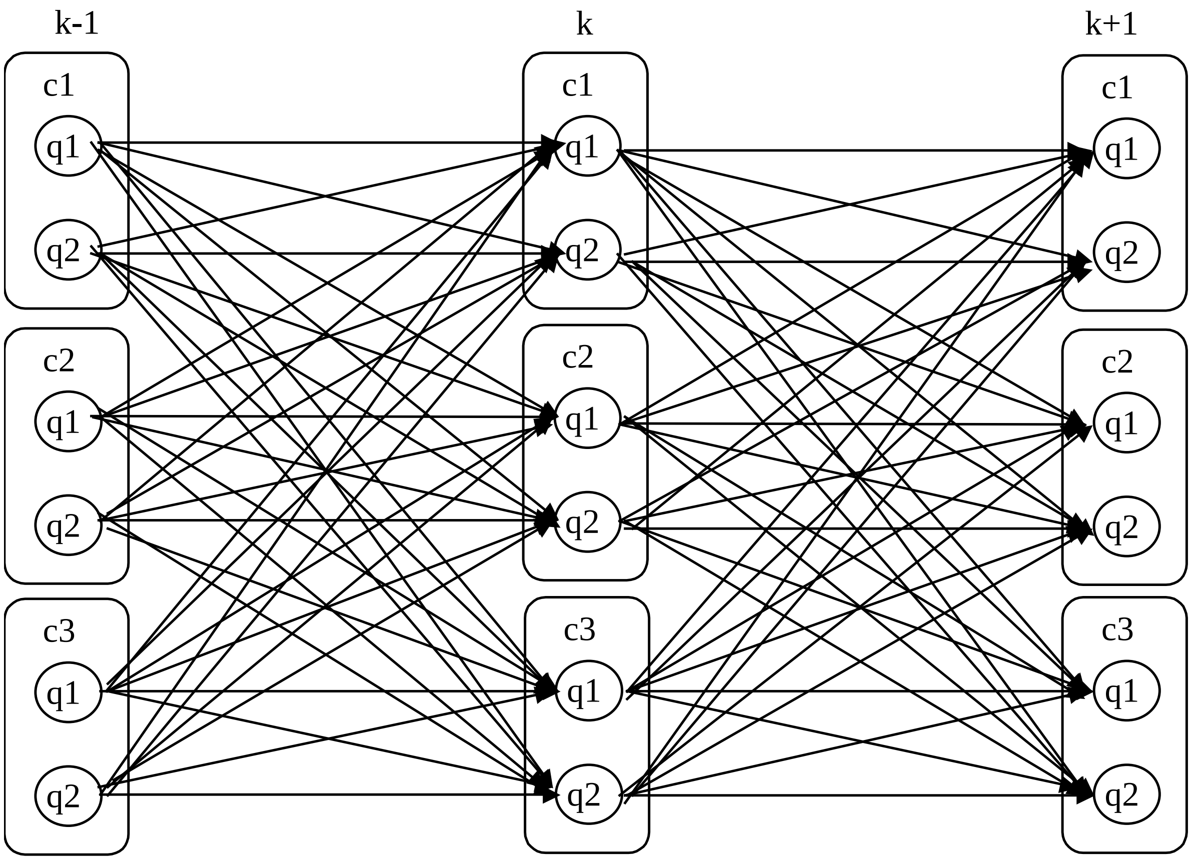
$$s.t. \quad \sum_{k=1}^{M} T_k \le T_0$$

- Dependency of the Lagrangian

$$\min_{\{\boldsymbol{\pi},\boldsymbol{\mu}\}} L(\boldsymbol{\pi},\boldsymbol{\mu},\lambda_1,\lambda_2) = \sum_{k=1}^{M} J_k$$

$$J_k = J_k(\pi^1,\mu^1,\ldots,\pi^{k-1},\mu^{k-1},\pi^k,\mu^k)$$

JSCPC— Solution Algorithm contd.



JSCPC— Solution Algorithm contd.

$$J_k = J_k(\pi_1, \mu_1, \ldots, \pi_{k-1}, \mu_{k-1}, \pi_k, \mu_k)$$

$$\mathrm{E}[D_k] = (1-\rho_k)\mathrm{E}[D_{R,k}] + \rho_k(1-\rho_{k-1})\mathrm{E}[D_{C,k}] + \rho_k \rho_{k-1} \mathrm{E}[D_{Z,k}]$$

$$J_k = J_k(\mu^{k-1}, \pi^{k-1}, \mu^k, \pi^k, w(k))$$

- Do DP with respect to "system state", w(k), instead of to choice of source coding parameters and service classes

---

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

$$\tau(k) = T_{\max} - \Delta T_{eb}(k) = T_{\max} - w(k) - \frac{B_k(\mu_k)}{R_k} = T_{\max} - w(k+1) + T_p$$

$$w(k) = w(k-1) + \frac{B_{k-1}(\mu_{k-1})}{R_{k-1}} - T_p$$

JSCPC— Solution Algorithm contd.

$$s(k) = \sum_{j=1}^{k} T(j): \quad \sum_{k=1}^{M} T(k) \le T_0 \;\Rightarrow\; s(M) \le T_0$$

$$w(k) = w(k-1) + T(k-1) - T_p$$

$$= w(1) + \sum_{j=1}^{k-1} T(j) - (k-1)T_p = w(1) + s(k-1) - (k-1)T_p$$

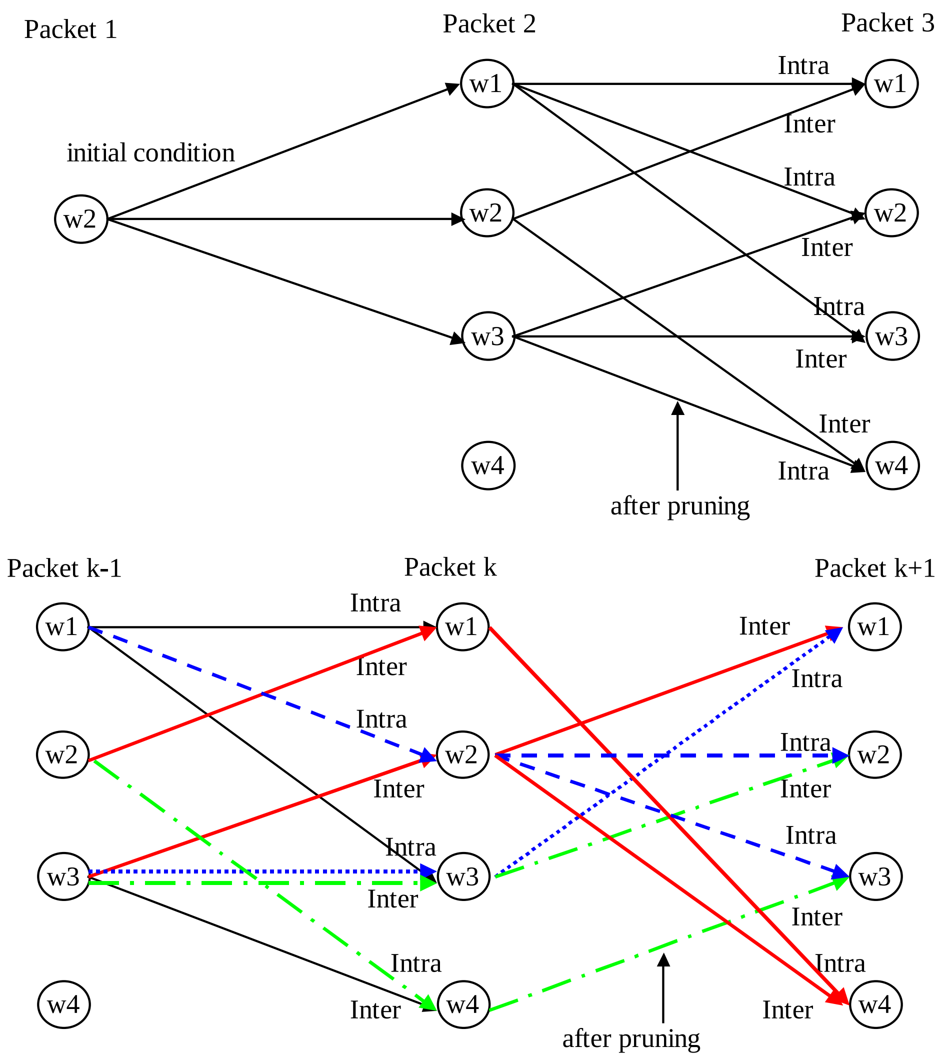
$$w(k) - w(1) + (k-1)T_p \le T_0 \quad \text{for } k = 1,\ldots,M$$

- DP solution

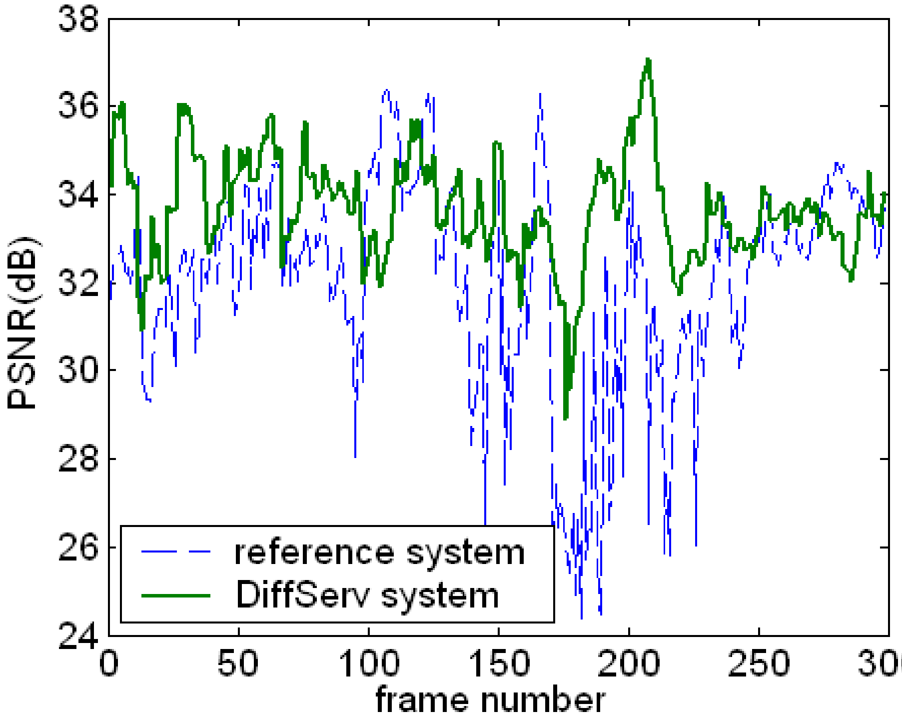
$$\min_{\{u(k)\in \mathbf{U}(w(k))\}} \sum_{k=1}^{M} J(k) = \sum_{k=1}^{M} J(\mu_{k-1}, \pi_{k-1}, \mu_k, \pi_k, w(k))$$

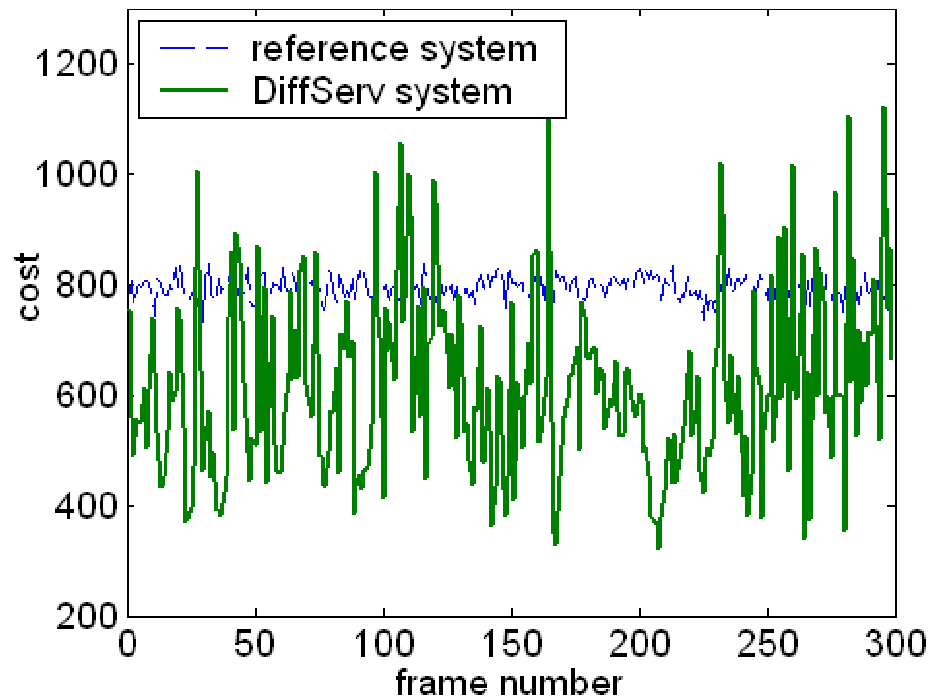
$$\mathbf{U}(w(k)) = \{u(k) \in \mathbf{\Pi} \times \mathbf{Q} : 0 \le \frac{B_k(\mu_k)}{R_k(\pi_k)} + w(k) - T_p \le \min(T_{\max}, T_0 + w(1) - kT_p)\}$$

JSCPC— Solution Algorithm contd.



JSCPC — Experimental Results





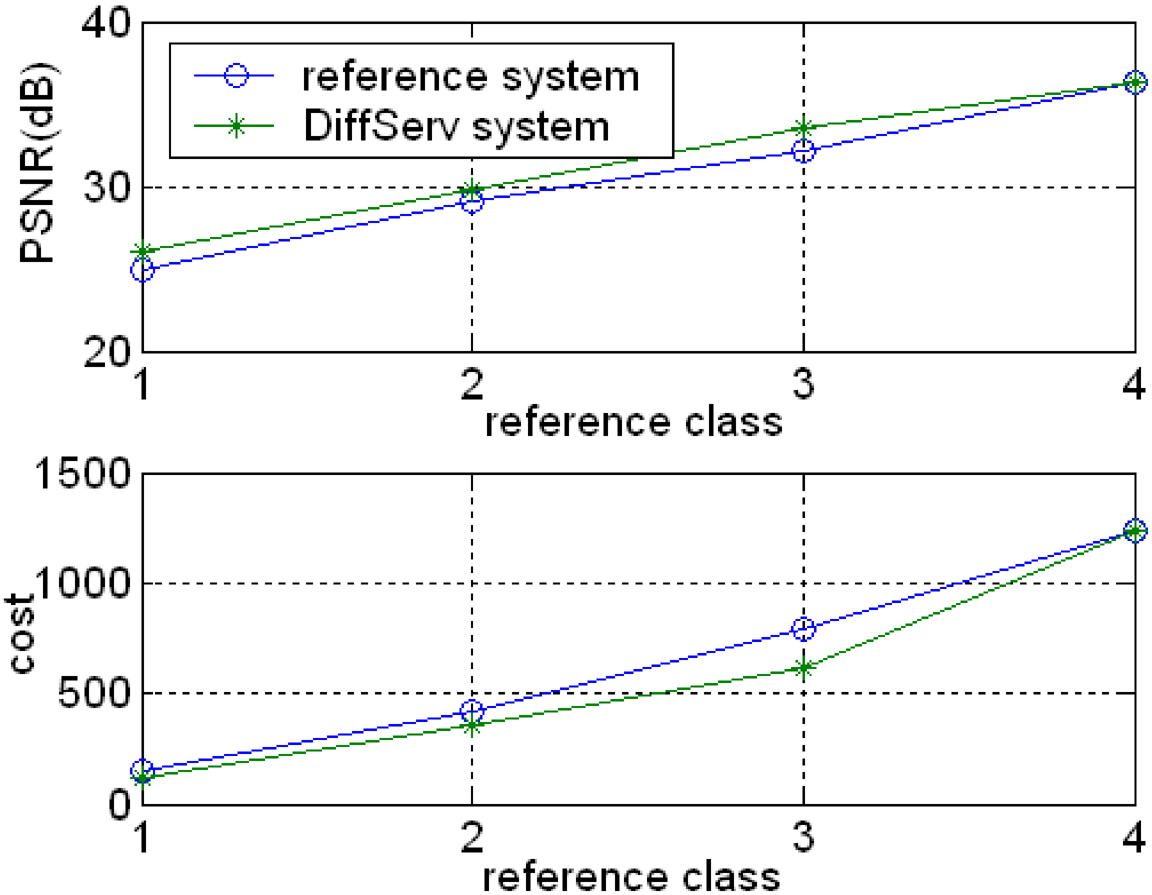
Minimum distortion approach Minimum cost approach

QCIF Foreman, *F*=30 fps, reference class = 3

JSCPC — Experimental Results

minimum distortion approach

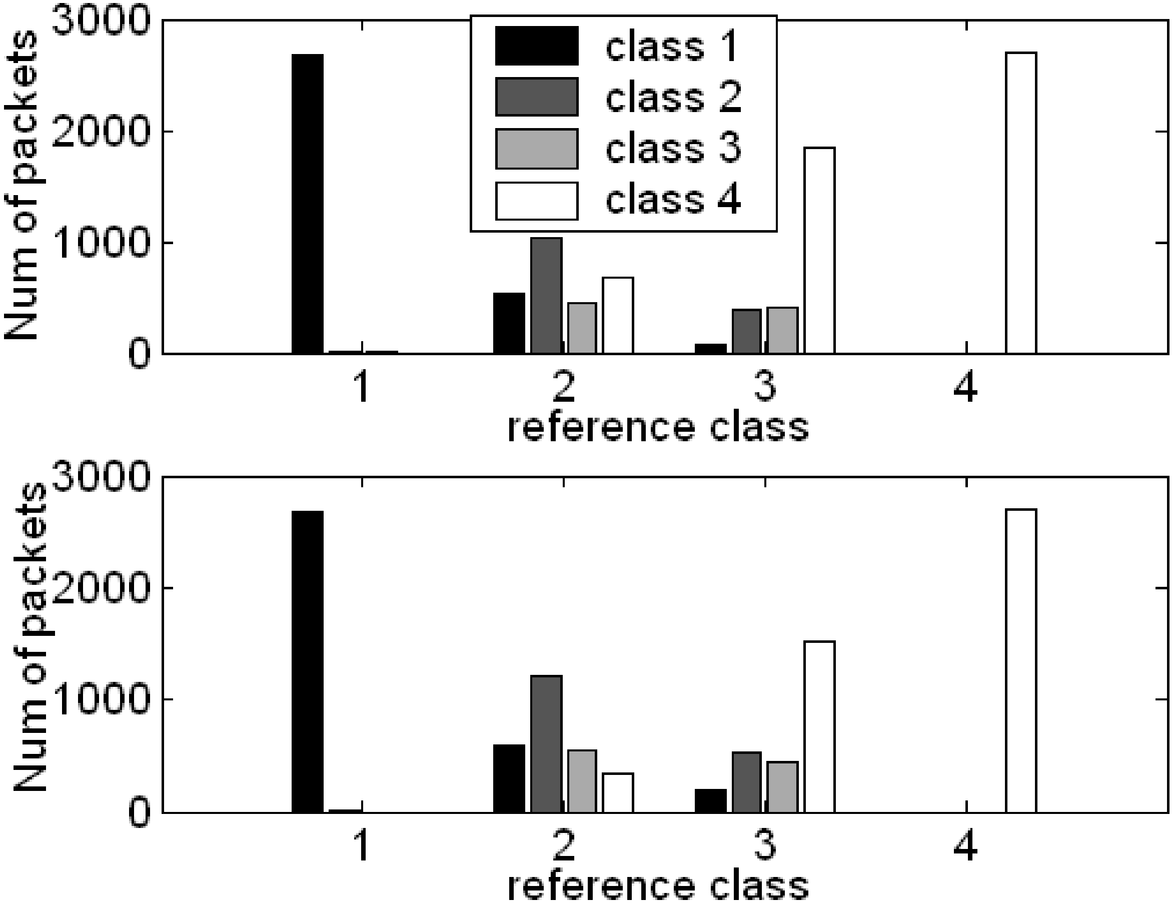
minimum cost approach



Distribution of packet classification

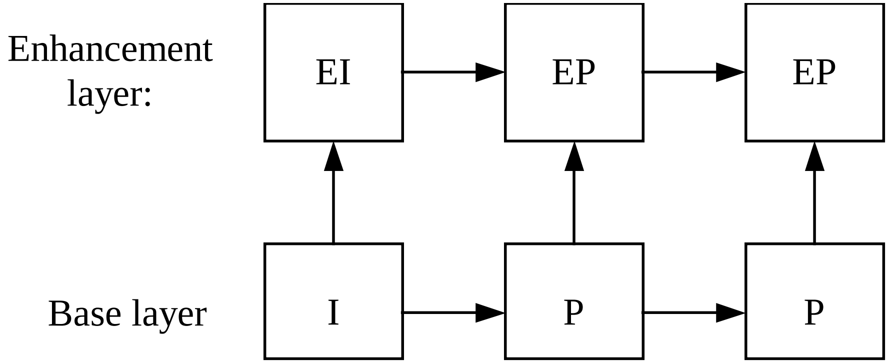
minimum distortion approach

minimum cost approach



Internet Scalable Video

- JSCC based on H.263+ SNR scalability codec



- Problem Formulation

$$\min_{\{\boldsymbol{\mu}^{(b)},\boldsymbol{\nu}^{(b)},\boldsymbol{\mu}^{(e)},\boldsymbol{\nu}^{(e)}\}} E[D] = E[D^{(b)}(\boldsymbol{\mu}^{(b)},\boldsymbol{\nu}^{(b)})] + E[D^{(e)}(\boldsymbol{\mu}^{(e)},\boldsymbol{\nu}^{(e)})]$$

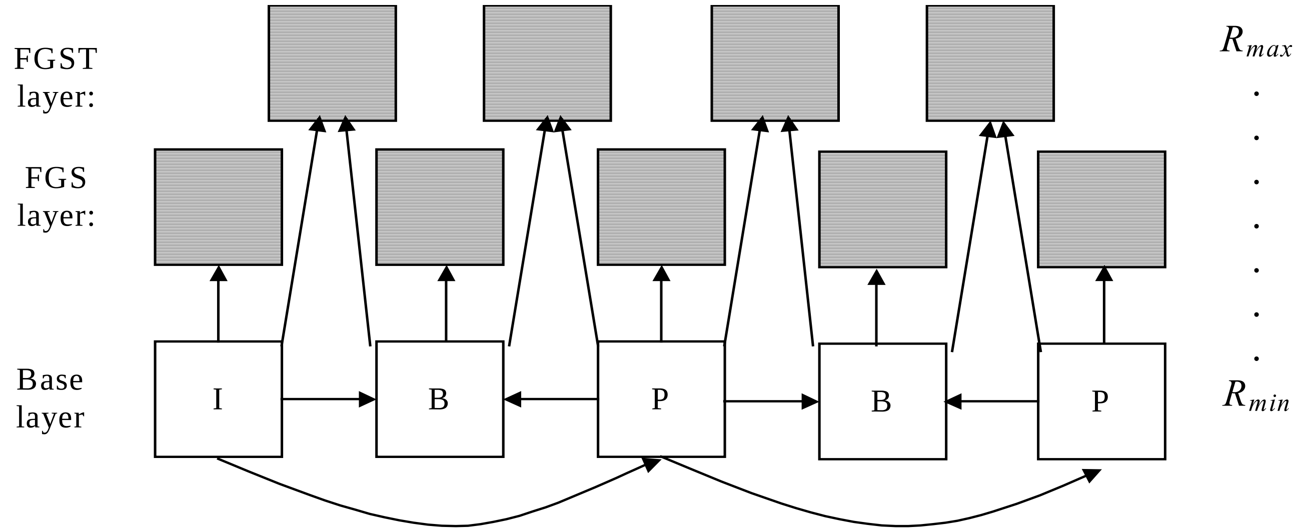
$$\text{s.t.} \quad T^{(b)} \leq T_0^{(b)}$$

$$T^{(b)} + T^{(e)} \leq T_0$$

F. Zhai, R. Berry, T. N. Pappas, and A. K. Katsaggeos, “Rate-distortion optimized error control scheme for scalable video streaming over the Internet,” *Proc. IEEE ICME*, July 2003

Wireless Scalable Video

- Optimal power allocation based on MPEG-4 FGS codec



- Problem Formulation

$$\min_{\{P_i^{(e)}\}} \left\{ D^{(b)} - E[\Delta^{(e)}] \right\}$$
$$\text{s.t.} \quad E_{tot} = E_0$$

- Assume base layer is pre-encoded and always received correctly

C. Costa, Y. Eisenberg, F. Zhai, and A. K. Katsaggelos, "Energy efficient wireless transmission of MPEG-4 Fine Granular Scalable video," *IEEE ICC'04*, 2004, accepted.

Conclusions

- Objective:
  - Optimal error control to achieve the best video quality
- Method:
  - Optimal cross-layer resource allocation
- Base:
  - Resource-distortion optimization framework
- Scope:
  - End-system design for video communication
- Have studied following applications:
  - JSCC—Internet video (Allerton'03, ICC'04, ICIP'04, TIM'04)
  - JSCCPA—wireless video (Allerton'03, ICASSP'04)
  - JSCPC– DiffServ video (ICIP'03, TMM'04)
  - Scalable video (ICME'03, ICC'04)

Future Work

- Rate control
  - Incorporate rate control into the optimization framework
- Channel model
  - Take into account the correlation of packet loss, e.g., Gilbert model
- Cross-layer Design
  - Consider modifying the current network protocols:
    Link-layer ARQ, UDPLite,

Q&A

Thank You!