Analysis of Rate-distortion Functions and Congestion Control in Scalable Internet Video Streaming

Min Dai

Electrical Engineering, Texas A&M University

Dmitri Loguinov

Computer Science, Texas A&M University

Motivation

- Scalable coding is widely applied in Internet streaming.
  - *Fine Granular Scalability* (FGS) has been chosen in MPEG-4 standard
  - Study the statistical properties of FGS encoder and propose a more accurate statistical model for it
- The Rate-distortion (R-D) theory is a powerful tool in Internet streaming.
  - Choose appropriate compression schemes
  - Optimally allocate bits in joint source-channel coding
  - Rate adaptation in the Internet

Motivation (cont.)

- R-D theory (cont.):
  - No current closed-form R-D model has been developed for scalable coding
  - Derive an R-D model for scalable video coding
- Constant quality control:
  - Another application of R-D model in Internet streaming
  - Human eyes are sensitive to quality fluctuation
  - Many video sequences have severe quality fluctuations

Motivation (cont.)

- Most existing CQ methods in scalable streaming are limited to the constant bitrate (CBR) case.
  - Channel bandwidth is varying in a real network
  - Congestion control is necessary to allow fair and efficient usage of network bandwidth
- Most existing congestion control methods (e.g., AIMD) are not proven to be asymptotically stable.
  - Kelly's continuous-feedback congestion control
  - Combine our R-D model with Kelly's control in Internet streaming

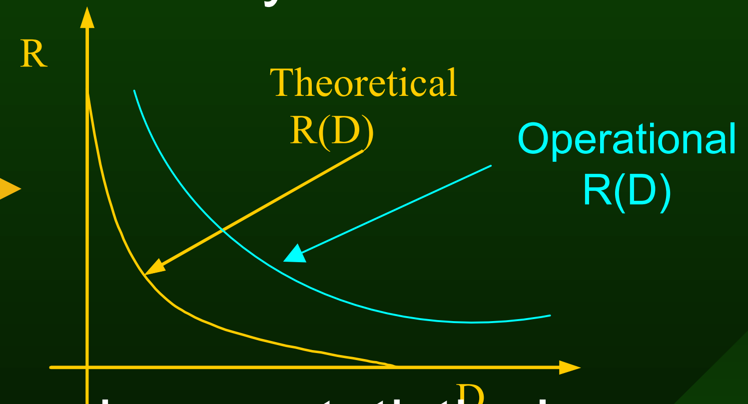
Overview of this Talk

- Background on rate-distortion theory and FGS scalable coding
- A big picture of this work
  - R-D modeling of an FGS encoder
  - Kelly's control
  - Constant quality control
- Experimental results
- Conclusion

Background

- Rate-distortion (R-D) theory
  - The theoretical discipline that treats data compression from the viewpoint of information theory

A typical R-D curve →

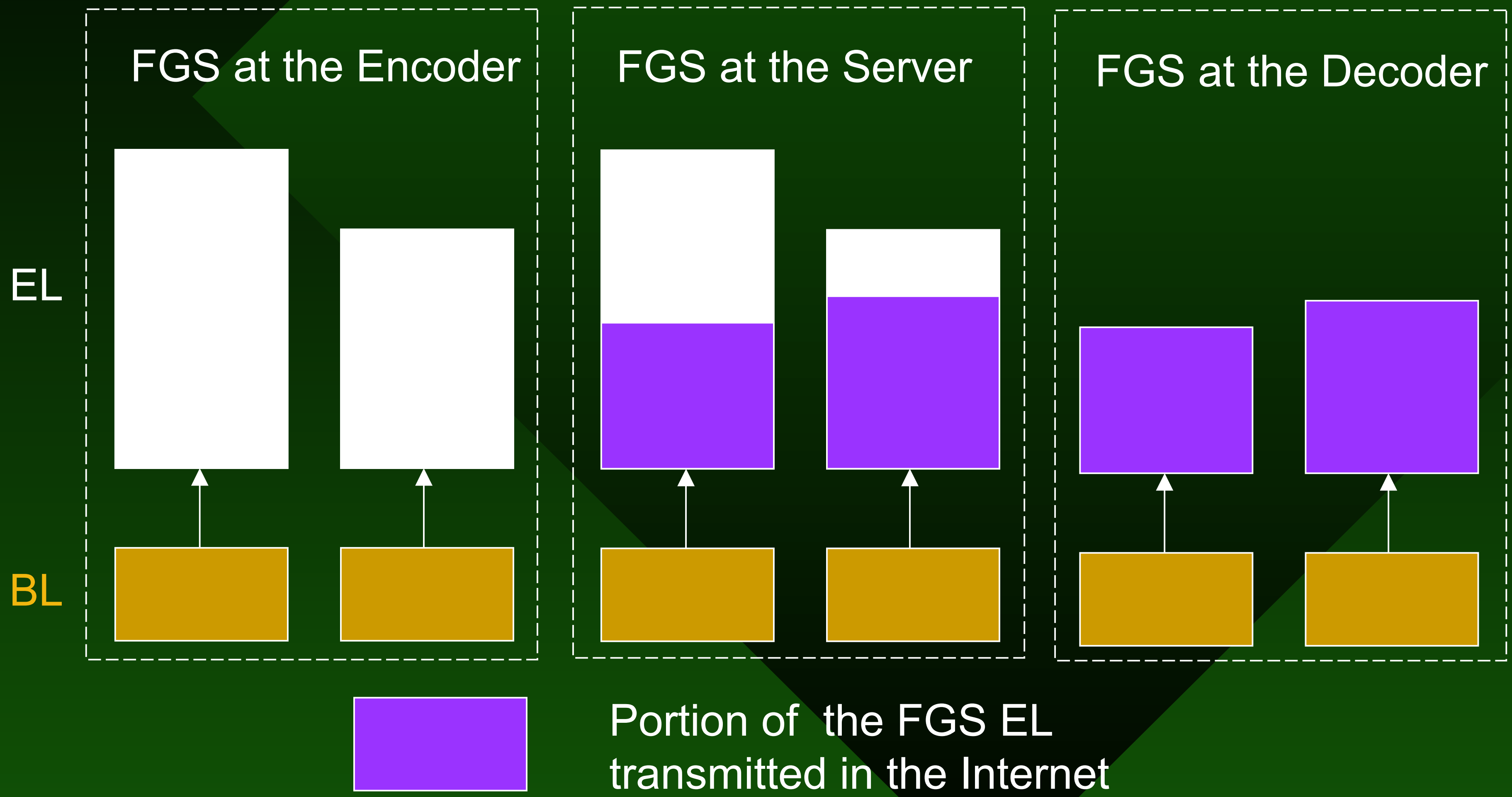


- Theoretical R-D model:
  - A lower bound for any encoder given a statistical distribution of the source
  - Often unachievable in the real world
- Operational R-D model
  - An achievable bound for a practical encoder

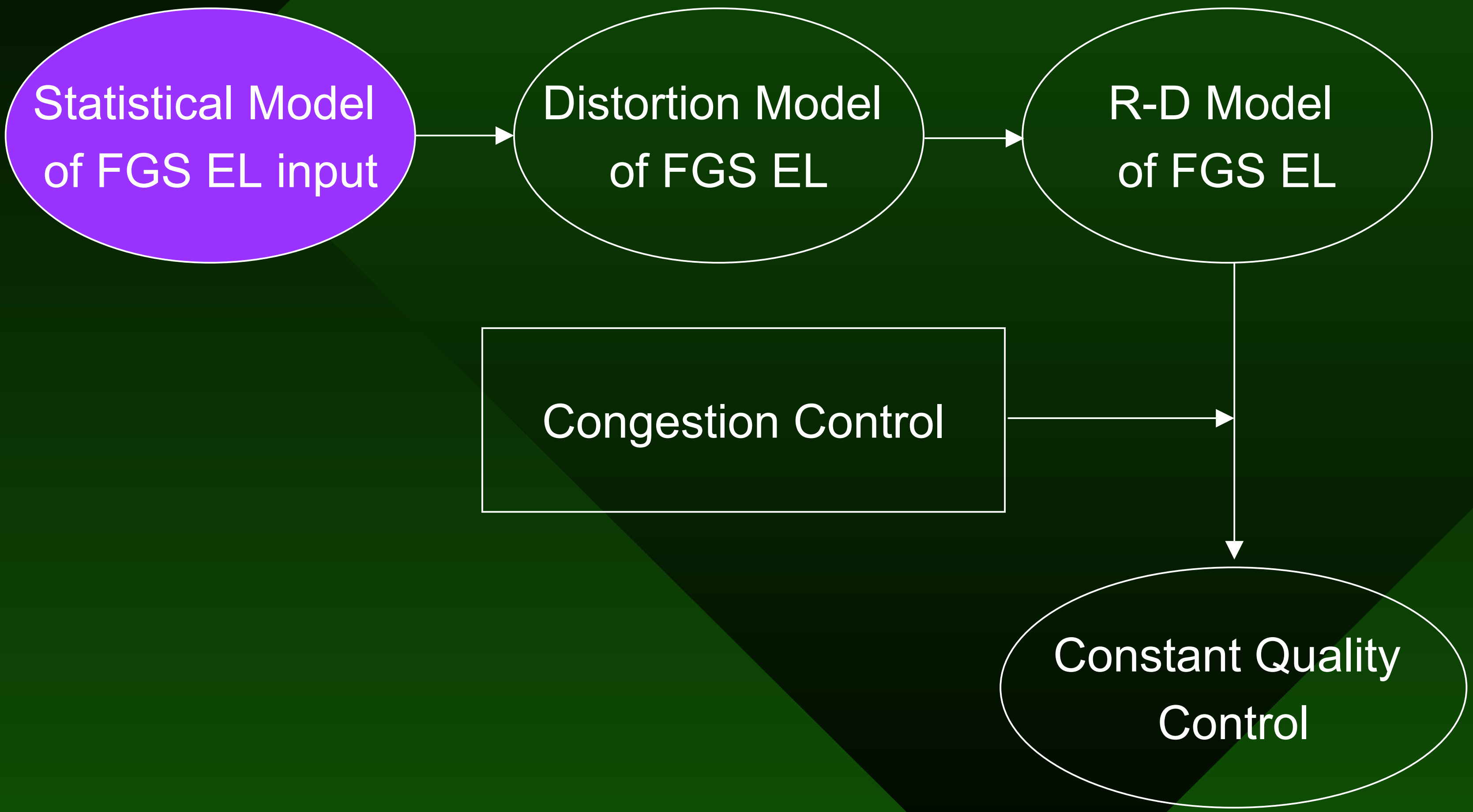
Background (cont.)

- Scalable coding:
  - Provides the capability of recovering image or video information by partially decoding the compressed bitstream
- Fine granular scalability (FGS):
  - One low bitrate base layer (BL) to provide a low but guaranteed level of quality
  - One high bitrate enhancement layer (EL) to provide finer quality improvement
  - EL can be truncated at any codeword

Background (cont.)



A Big Picture



Related work on Statistical Models

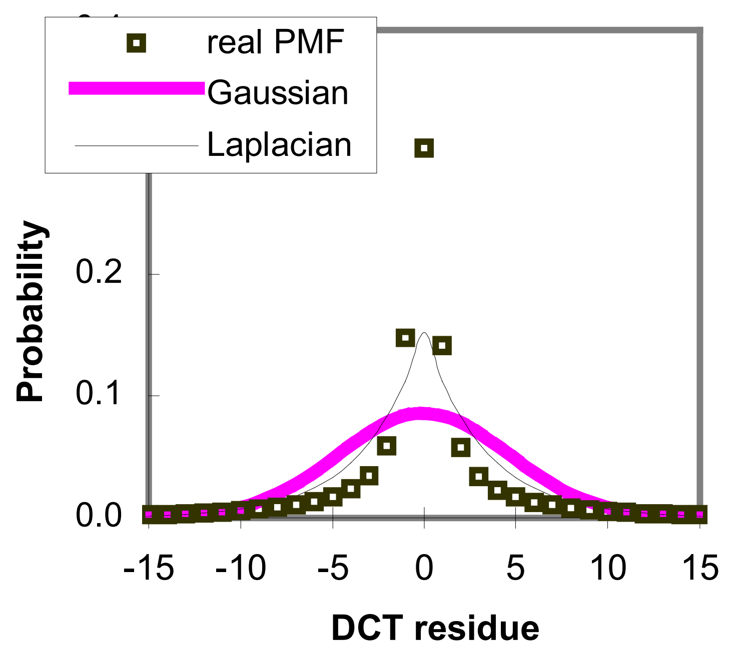
- Input to FGS EL:
  - DCT residue between the original image and the reconstructed image from BL
- The two most popular models for DCT residue:
  - Zero-mean Gaussian distribution:

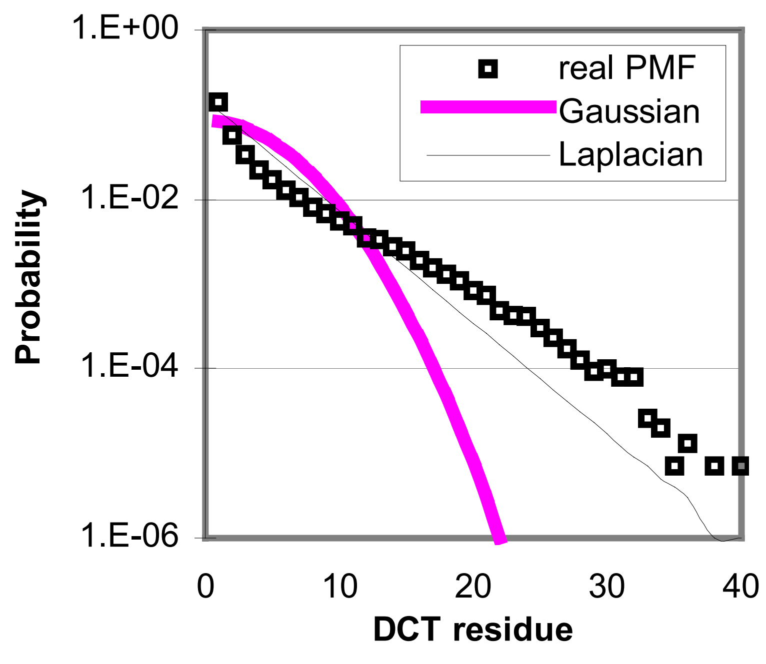
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

  - Laplacian distribution (double exponential):

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$$

Related work on Statistical Models





- The PMF of DCT residue with Gaussian and Laplacian estimations (left). Logarithmic scale of PMFs for the positive residue (right). All testing sequences shown in this paper are coded at 10fps and 128 kb/s in the base layer

Proposed Statistical Model

- Mixture Laplacian model:

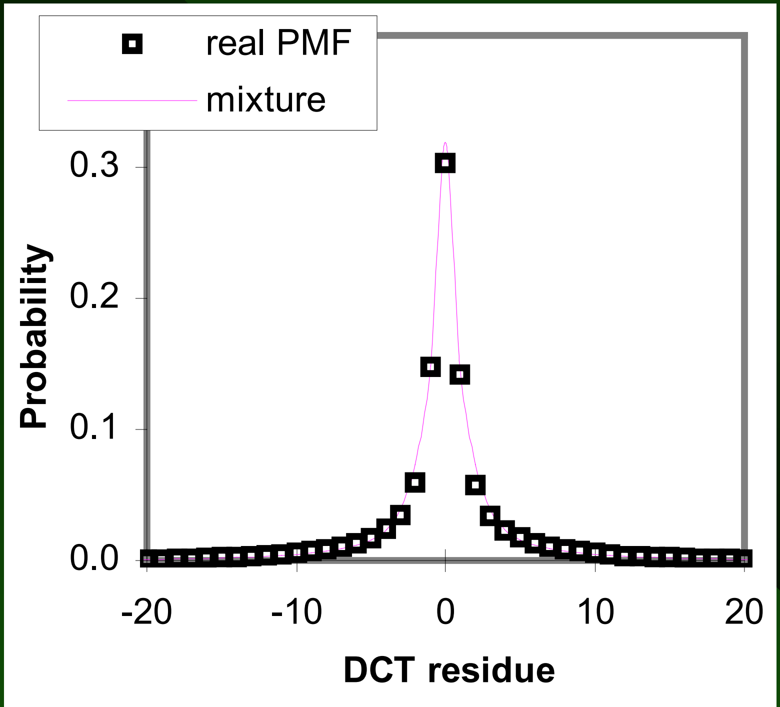
$$f(x) = p\frac{\lambda_0}{2}e^{-\lambda_0|x|} + (1-p)\frac{\lambda_1}{2}e^{-\lambda_1|x|}$$

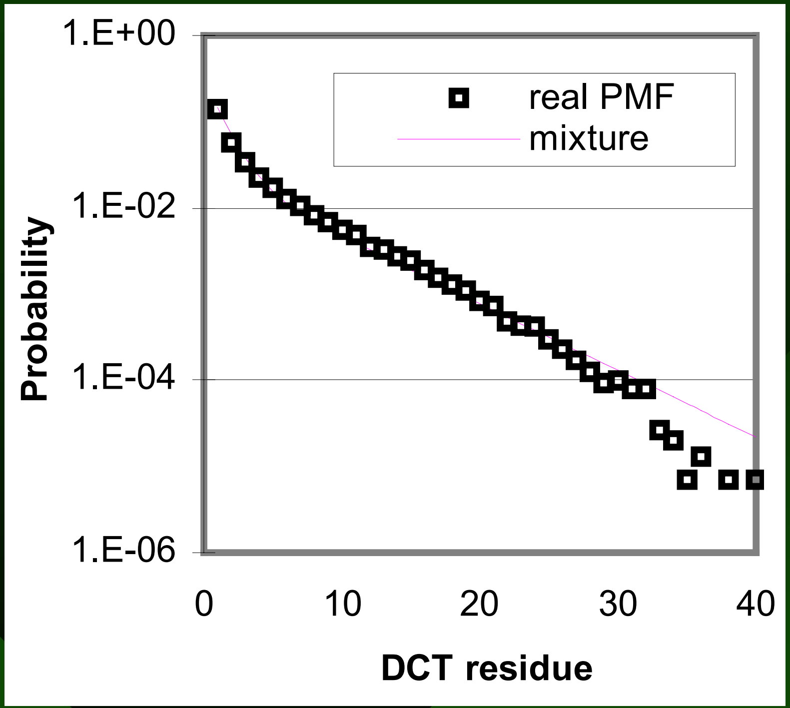
where λ_0 denotes the small variance Laplacian distribution and λ_1 denotes the large variance Laplacian distribution

- Use Expectation-Maximization (EM) algorithm to give Maximum-likelihood (ML) estimation for parameters { p, λ_0, λ_1 }

Proposed Statistical Model (cont.)

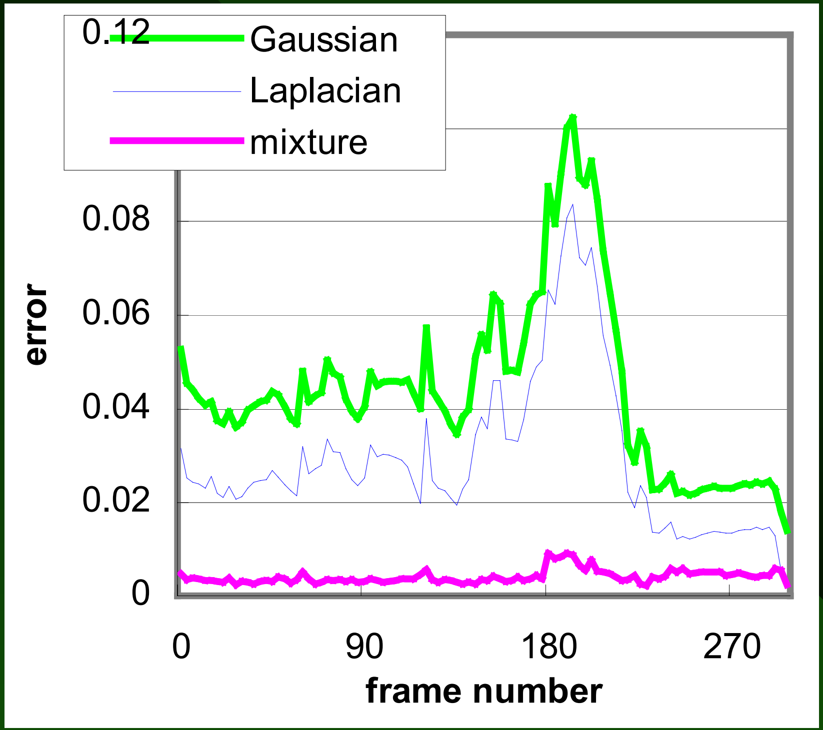
- Real PMF and mixture Laplacian (left) and Logarithmic scale of the positive part (right)

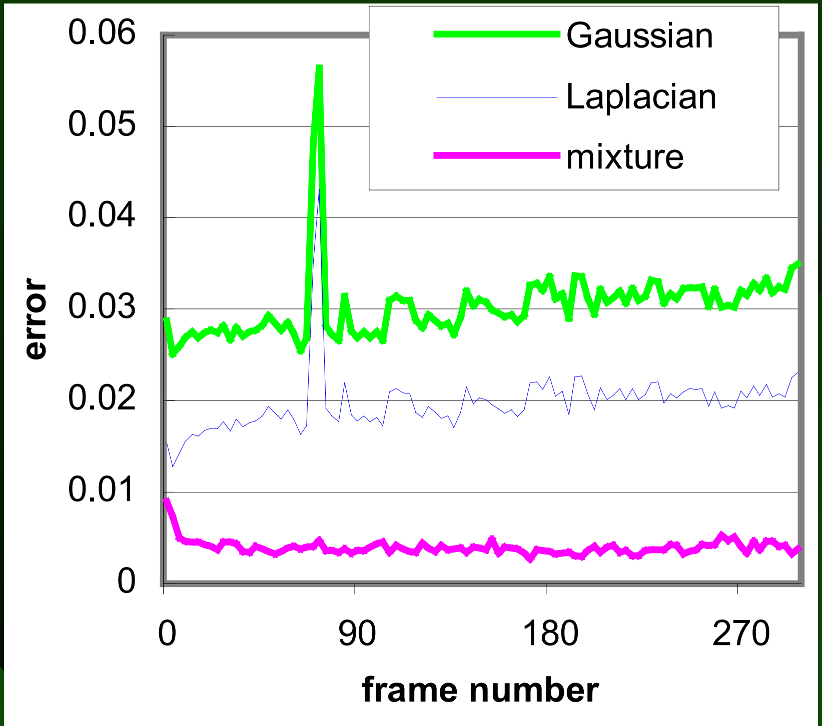




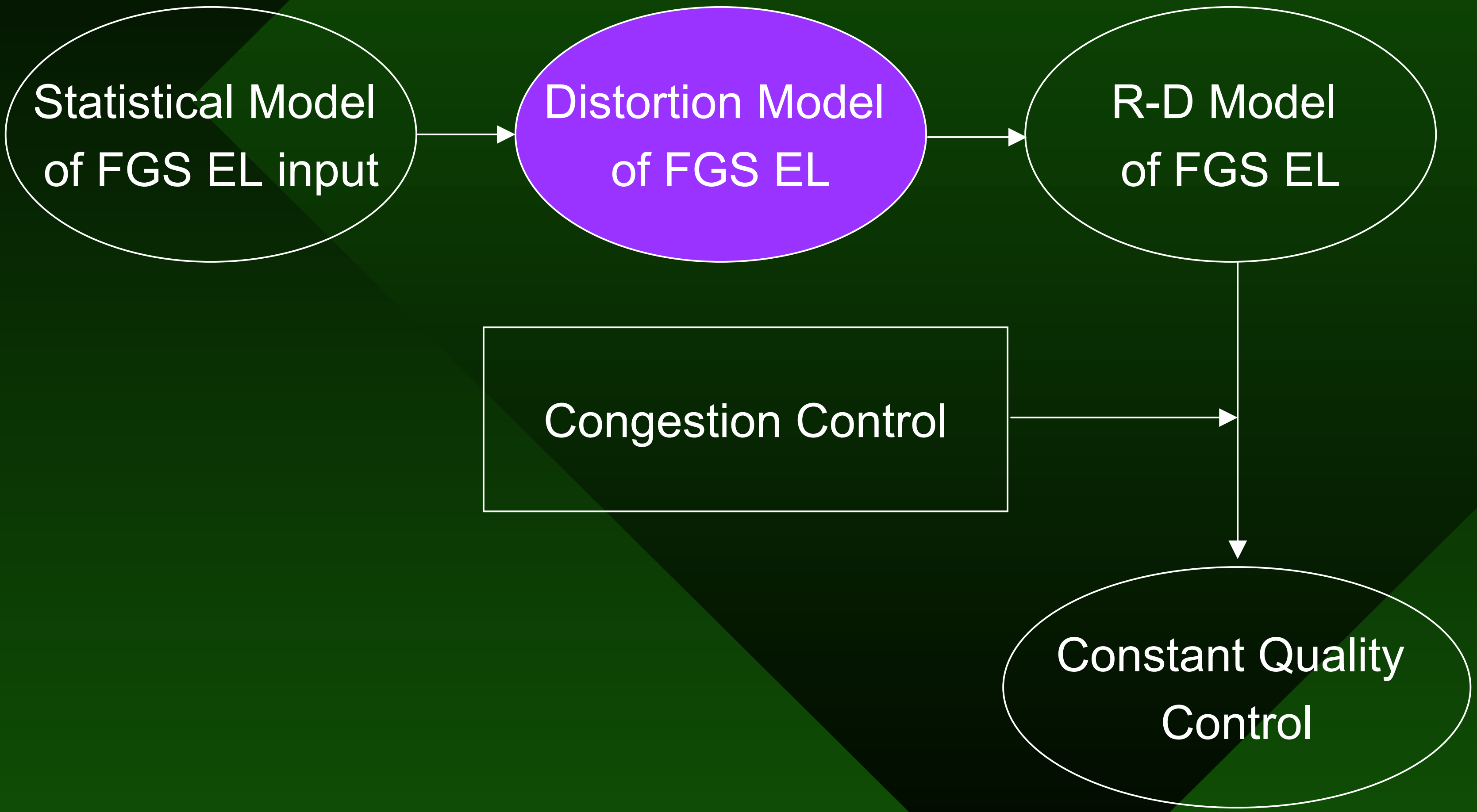
More Results

- The weighted absolute error of estimations in Foreman CIF (left) and Coastguard CIF (right)





A Big Picture



Current Distortion Models

- Classical model:

$$D = \varepsilon^2 \sigma_X^2 2^{-2R}$$

  – where ε^2 is a signal-dependent constant, σ_X^2 denotes the signal variance and R is the bitrate

- A variation of the classical model (proposed by Chiang *et al*. in 1997):

$$R = aD^{-1} + bD^{-2}$$

  – where parameters a, b are obtained empirically

Current Distortion Models (cont.)

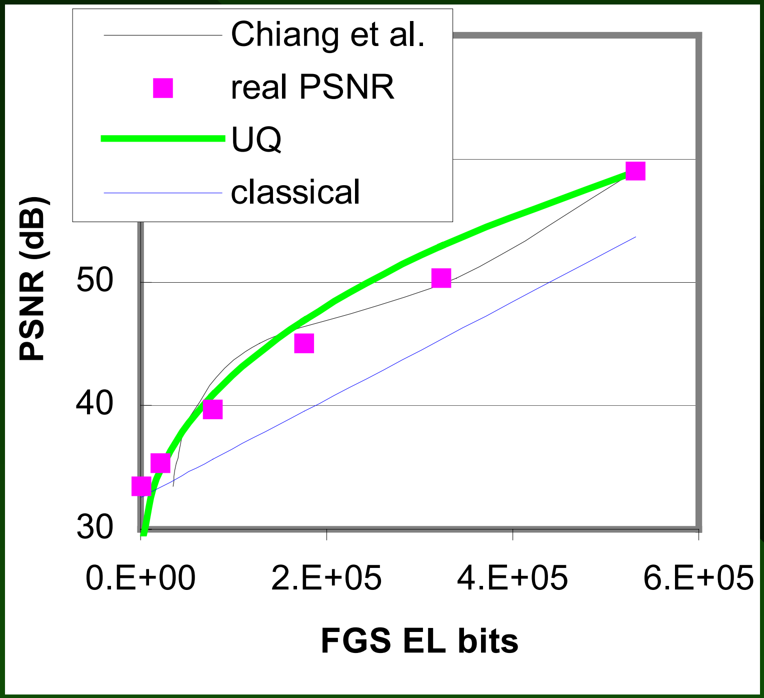
- Distortion model for Uniform Quantizer (UQ):

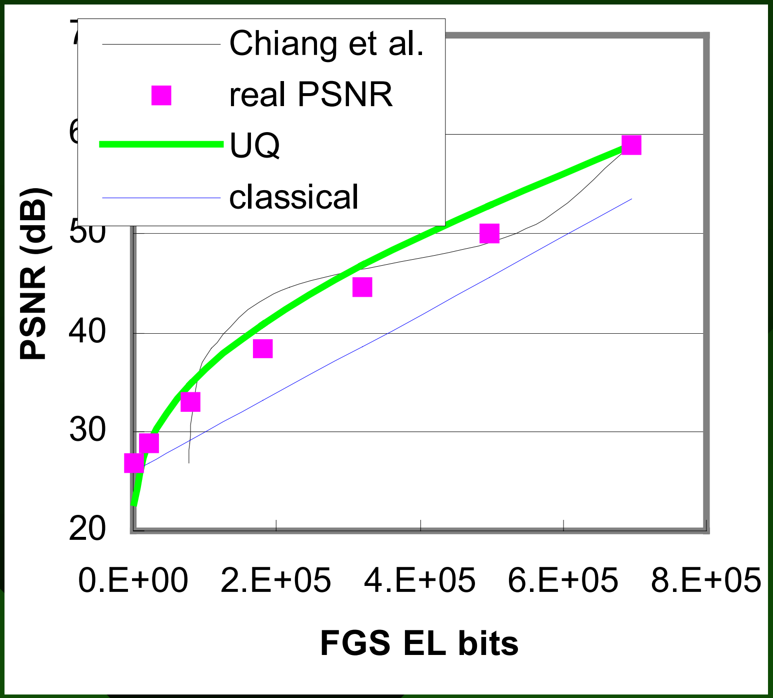
$$D(\Delta) = \frac{\Delta^2}{\beta}$$

where Δ is quantization parameter (QP) and β equals 12

Current Distortion Models (cont.)

- Performances of current models in frame 0 (left) and frame 252 of Foreman CIF (right)





A more Accurate Distortion Model

- For each Laplacian component in the mixture Laplacian model, the distortion is:

$$D_n(\Delta) = \frac{-1}{(1-e^{-\lambda_n \Delta})}\left[e^{-\lambda_n(\Delta-1)}\left[\left(\Delta - 1 + \frac{1}{\lambda_n}\right)^2 + \frac{1}{\lambda_n^2}\right] - \frac{2}{\lambda_n^2}\right]$$

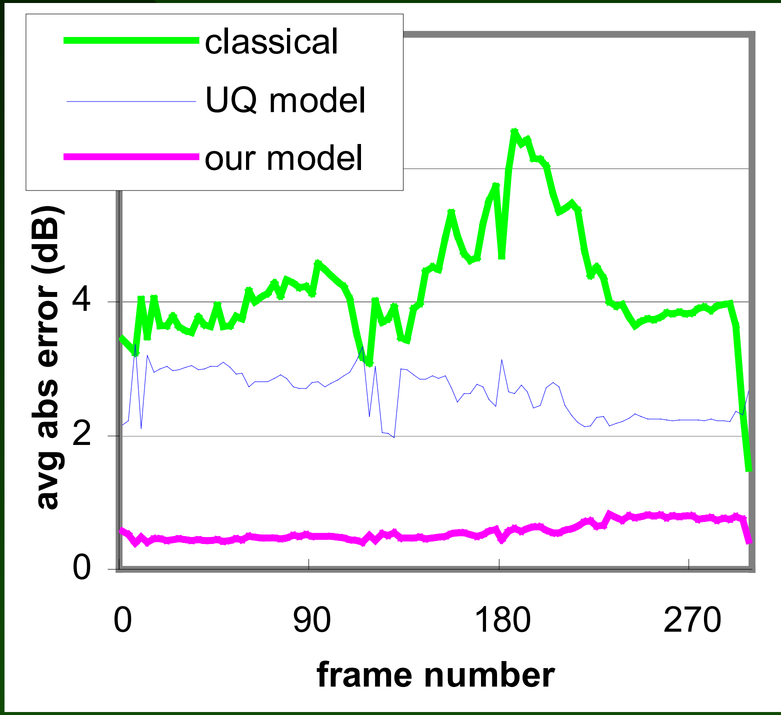
- Final version:

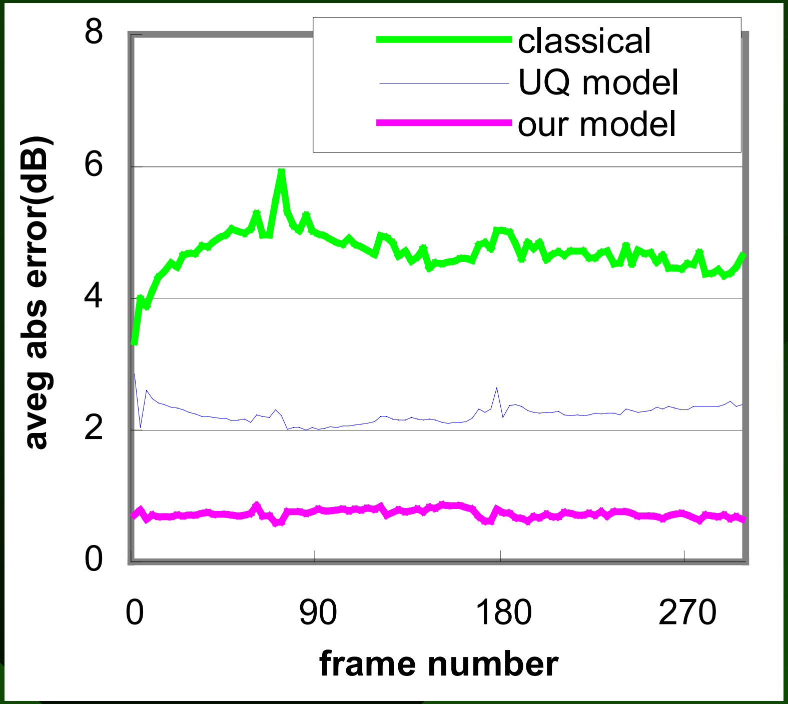
$$D(\Delta) = p \cdot D_0(\Delta) + (1-p) \cdot D_1(\Delta)$$

where Δ is the quantization step of each bitplane in the FGS EL and *p* is the probability of related Laplacian component

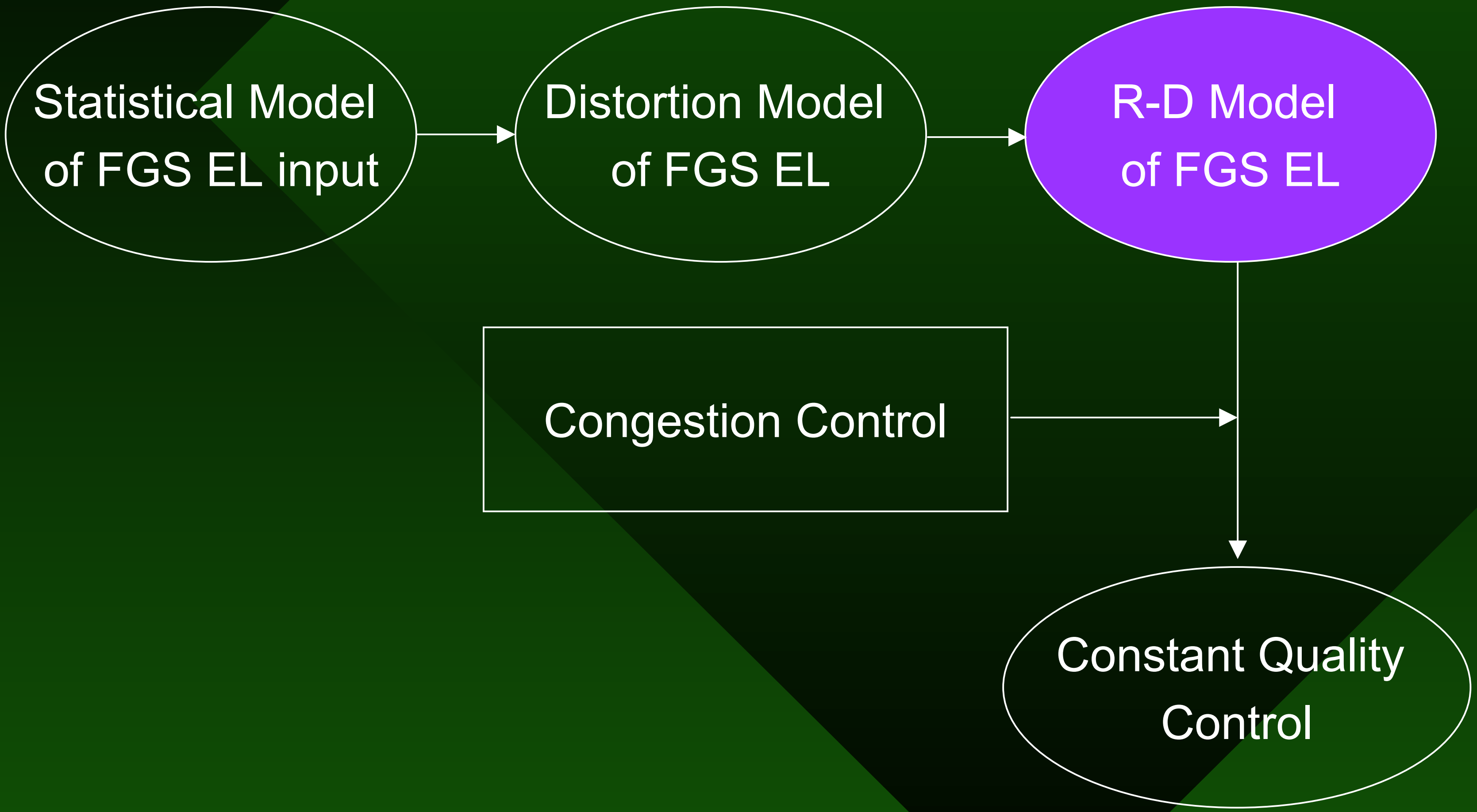
Results of Distortion Model

- The average absolute errors in Foreman CIF (left) and Coastguard CIF (right)





A Big Picture



Operational R-D model

- Peak Signal-to-Noise Ratio (PSNR) is the most popular quality measurement in video coding

$$PSNR = 10\log_{10}(255^2 / D)$$

- Based on our distortion model, we found that PSNR could be described with a quadratic function of bitplane number *z* ($z = \log_2(\Delta)$)

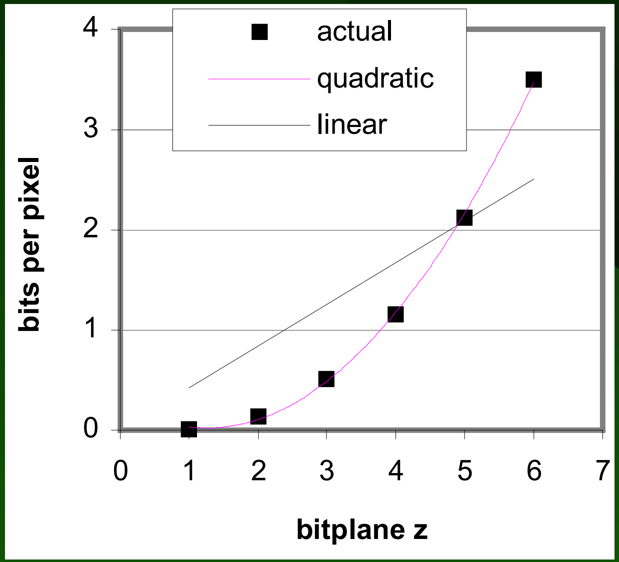
$$PSNR(z) \approx d_1 z^2 + d_2 z + d_3$$

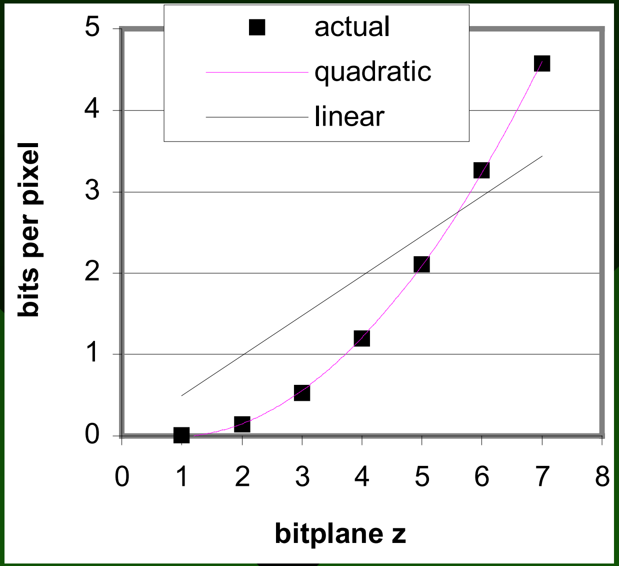
Operational R-D model (cont.)

- In traditional R-D models, bitrate *R* is a linear function of the bitplane number *z*

$$R(z) = az + b$$

- Experimental results show that a quadratic function of z is a much better model of *R*





Operational R-D model (cont.)

Distortion model $D(\Delta)$

PSNR ~ $\log(D)$

Quality PSNR $(z) = d_1z^2 + d_2z + d_3$

Rate $R(z) = a_1z^2 + a_2z + a_3$

$$PSNR = AR + B\sqrt{R} + C$$

$$D(R) = 2^{aR + b\sqrt{R} + c}$$

Operational R-D model (cont.)

- The proposed R-D function is:

$$D(R) = 2^{aR + b\sqrt{R} + c}$$

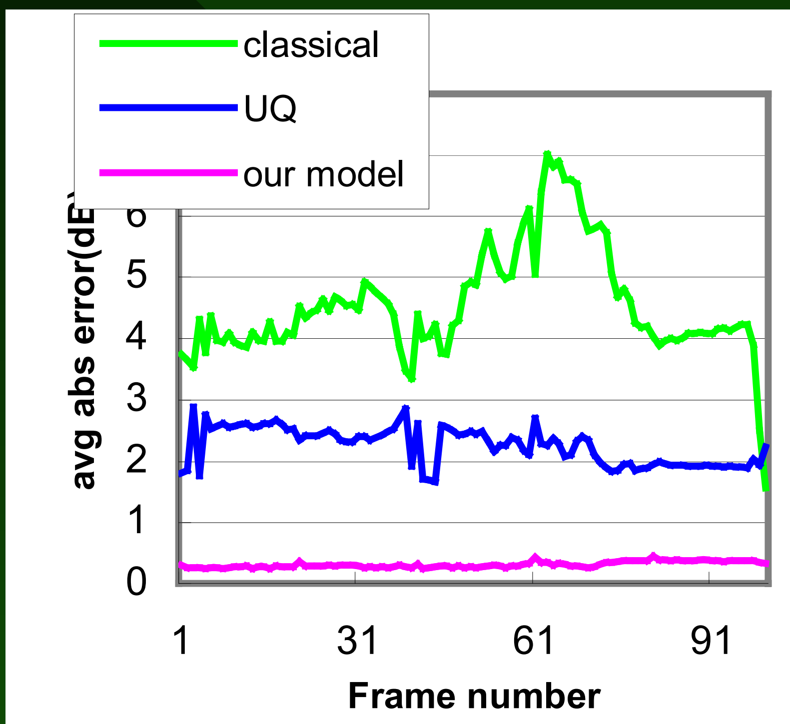
  – where a, b, c are constants

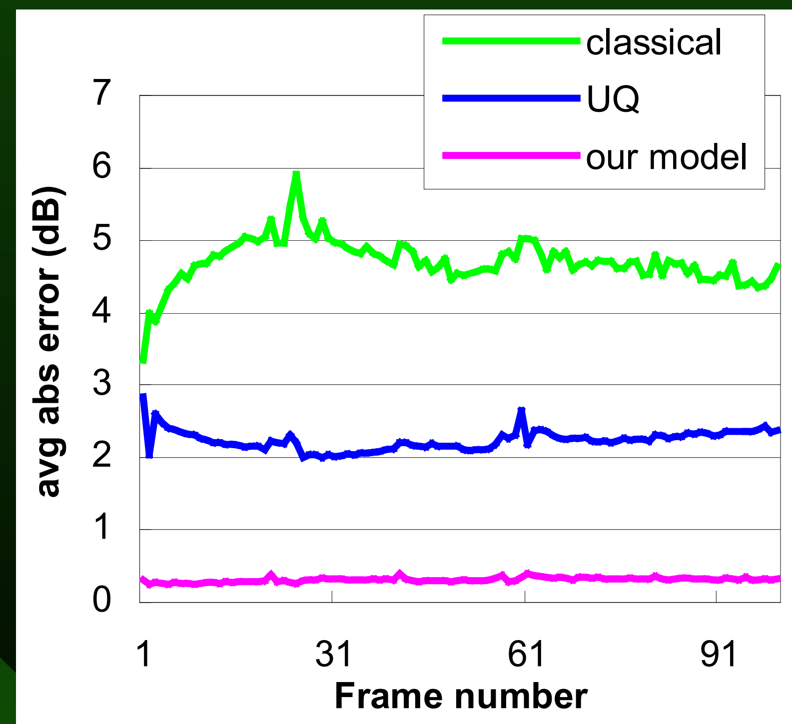
- Notice that the classical model is a special case of our model with *a* = -2 and *b*=0.

$$D(R) = 2^{-2R + \log_2(\varepsilon^2 \sigma_x^2)}$$

Results of R-D models (1)

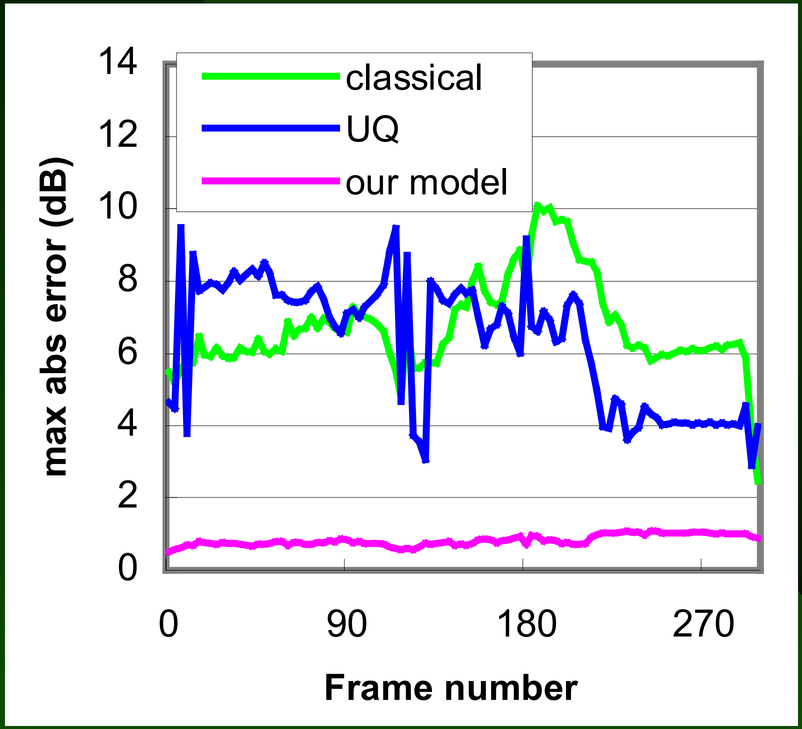
- The average absolute errors in Foreman CIF (left) and Coastguard CIF (right)

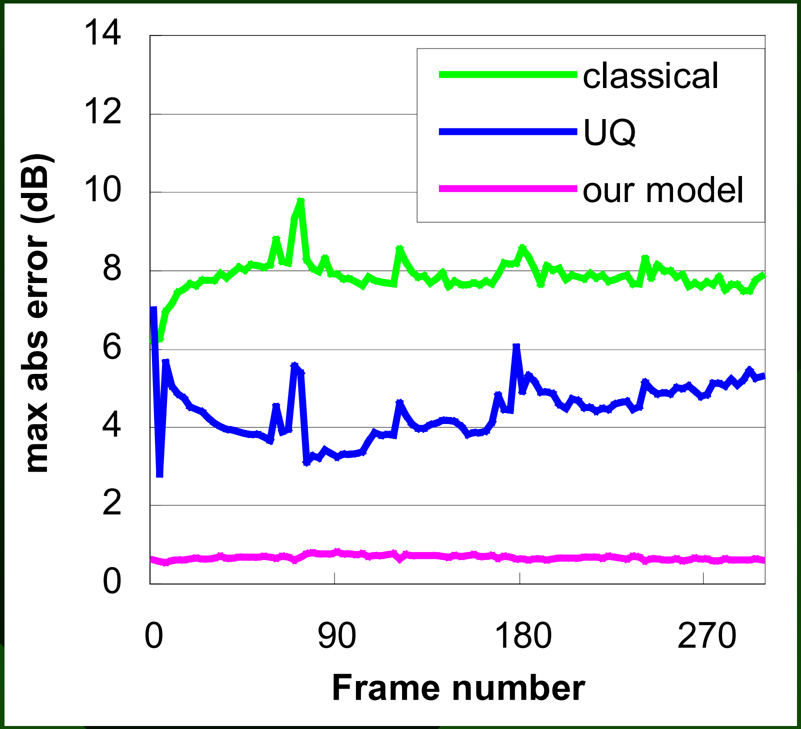




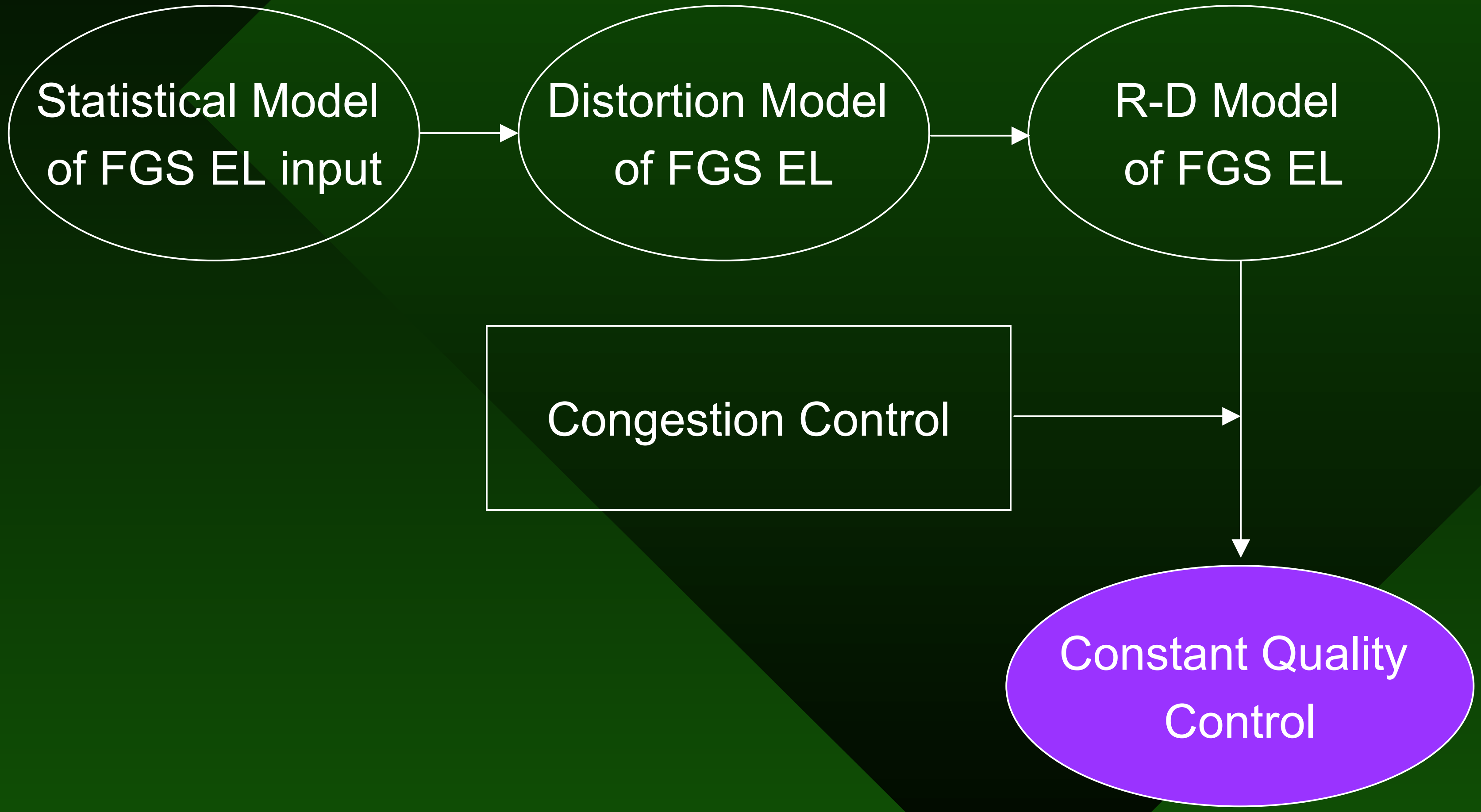
Results of R-D models (2)

- The maximum absolute errors in Foreman CIF (left) and Coastguard CIF (right)

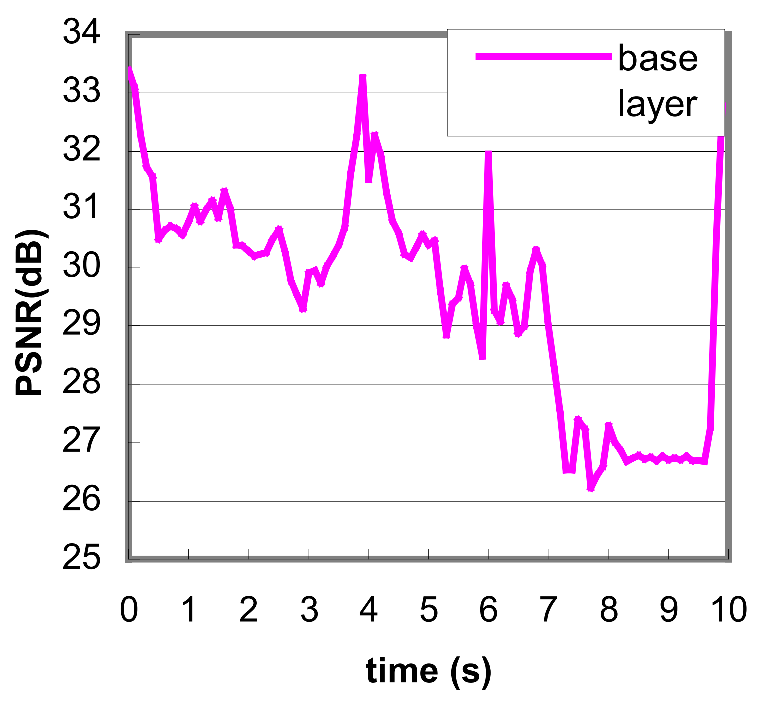


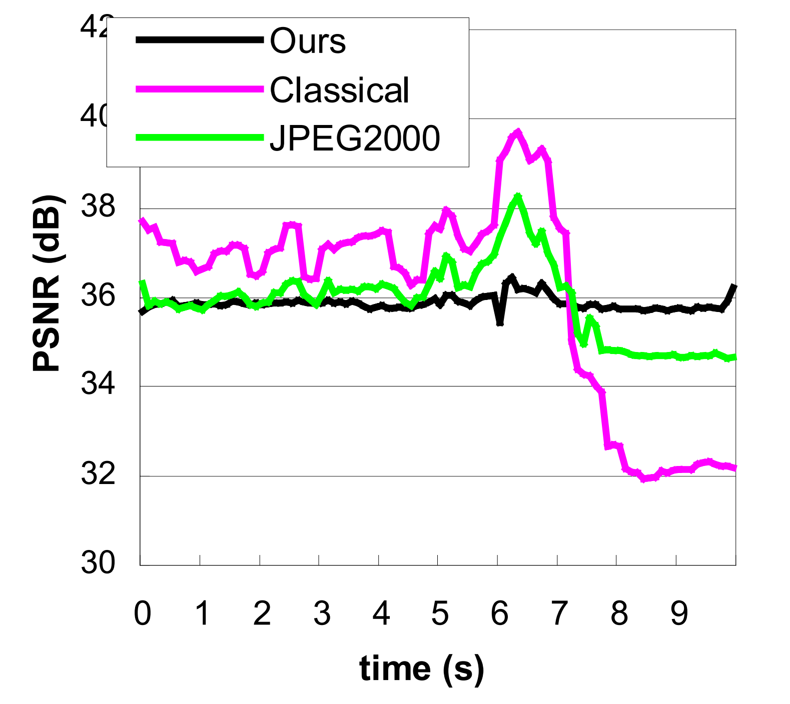


A Big Picture



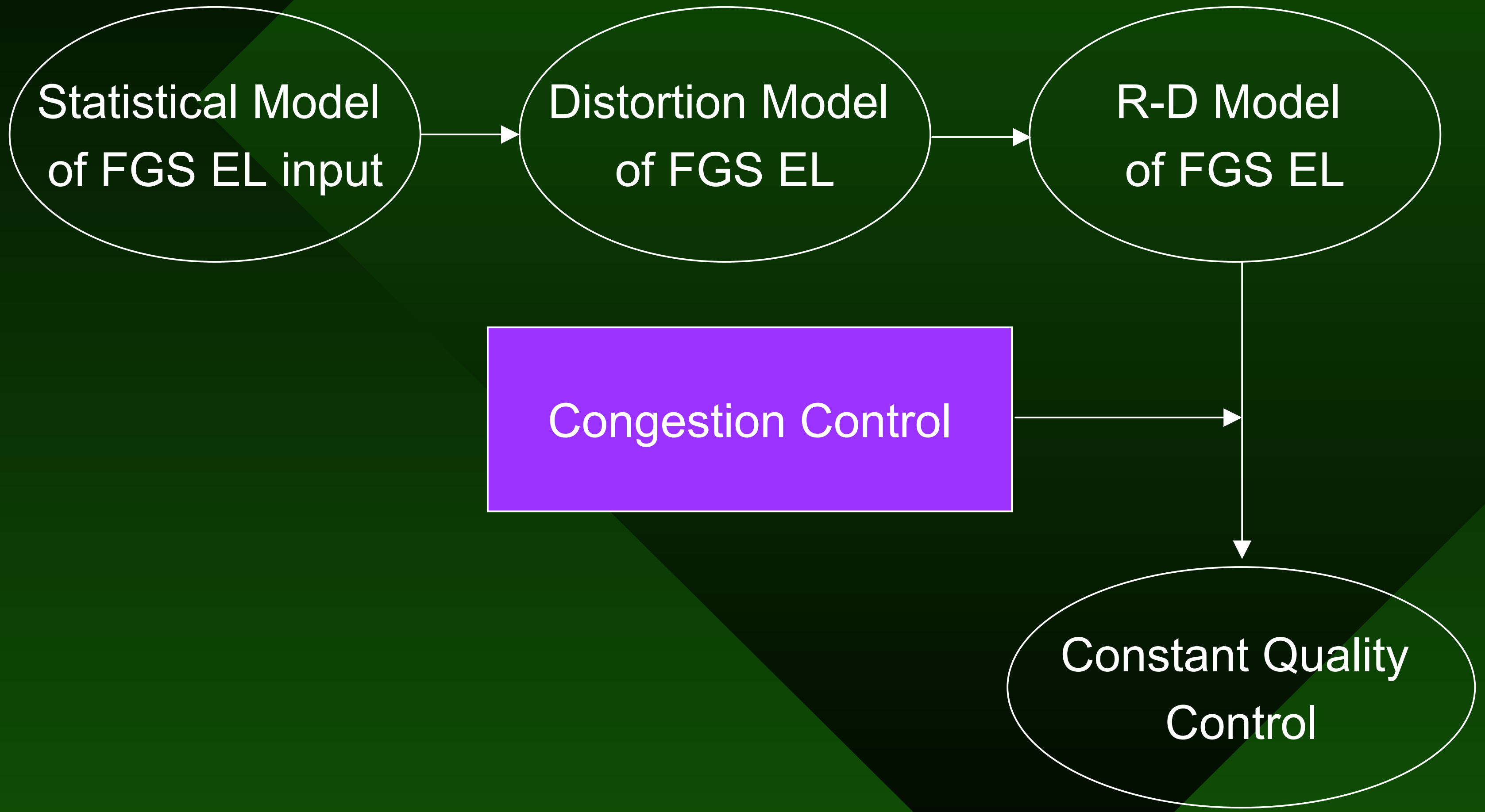
Constant Quality in CBR





- Notice that most CQ papers stop here, while the available bandwidth is varying in the real case

A Big Picture



Congestion Control

- Current status:
  - AIMD, TFRC and binomial algorithms oscillate around the average rate
- Continuous-feedback controller is proposed by Kelly *et al*. in 1998 :

$$\frac{dr(t)}{dt} = r(\alpha U'(r) - \beta \sum_{l \in P} p_l), \ \ where\ U(r) = \log(r)$$

where r is the current sending rate and α, β are constants. U is the utility function of the end user and p_l is the price that the flow pays for using router l along the end-to-end path P

Congestion Control (cont.)

- An application-friendly version:

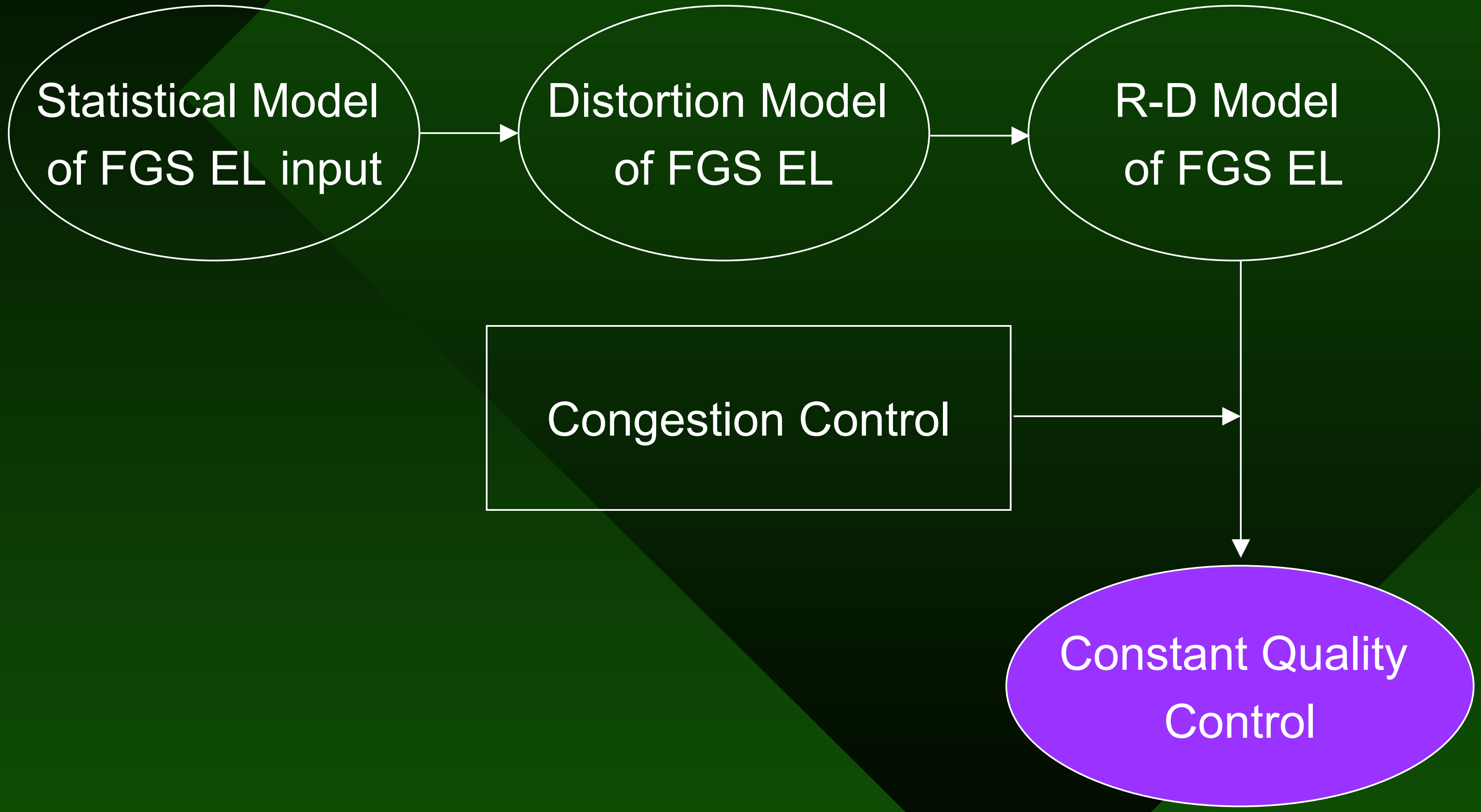
$$\frac{dr(t)}{dt} = \alpha - \beta \cdot p(t) \cdot r(t)$$

- Bottleneck packet loss *p* is used as the feedback instead of prices

$$p_l(t) = \frac{\sum_i r_i(t) - C_l}{\sum_i r_i(t)}$$

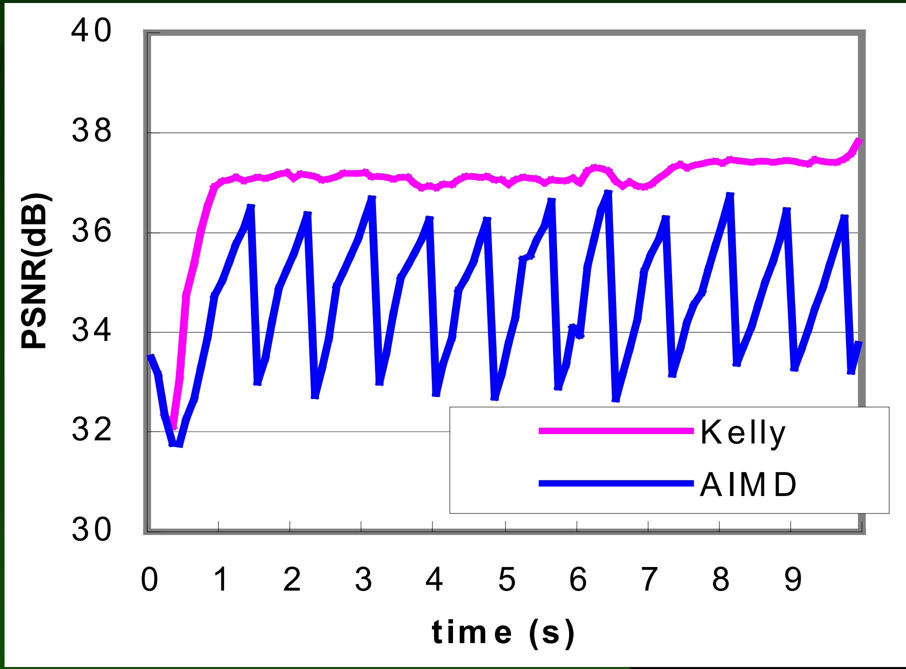
where r_i is the sending rate of the *i*-th flow passing through the bottleneck router *l*, C_l is the speed of router *l*

A Big Picture



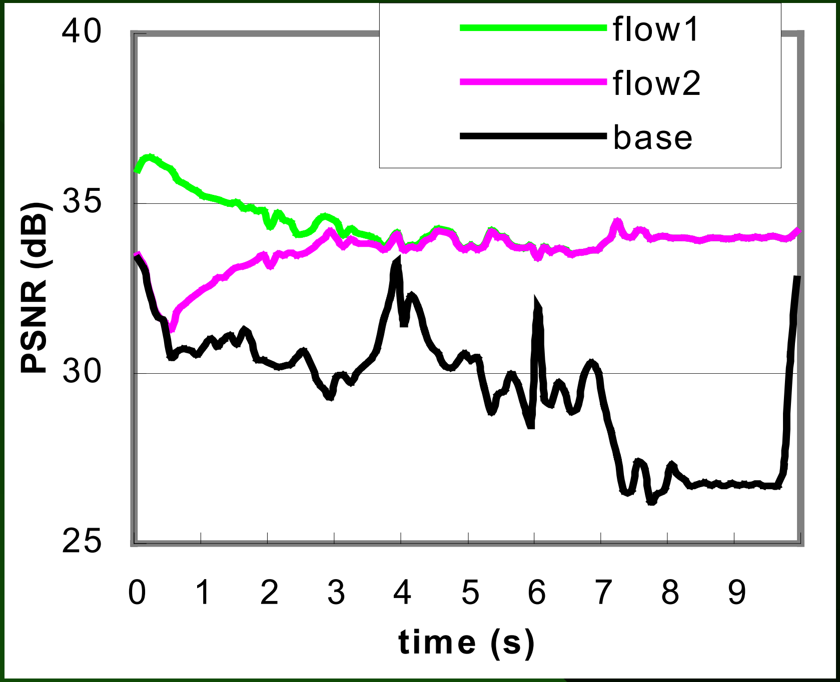
Experimental Results (1)

- Comparison between a single AIMD flow and a single Kelly's flow. Bottleneck bandwidth *C* is 1mb/s and RTT=100ms



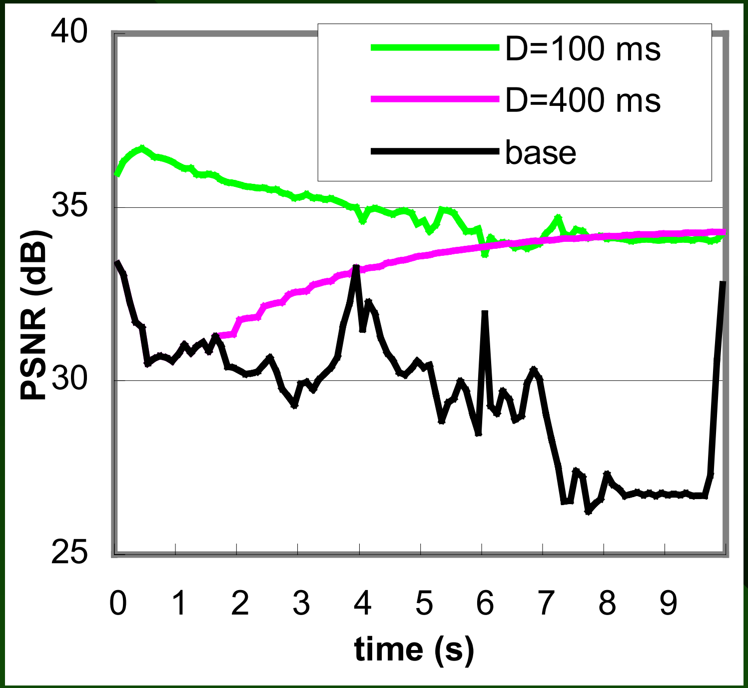
Experimental Results (2)

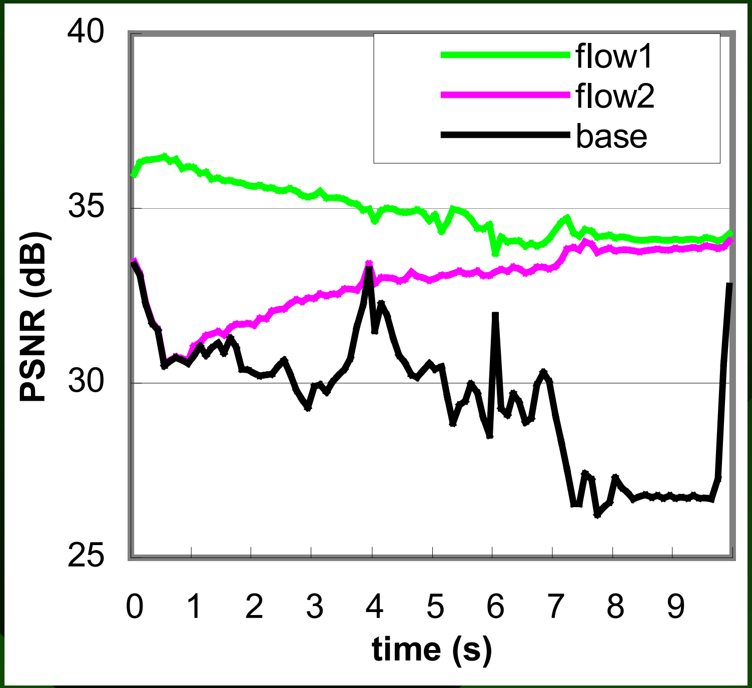
- Two Kelly flows are sharing the same bottleneck link *C* under identical delay



Experimental Results (3)

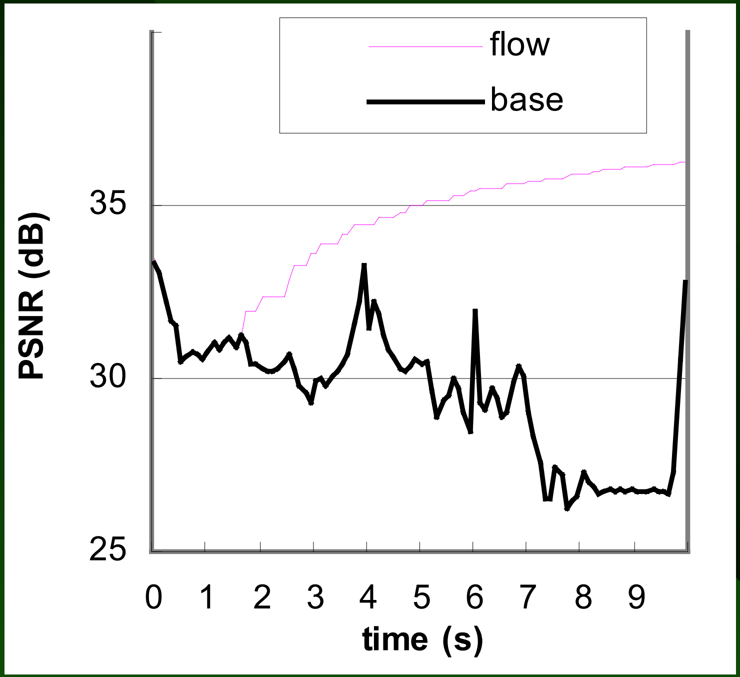
- Examine the effect of different round-trip delays (fixed and random) on fairness

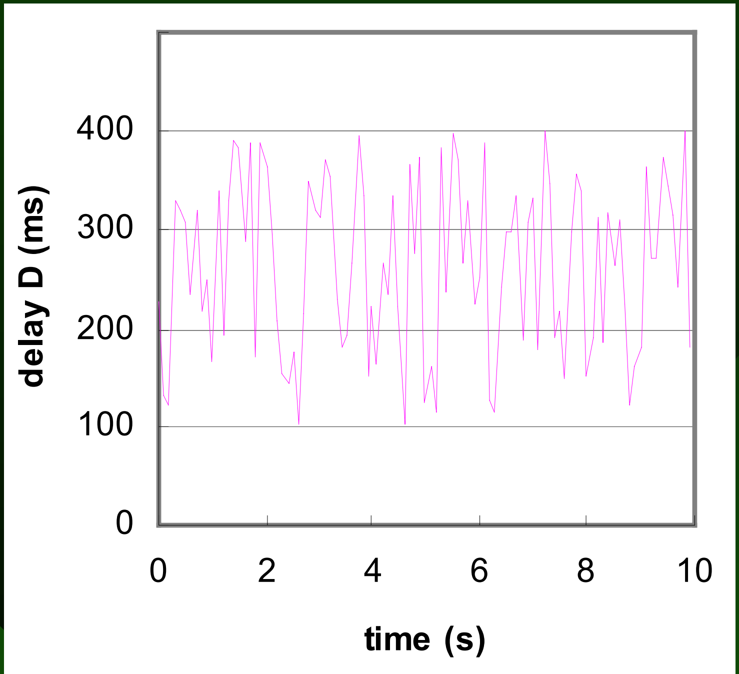




Experimental Results (4)

- Examine the situation *n* flows sharing the bottleneck bandwidth and each flow has a random delay





A single-flow PSNR when *n* = 10 flows share a 10 mb/s bottleneck link (left), Random delay for the flow (right)

Conclusion to this Work

- This paper derives a simple but accurate operational R-D model based on the properties of FGS encoders
- Based on this R-D model, we show a simple algorithm that can achieve better constant quality in CBR for scalable streaming than many other CQ methods
- Another contribution of this work is the successful combination of our R-D model with Kelly's congestion control in Internet streaming

Thank You!