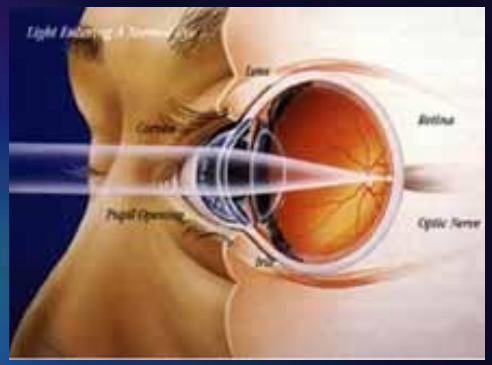
Foveated Motion Estimation for H.263 Video Encoder

Serene Banerjee

serene@ece.utexas.edu

Embedded Signal Processing Laboratory
Dept. of Electrical and Computer Engineering
The University of Texas at Austin
December 6th, 2000

Objective

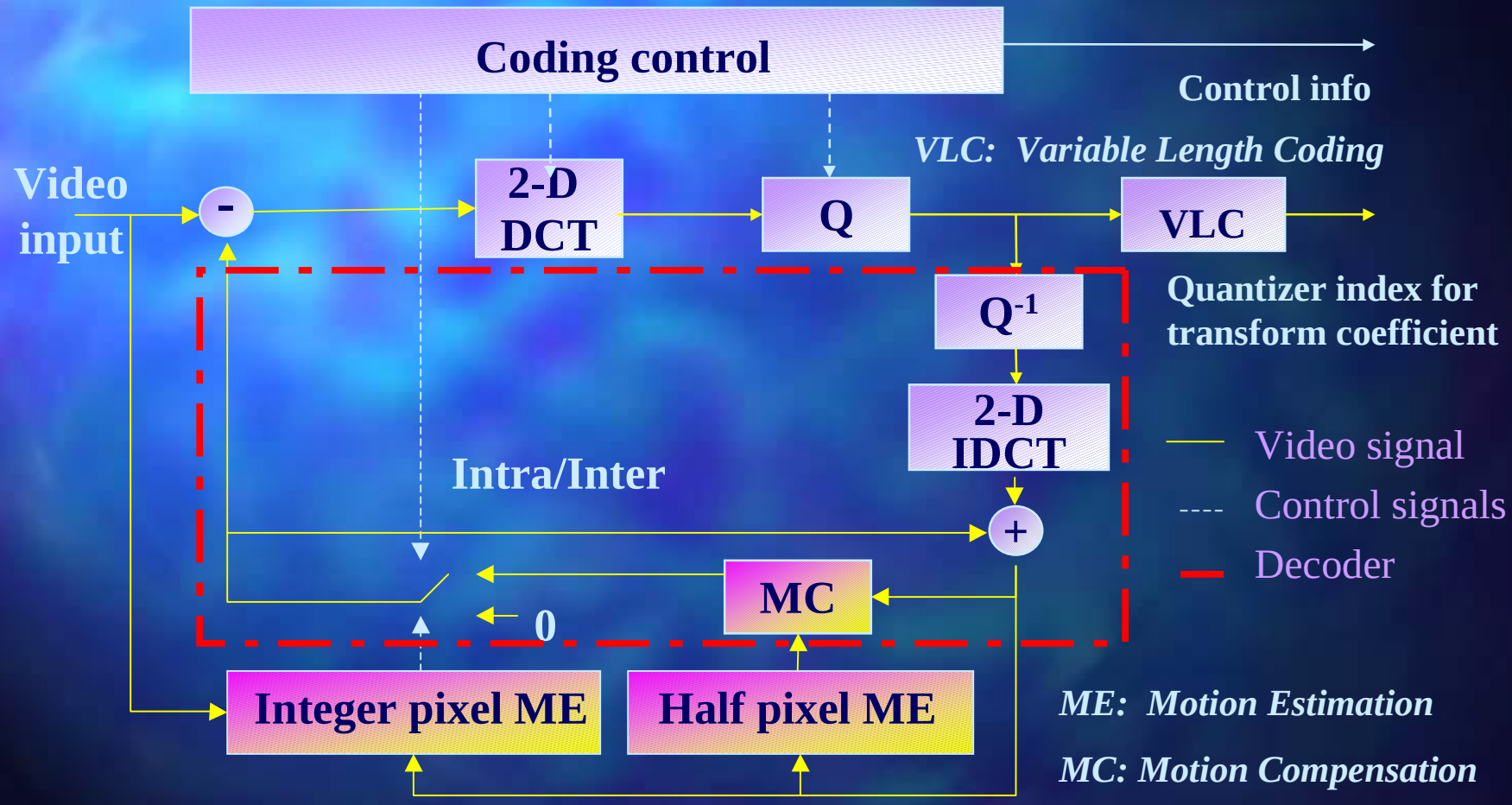


- Reduce bit-rate of H.263 video via foveation
  - Human visual system (HVS) samples non-uniformly
    - Spatial resolution reduces by half at about 2.3 degrees
  - Point of fixation is known as *foveation point*
    - Estimated from models
    - Obtained from eye trackers
- Applications
  - Image compression
  - Video compression
  - Thinwire visualization
    - Servers having large databases, *fovea-first* transmission

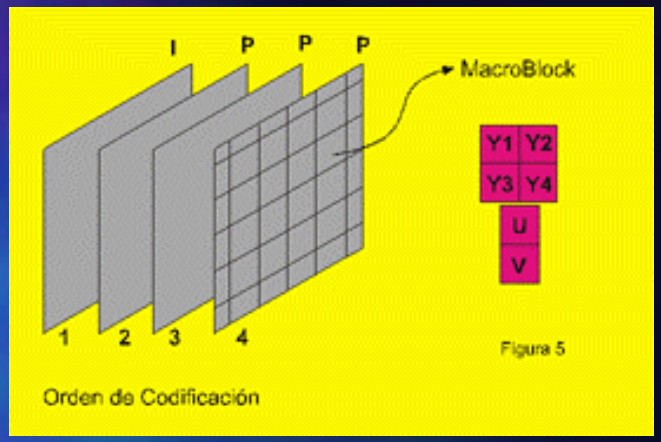
Foveation Methods

- Preprocessing [Chang and Yap, 1997; Lee and Bovik, 1999]
- DCT domain [Tsumura, Endo, Miyake, 1996; Sheikh, Liu, Evans, Bovik, 2000]
- Motion vector [Lee and Bovik, 1999; Bonmassar and Schwartz, 2000]

H.263 Video Encoder



H.263 Encoding



- Intra frame
  - Use DCT and quantization to reduce spatial redundancy
  - Manipulate quantization coefficients for foveation
- Predicted frame
  - Integer motion estimation via sum of absolute differences
  - Motion compensation to get predicted macroblock
  - Computation of difference between current and predicted block
    - Large difference: Intra code that block
    - Small difference:
      - DCT, quantize and foveate prediction error
      - Send information about motion vectors

Motion Foveation

- Bits allocated for motion vectors is about 5%
  - Cannot be used independently to get comparable results
  - Good for rate-distortion optimization
- Blur motion with chirp transform [Bonmassar and Schwartz, 2000]
- Problem: Changing motion vectors increases prediction error and blocks are intra coded increasing bit rate
- Solution:
  - Away from fovea, allocate less bits for motion vectors
  - Use same motion vector for entire row of macroblocks for better compression using variable length encoding

Temporal Foveation

- Update background less often [Robinson and Reeves, 1996]
- H.263: Randomly chosen macroblocks intra coded
- Modification: Block update rate inversely proportional to distance from fovea
- Gain: 3% lower bit rate
- Progressive transmission

Results

- Methods
  - Method #1: Foveation as preprocessing
  - Method #2: DCT domain foveation
  - Method #3: Foveation of motion vectors
- Sequence (60 frames): Mobile (293 kB), News (29.4 kB)

Sequence	#1	#1,3	#2	#2,3	#1,2	#1,2,3
Mobile	91.8	90.1	133	131	84.6	82.9
News	20.4	20.4	23.9	23.8	20.3	20.2

File sizes reported in kB

Quality Measures

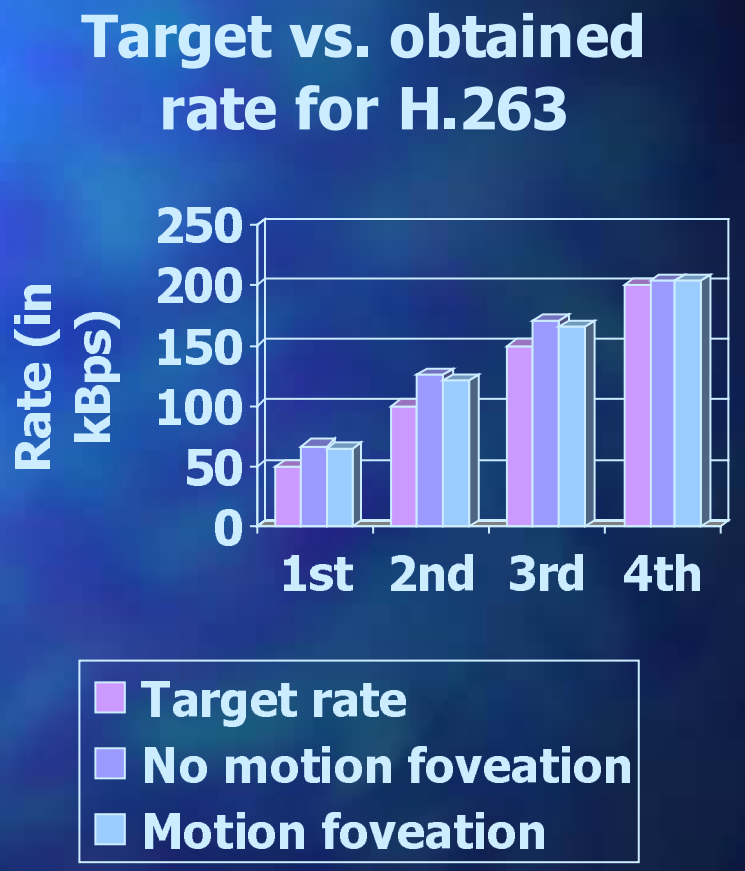
- PSNR: Peak signal to noise ratio and mean squared error (MSE) only good for additive noise
- Distortions: Frequency distortion and noise injection
- FPSNR: Based on weighted (foveated) MSE (FMSE)

$$FMSE = \frac{1}{\sum_{n=1}^{N} f_n^2} \sum_{n=1}^{N} [a(x_n) - b(x_n)]^2 f_n^2$$

$$FPSNR = 10 \log_{10} \frac{\max[a(x_n)]^2}{FMSE}$$

Rate-Distortion

- FPSNR over PSNR
  - PSNR
    - Foveate: 22.8 dB
    - Unfoveated: 27.8 dB
    - Difference 18 %
  - FPSNR
    - Foveate: 49.8 dB
    - Unfoveated: 51.5 dB
    - Difference 3.3 %
- Closer to target rate with motion foveation



Comparison [For mobile sequence]

Merit figures	#1	#2	#3
Computation	O(N^2)	O(1)	O(1)
Complexity	No encoder modification	Modify encoder	Modify encoder
Compression from foveation	68.7 %	54.6 %	+1 %
Quality	Good	Block artifact	Motion artifact

Conclusions

- Summary
  - Foveating motion vectors give an 1-2% lower bit rate
  - For same subjective quality, foveated video has lower bit rate
  - At same bit rate foveated video has higher subjective quality
  - FPSNR is a better image quality assessment criteria
- Future research
  - Appropriate model development for image quality assessment
  - Find correlation between motion vectors to remove artifacts
  - Object tracking using the foveal system as a magnifier