Technion - IIT
Dept. of Electrical Engineering
Signal and Image Processing lab

Transrating of Coded Video Signals via Optimized Requantization

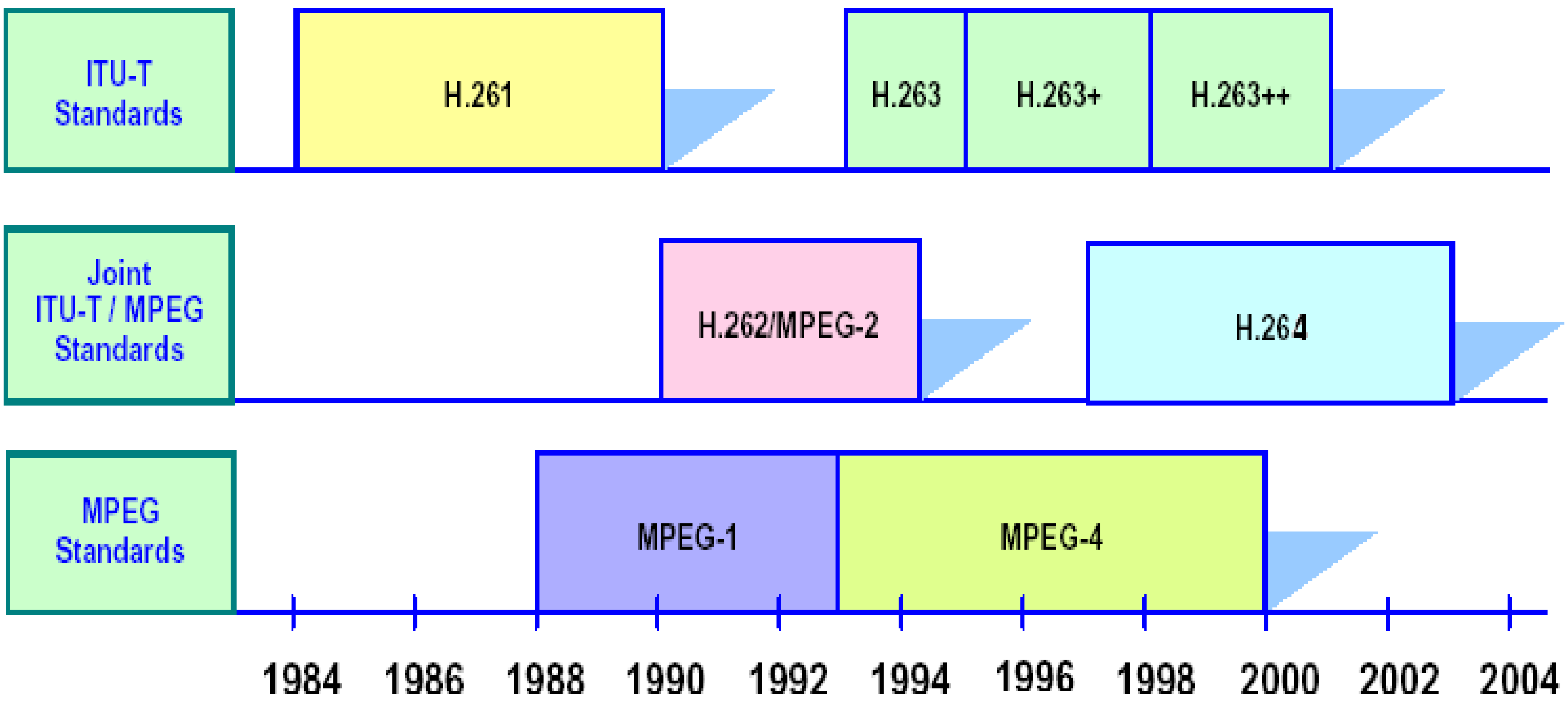
Michael Lavrentiev
supervisor : Prof. David Malah

March 2004

Outline

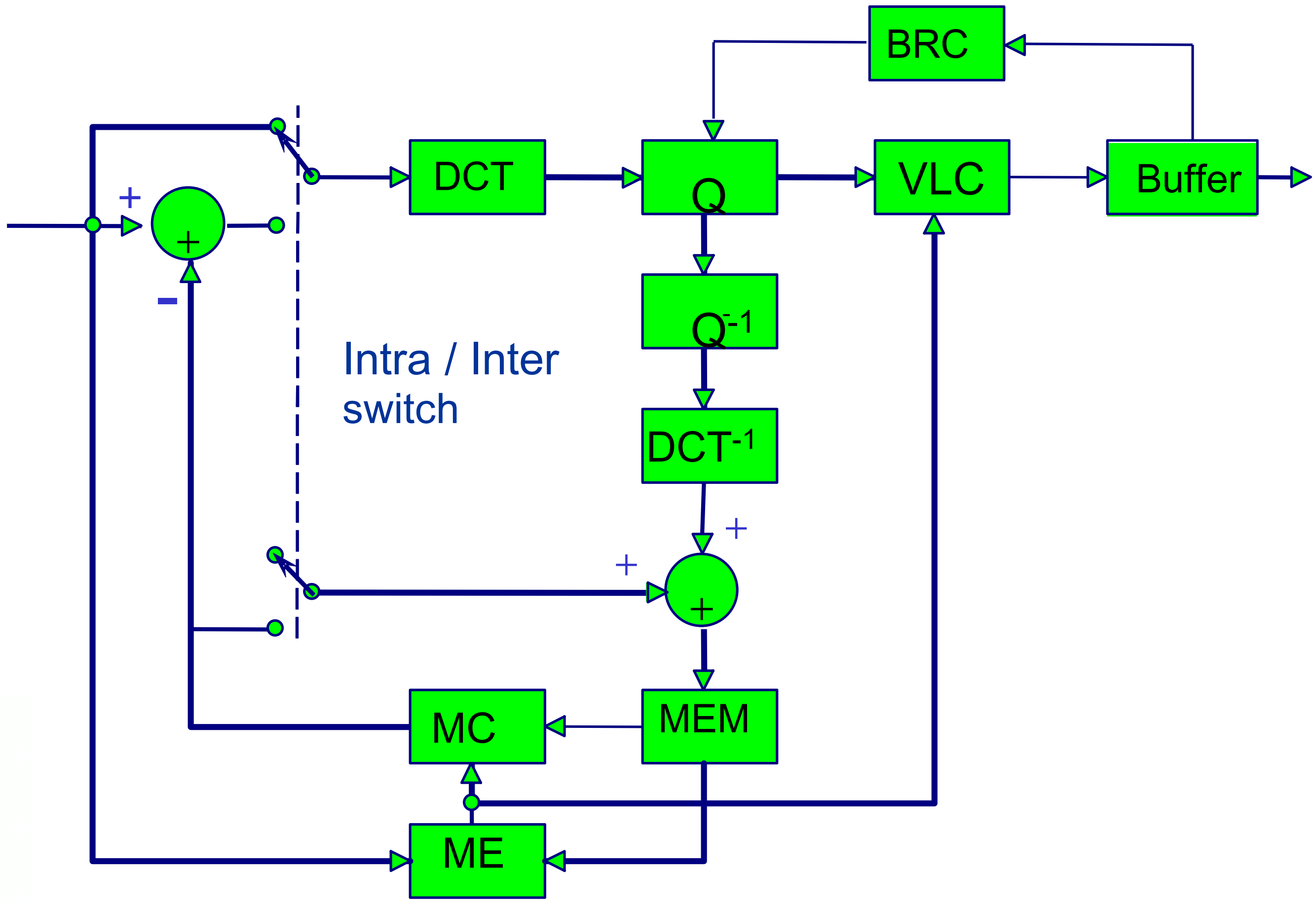
- Video encoding and transrating review
- Requantization methods:
  - “Simple” requantization
  - MAP/MSE requantization
  - Lagrangian optimization (+MAP/MSE)
  - Proposed Extended Lagrangian optimization (+MSE)
- HVS-based modification of requantization methods
  - Frame segmentation and tracking
  - Cost function modification
- Summary and future directions

Video Coding Standards

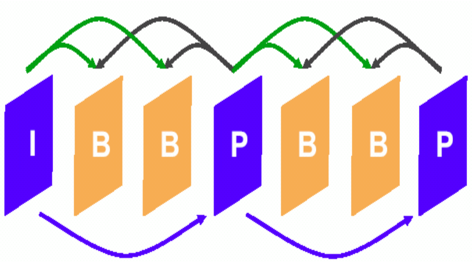


Generic Hybrid Video Encoder

Image sequence



MPEG GOP Structure



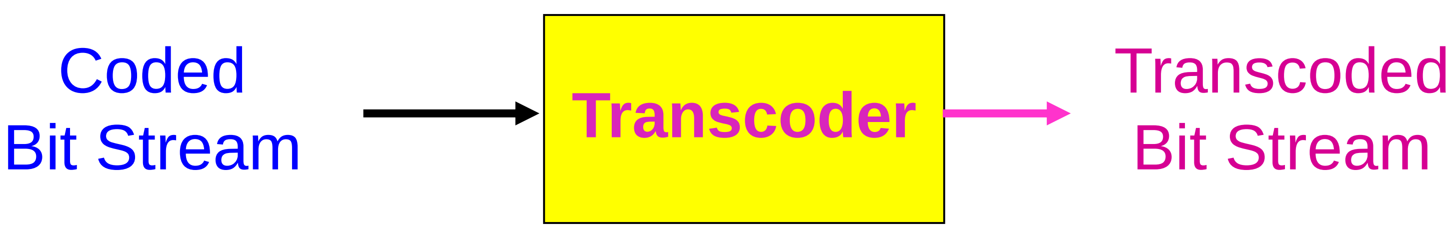
BRC - Bit Rate Control
MC - Motion Compensation
ME - Motion Estimation
MEM - Frame store
DCT - Discrete Cosine Transform
Q - Quantization
VLC - Variable Length Coding

Transcoding

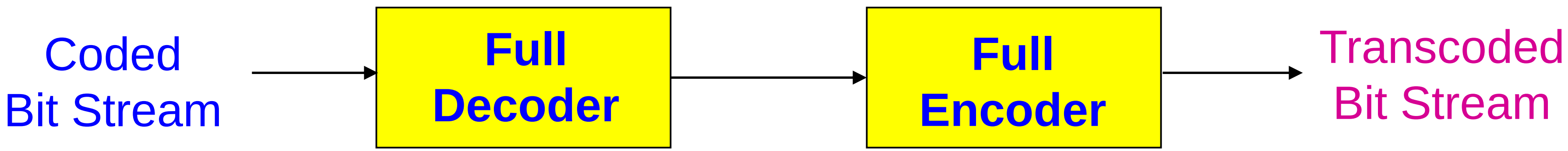
Converting a coded bit stream to another bitstream to match a destination profile in terms of:

- Media Formats {e.g., MPEG-2 ⟶ H.264 }
- Resolutions (Spatial, Temporal)
- Bit Rates (Transrating)
- Color depths/formats { e.g., 4:4:4 ⟶ 4:2:0}

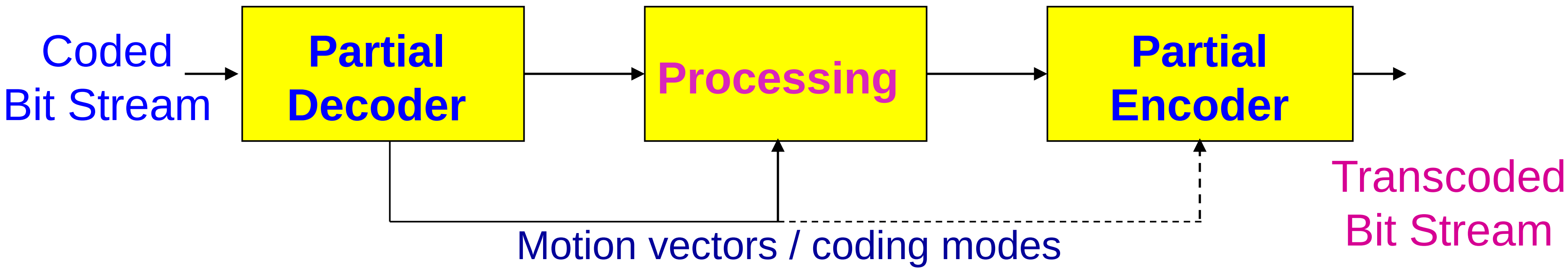
⋮



Transcoder Architectures



Cascading Decoder and Encoder (Re-encoding)



Complexity-Reduced Transcoding
(Compressed Domain)

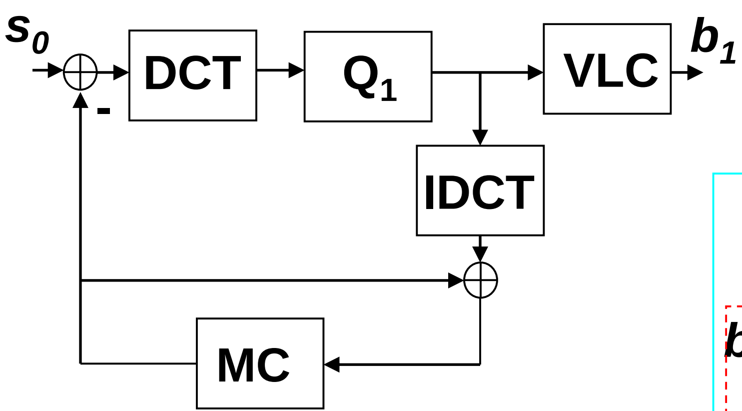
Transrating Methods

(Bit-Rate Reduction - BRR)

- Frame dropping (B, P)
- Color suppression
- Discarding high-frequency DCT coefficients
- Reducing spatial resolution (size reduction)
- DCT coefficients Requantization

Open and Closed Loop Schemes

(Keesman, 1996)



VLC = Variable Length Coding

DCT = Discrete Cosine Transform

IDCT = Inverse DCT

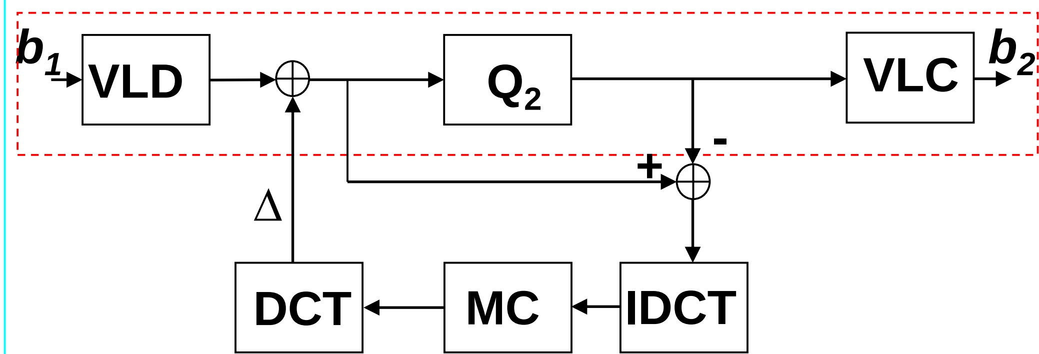
MC = Motion Compensation

Q_1 = first generation quantizer

Encoder

Transrater

Open-Loop



Drift Compensation

VLD = Variable Length Decoding

Q_2 = second generation quantizer

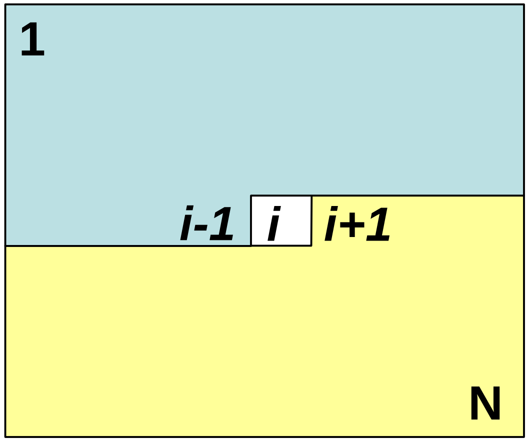
Δ = correction signal

“Simple” Transrating -1

- GOP-level BRC: $T_{Frame} = B_{in}^{Frame} \frac{R_{out}}{R_{in}}$
  - No buffer considerations
- MB complexity: $X_i \propto Q_i \cdot B_i$

  B_i – i-th MB bit-rate
- Frame-level BRC:
  - $Q_2^1 = \frac{1}{T_{Frame}} \sum_{i=1}^{N} X_i$
  - Rate-adaptive quantization:
    - Increase Q_2 if a < b
    - Decrease Q_2 if a > b
- One-path algorithm
- Not optimal in any sense

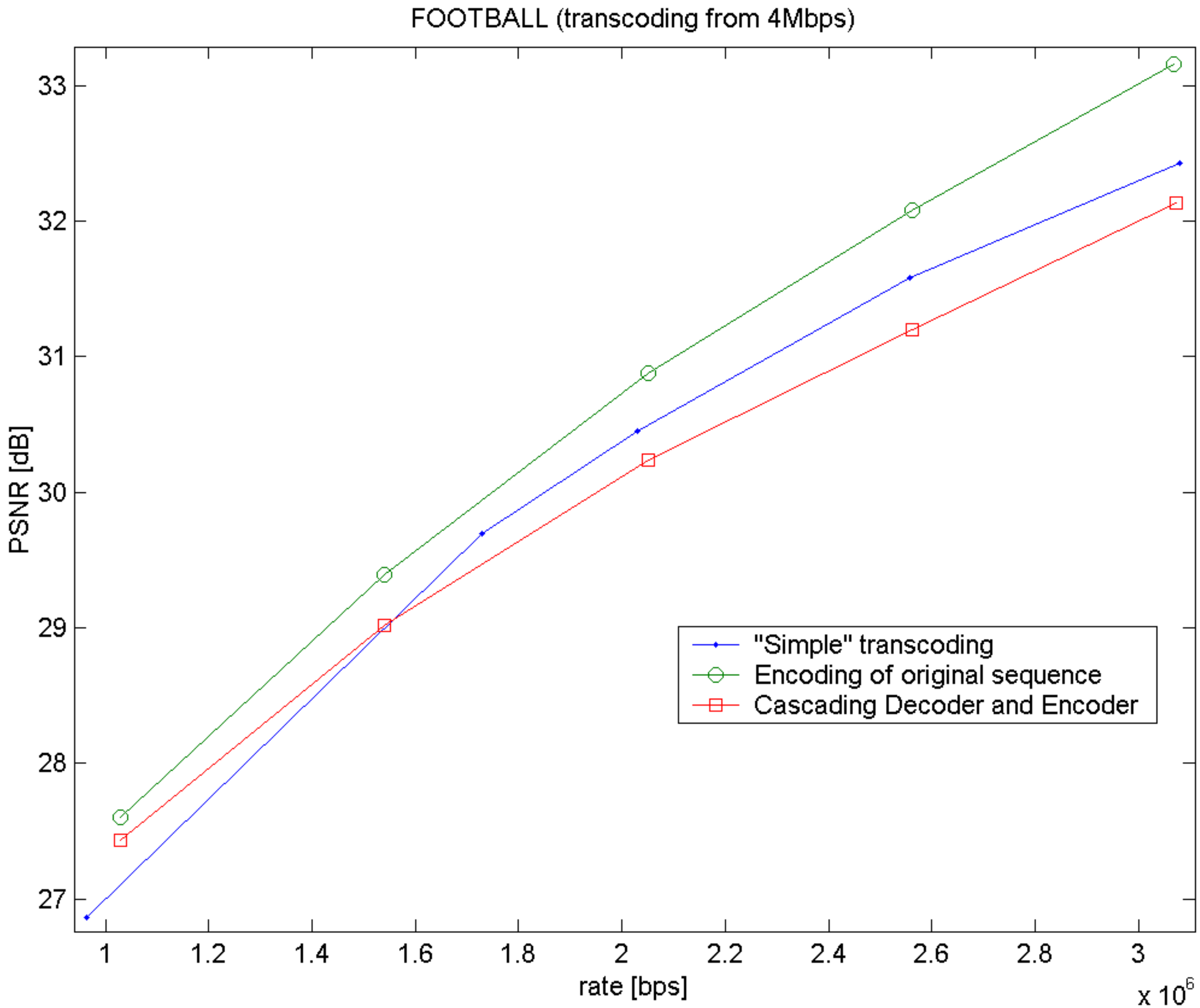
MB processing order:



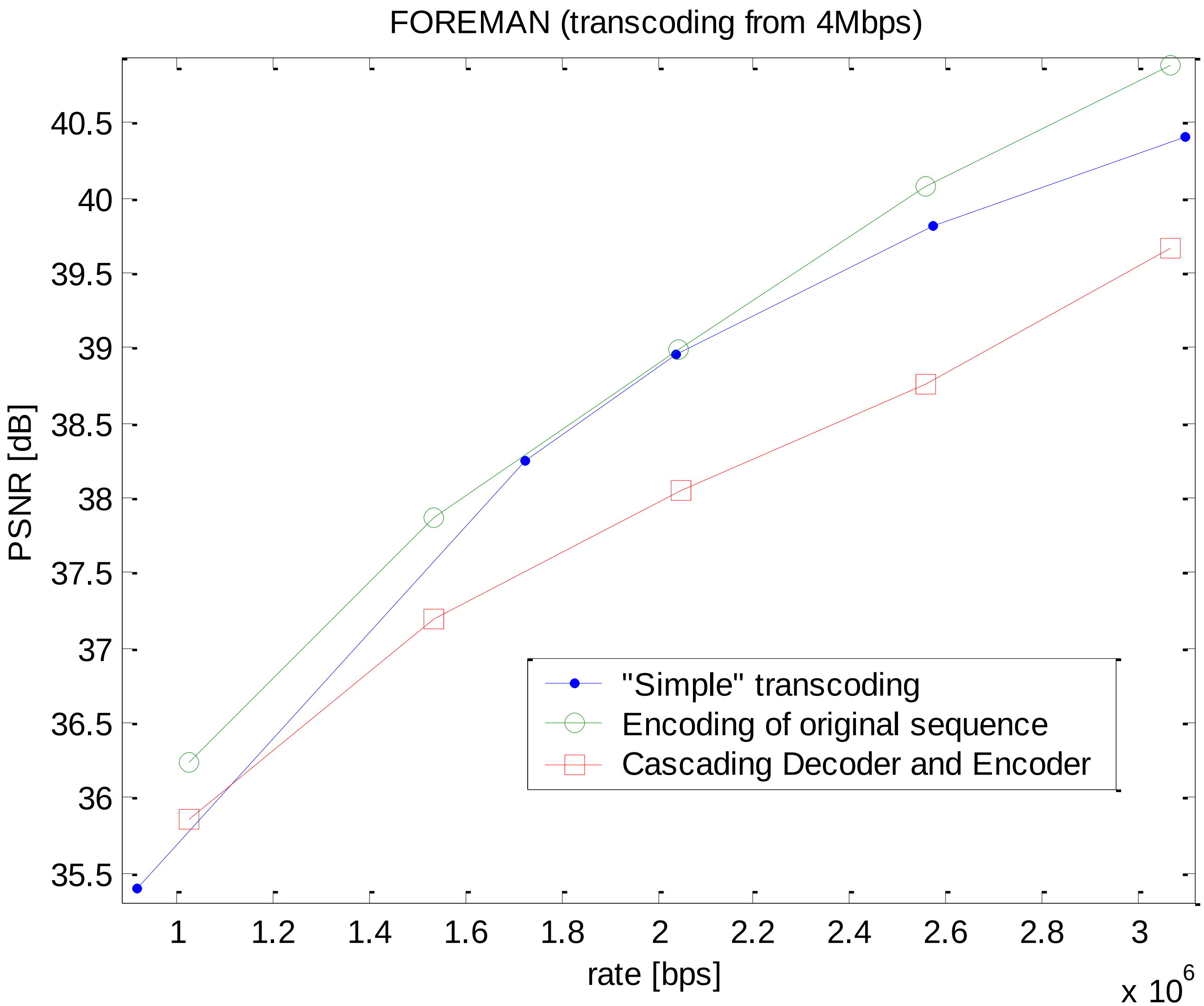
$$a = T_{Frame} - \sum_{j=1}^{i-1} B_j^{out}$$

$$b = \frac{B_{i-1}^{out}}{X_{i-1}} \sum_{j=i}^{N} X_j$$

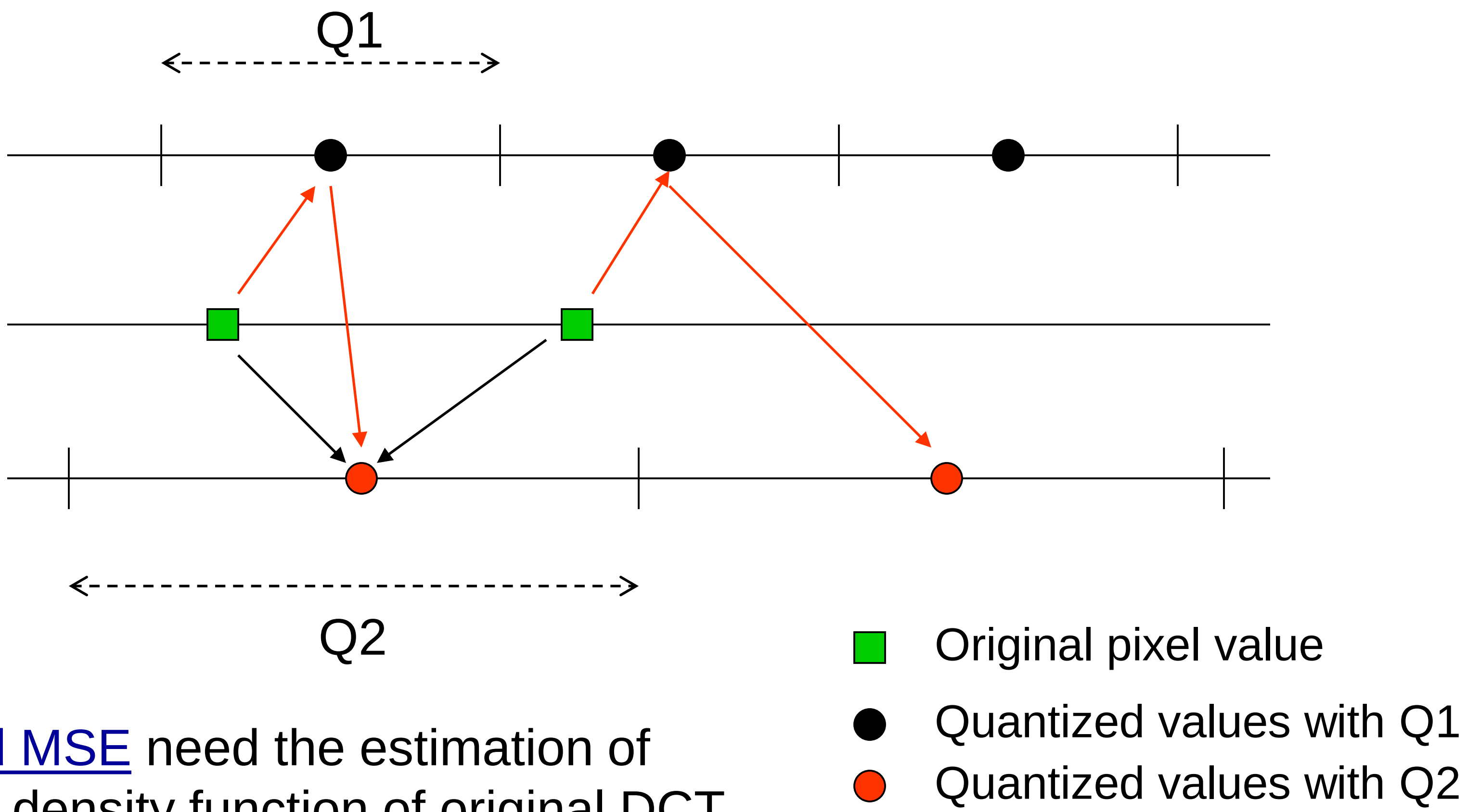
“Simple” Transrating -2



“Simple” Transrating -3

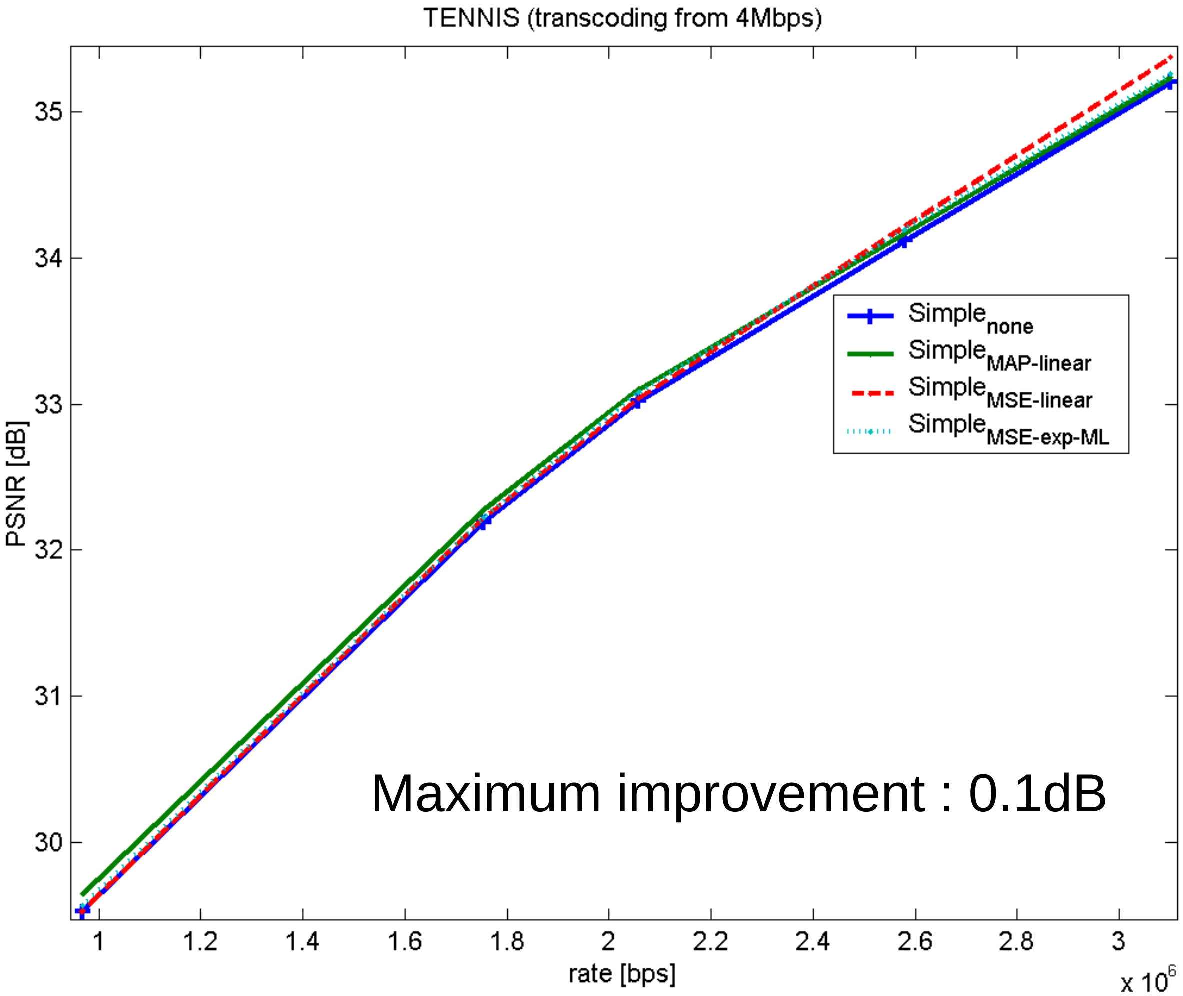


Requantization Error -1



• MAP and MSE need the estimation of probability density function of original DCT coefficients from quantized values received

Requantization Error -2



Requantization by Lagrangian Optimization -1

(Assunção and Ghanbari, 1997)

Constrained Minimization problem:

$$\text{Min } D, \text{ under the constraint } R \le R_T \;;$$

$$D = \sum_{k=1}^{N} d_k(q_k)\;; \qquad R = \sum_{k=1}^{N} r_k(q_k)$$

N – number of MBs in picture; d_k – distortion introduced into k-th MB

q_k – quantization step for k-th MB; r_k – rate of k-th MB after transcoding

Rate and Distortion are merged using Lagrangian parameter, $\lambda \ge 0$, :

$$J = D + \lambda R$$

Requantization by Lagrangian Optimization -2

(Assunção and Ghanbari, 1997)

Lagrangian cost function becomes sum of independent MB level calculated parts:

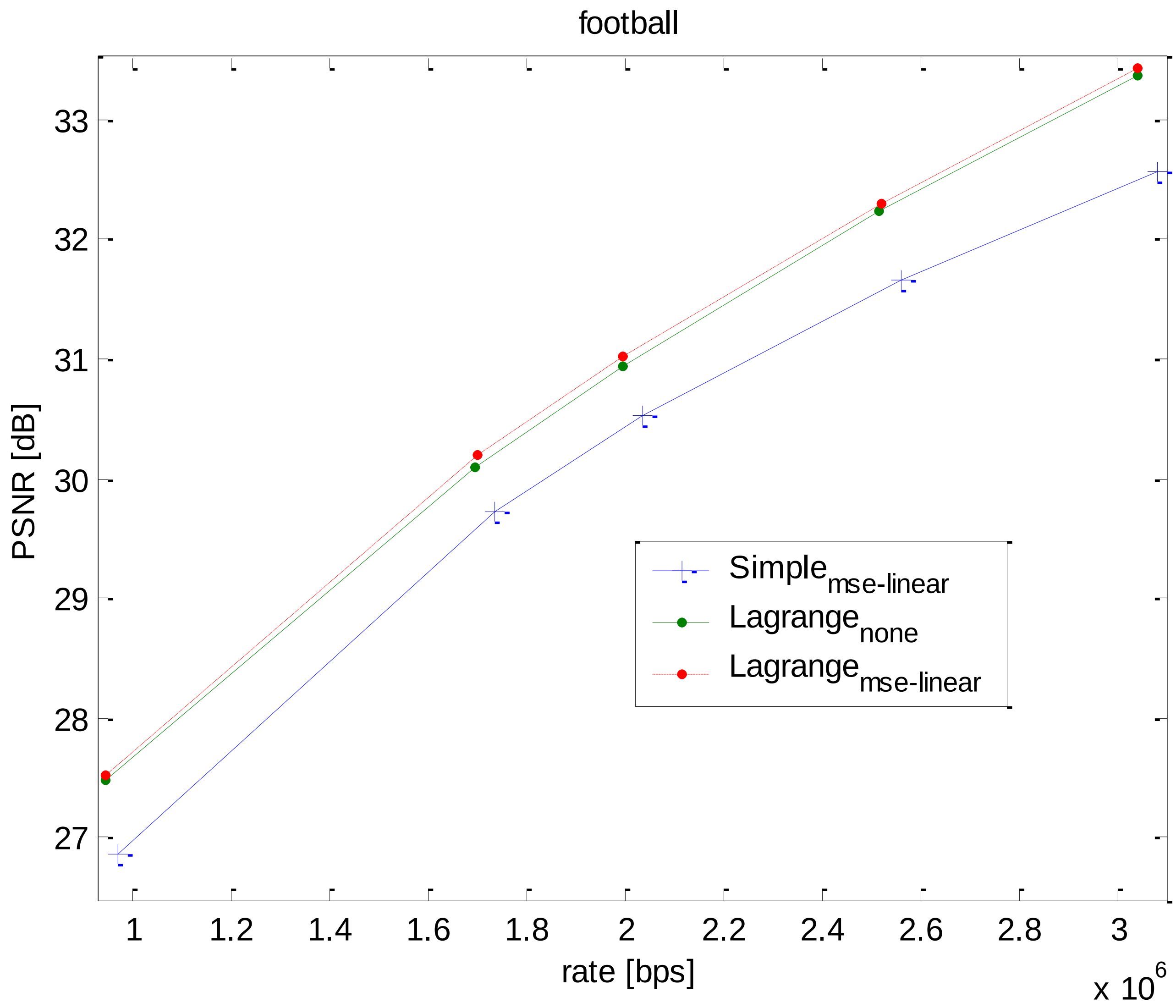
$$J_k(\lambda) = \min_{q_k}\{d_k(q_k) + \lambda r_k(q_k)\}$$

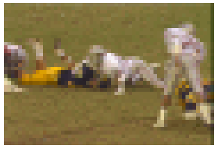
Lagrangian parameter, $\lambda \geq 0$, is iteratively updated to achieve desired bit-rate R_T:

1. Calculate all $\mathbf{J_k(\lambda)}$ for all **k**
2. Compute total rate $R_{total}=\Sigma r_k(q_k)$ and check:
   - If $R_{total} = R_T$, transmit transrated frame and go to next frame
   - If $R_{total} < R_T$ => decrease λ ; If $R_{total} > R_T$ => increase λ
3. Goto 1 with new λ

- Recent result (Leventer and Porat – ICIP-03): Optimal quantizer steps are obtained at values close to specific multiples of the input quant. step.

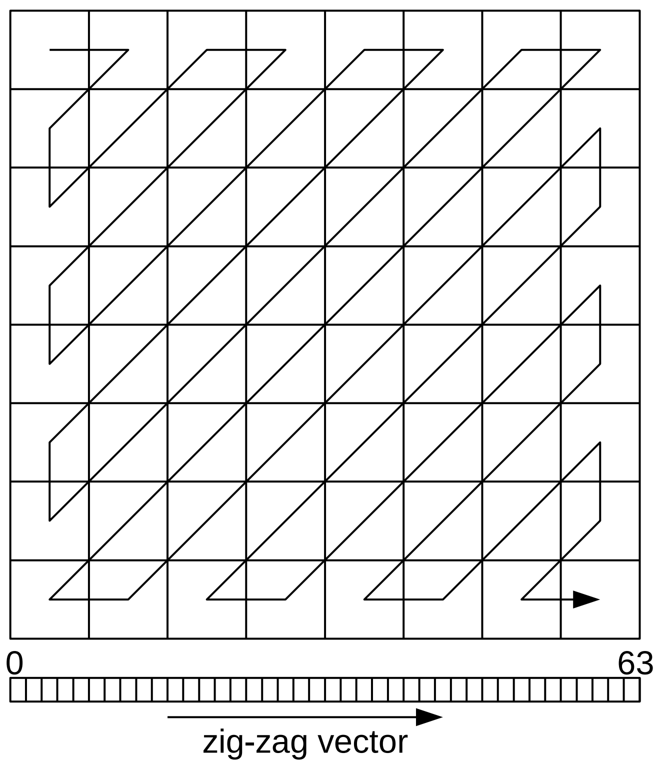
Requantization by Lagrangian Optimization -3



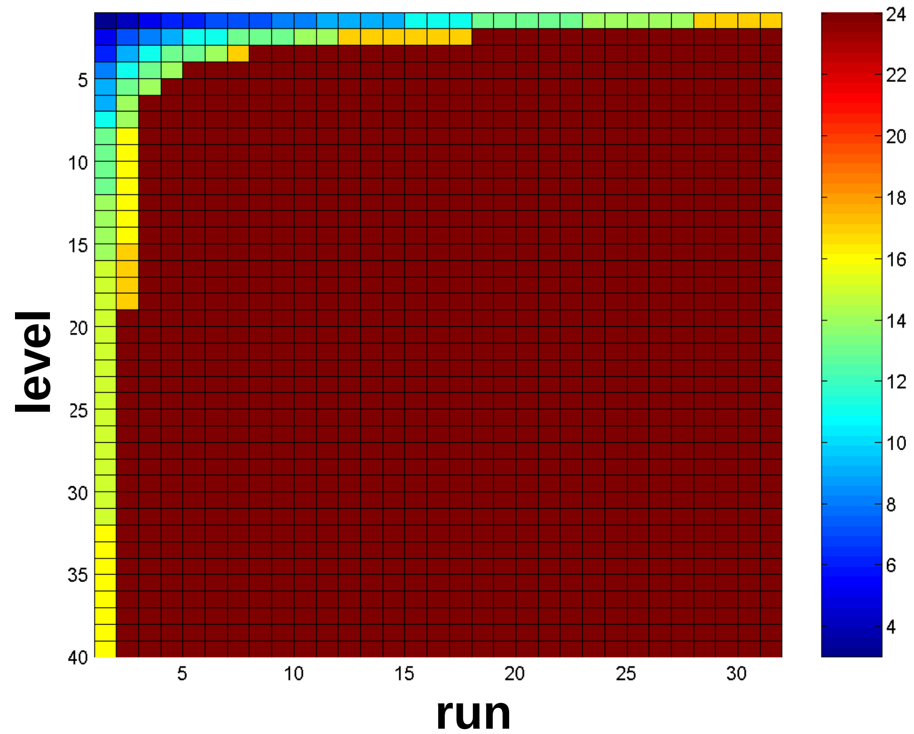


Variable Length Coding

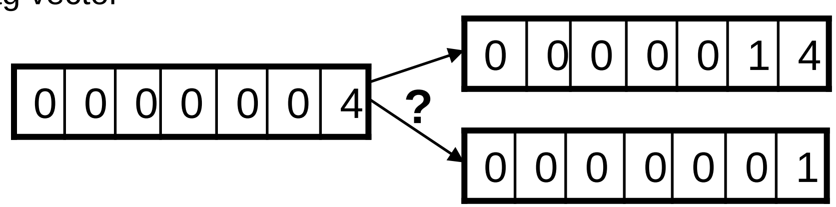
Zig-zag scan of quantized DCT coefficient indices in block



VLC table



rate = 24 ; D = 0



rate = 13 ; D = 1

rate = 8 ; D = 9

Extended Lagrangian Optimization

We propose to extend Lagrangian optimization by the modification of quantized coefficients index values :

Min D, under the constraint $R \leq R_T$;

$$D = \sum_{k=1}^{N} d_k(q_k, \mathrm{v}) \; ; \qquad R = \sum_{k=1}^{N} r_k(\mathrm{v})$$

$N -$ number of MBs in picture; $d_k -$ distortion introduced into k-th MB

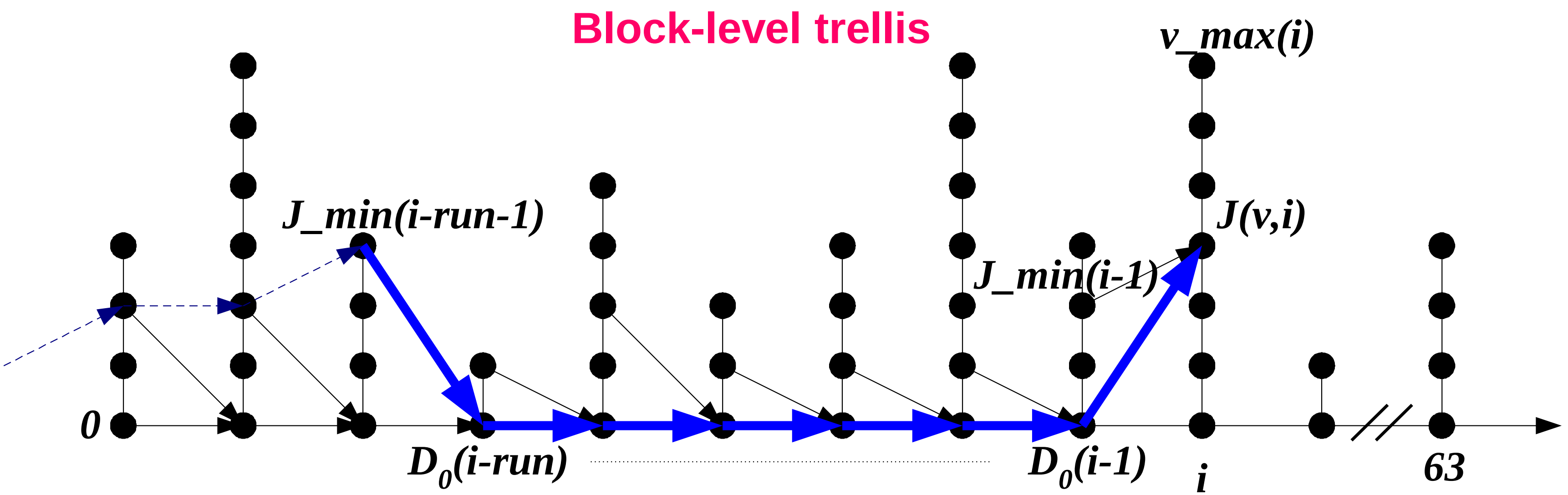
$q_k -$ quantization step for k-th MB; $r_k -$ rate of k-th MB after transcoding

v - vector of quantized DCT coefficients indices

Lagrangian cost function becomes sum of independent MB level calculated parts:

$$J_k(\lambda) = \min_{q_k, \mathrm{v}} \{ d_k(q_k, \mathrm{v}) + \lambda r_k(\mathrm{v}) \}$$

Trellis-based Re-quantization



$$J(v,i) = \min_{run} \left\{ J\_\min(i-run-1) + \sum_{j=i-run}^{i-1} D_0(j) + \lambda R(run,v) + D(v,i) \right\}$$

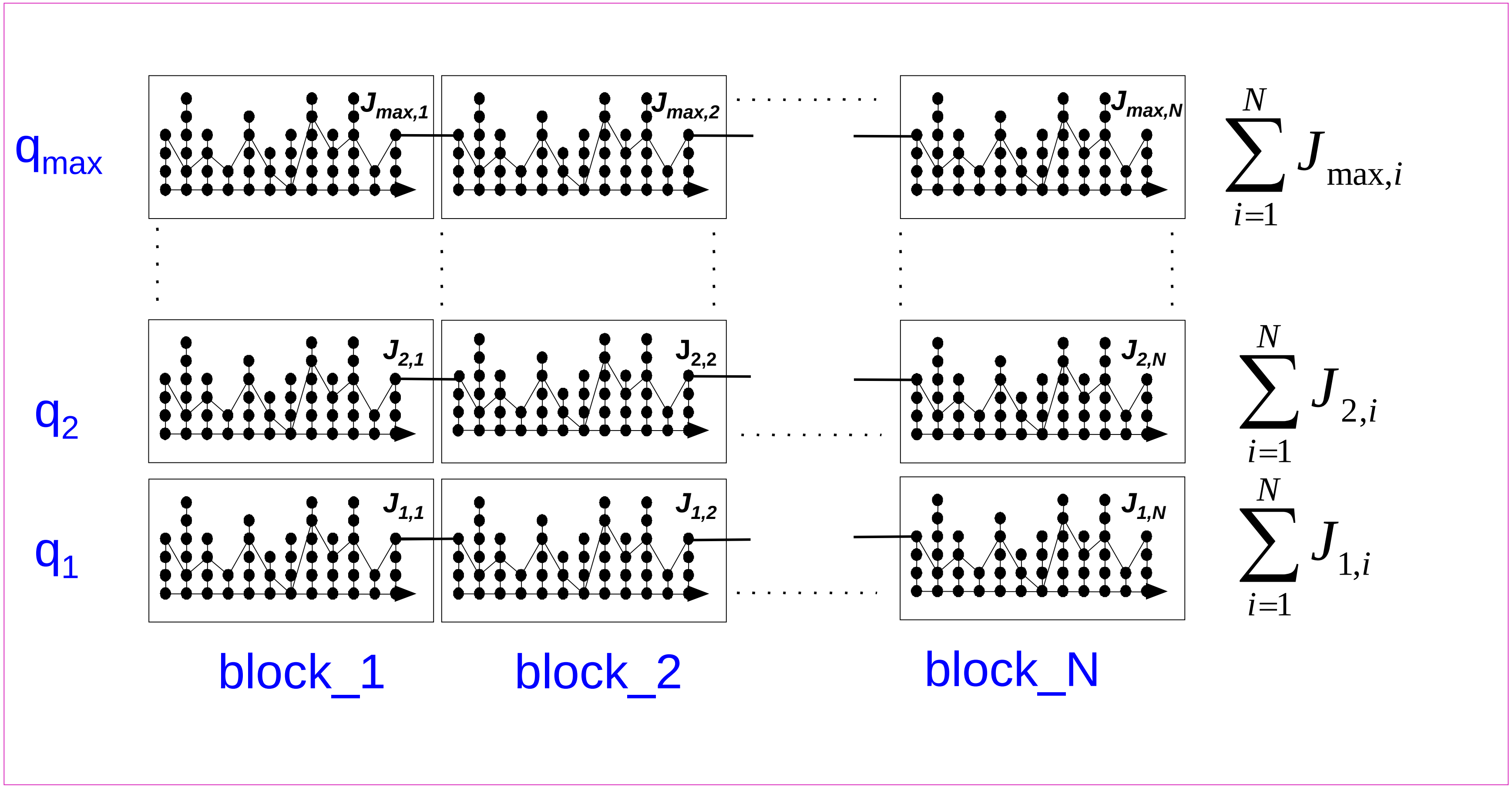
$$J\_\min(i) = \min_{v} J(v,i)$$

Complexity reduction:

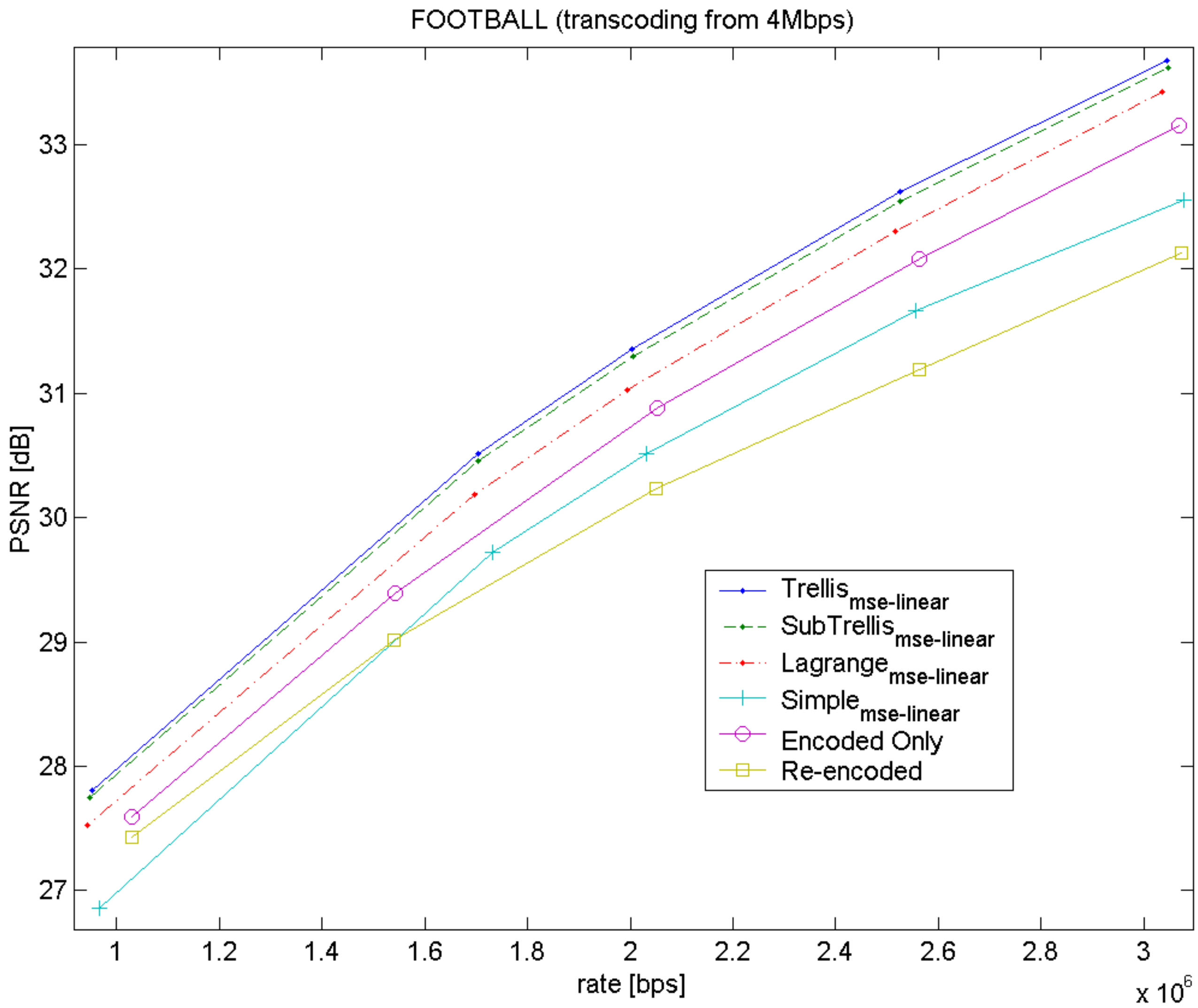
- When *R(run,v)* reaches the maximum, *J(v,i)* is determined
- Sub-optimal: run-level pair splitting is forbidden

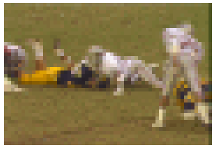
Trellis-based Re-quantization

Macro-Block

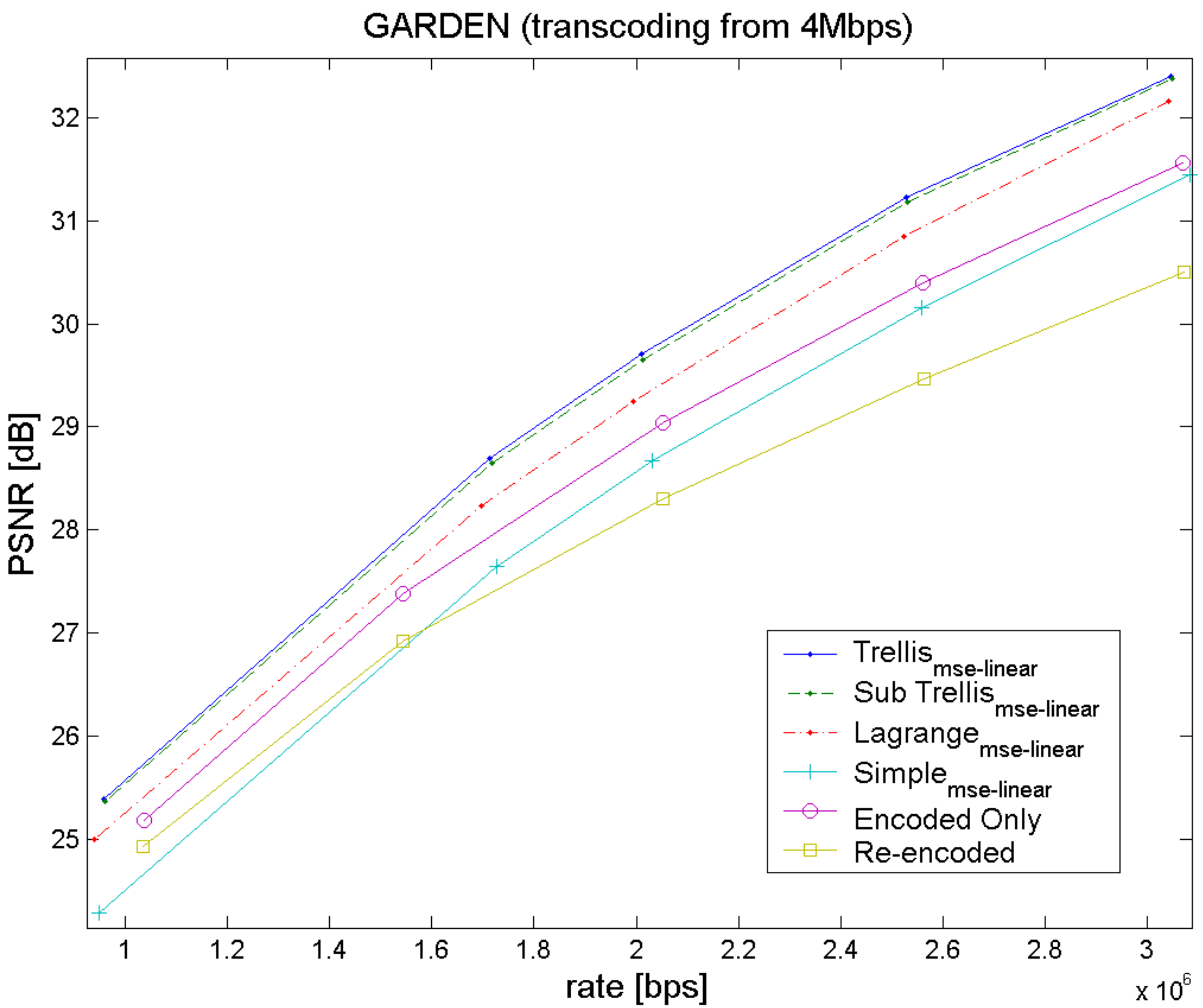


Experimental Results





Experimental Results



Experimental Results

FOOTBALL sequence from 4Mbps

method	Run-time ratio	PSNR Improvement*
“Simple”	0.15	-0.4dB
Run-time optimized Lagrangian	1	0.3dB
Extended Lagrangian	7.9	0.7dB
Sub-optimal Extended Lagrangian	3.3	0.6dB

*** Relative to Source Encoding**

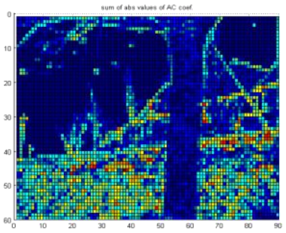
HVS-based Bit Allocation - 1

- Video frames can be partitioned (segmented) into regions having different characteristics:
  - Textured regions
  - Smooth regions
  - Edges
  - Moving objects
- Different regions have different perceptual importance and may be encoded at different rates for similar subjective quality.
- For efficiency, segmentation should be based on compressed-domain data: Block DCT Coeff., MB type, MB q-step, MB rate, and Motion Vectors (MV).
- Bit allocation obtained reflects the perceptual importance.

HVS-based Bit Allocation -2

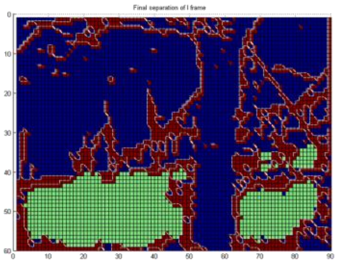
- Segmentation of I-frames:

  - Sum of AC coefficients absolute values as indication of block activity



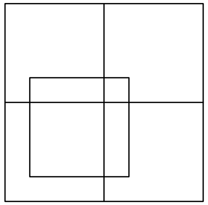
  - Binarization by Frame-adaptive thresholding (Otsu 1979)

  - Partition into Textured, Smooth and Boundary regions using morphological operators



HVS-based Bit Allocation -3

- Segmentation Tracking (P, B frames) :
  - Reference blocks are found using MVs
  - Weighted sum (proportional to each reference block prediction area) is assigned to predicted block

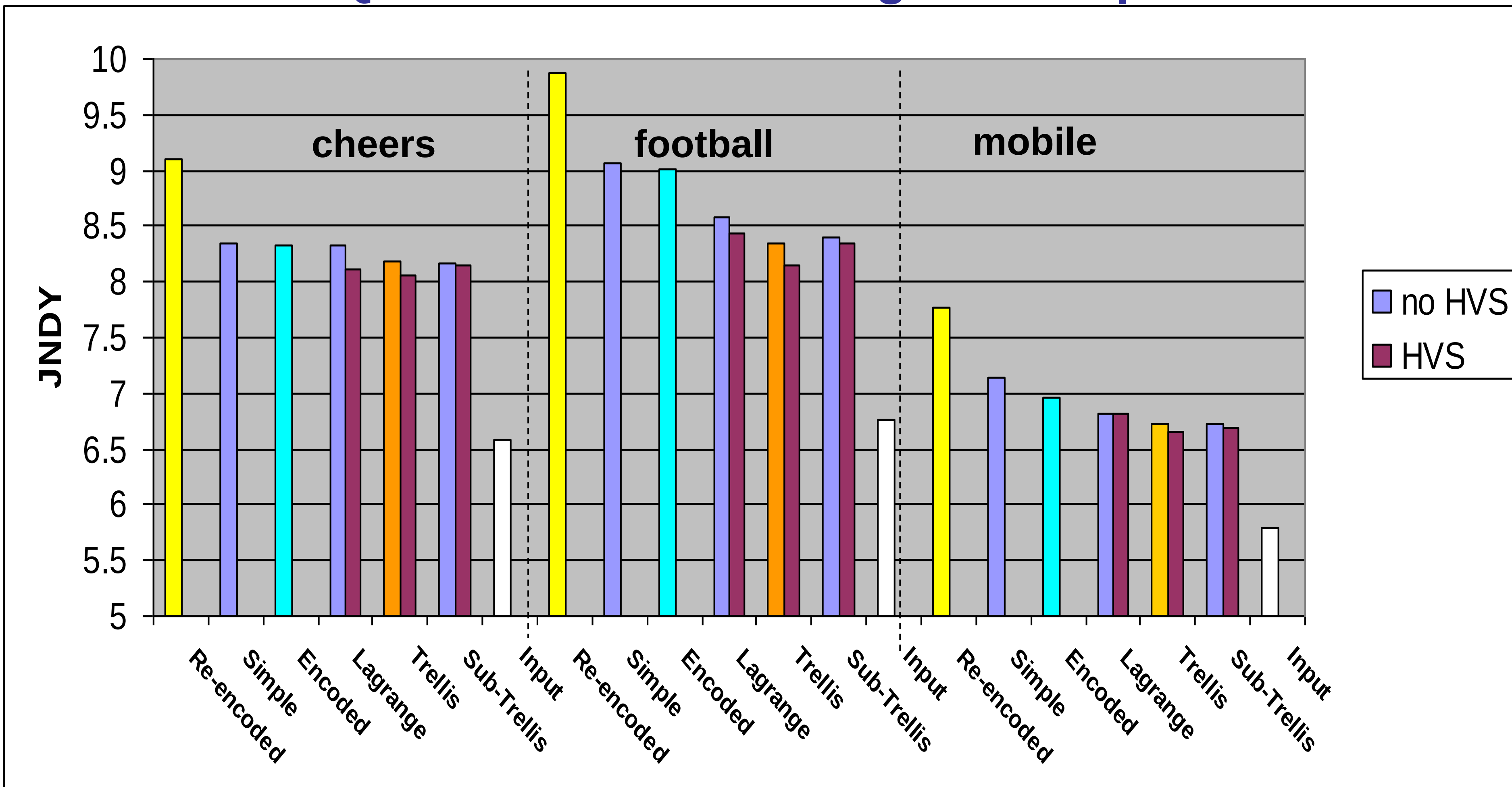


  - Segmentation is also applied to difference pictures and used if provides new information about textures and boundaries

- Perceptual Weighting:
  - Squared Error for each block is weighted according to its type and motion
  - Lagrangian optimization is applied with modified Distortion function

HVS-based Bit Allocation -4

PQA results: Transrating 6->4 Mbps



Summary

- An Extended Lagrangian Optimization requantization is proposed and implemented using a trellis-based scheme. Its PSNR is always better than PSNR of original sequence encoding to the same rate
- Utilization of MAP and MSE methods with Lagrangian optimization is examined
- Developed an approach for integrating HVS-based models into the optimization procedure is proposed and results in improved performance

Future directions

- Smart GOP-level BRC can further improve the performance
- Other methods like frame-rate reduction, resolution reduction and frame cropping can be combined with proposed requantization methods
- HVS-based distortion weight-allocation can be further improved

END