Towards Model-based Transrating of H.264 Coded Video

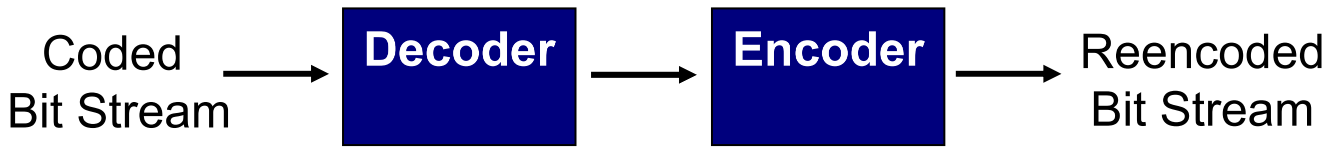
Naama Hait and David Malah
Department of Electrical Engineering, Technion

Presentation Outline

- Introduction
- Optimal requantization in H.264
- ρ domain rate-distortion modeling
- Proposed algorithm summary
- Results
- Conclusion

Transrating of H.264 coded video

- H.264 (a.k.a. MPEG-4/AVC): State of the art standard coder
- Transrating: Bit rate reduction
- Trivial solution – reencoding (high computational complexity):



- Compressed-domain transrating

doesn't perform full decoding and full encoding.

Transrating via optimal transform coefficients requantization

$$\min_{\{QP_i\}} \quad D, \quad subject \quad to \quad R \le R_{target}$$

where $D = \sum_{i=1}^{N_B} d_i(QP_i)$ and $R = \sum_{i=1}^{N_B} r_i(QP_i)$

N_B - number of blocks in the frame

QP_i - quantization parameter for the i-th block

d_i - requantization distortion

r_i - requantization rate

- Common Lagrangian optimization approach:
  - Merge rate and distortion using Lagrangian parameter, $\lambda \geq 0$

    $$J = D + \lambda(R - R_{target})$$ (Assunção and Ghanbari, 1997)
  - Decompose cost into sum of independent costs for each block

Optimal requantization in H.264

- ΔQP limitation in H.264: $|QP_{i+1} - QP_i| \leq 2$
  introduces dependencies in $\{QP_i\}$ choices.
  → Common Lagrangian optimization approach cannot be applied.

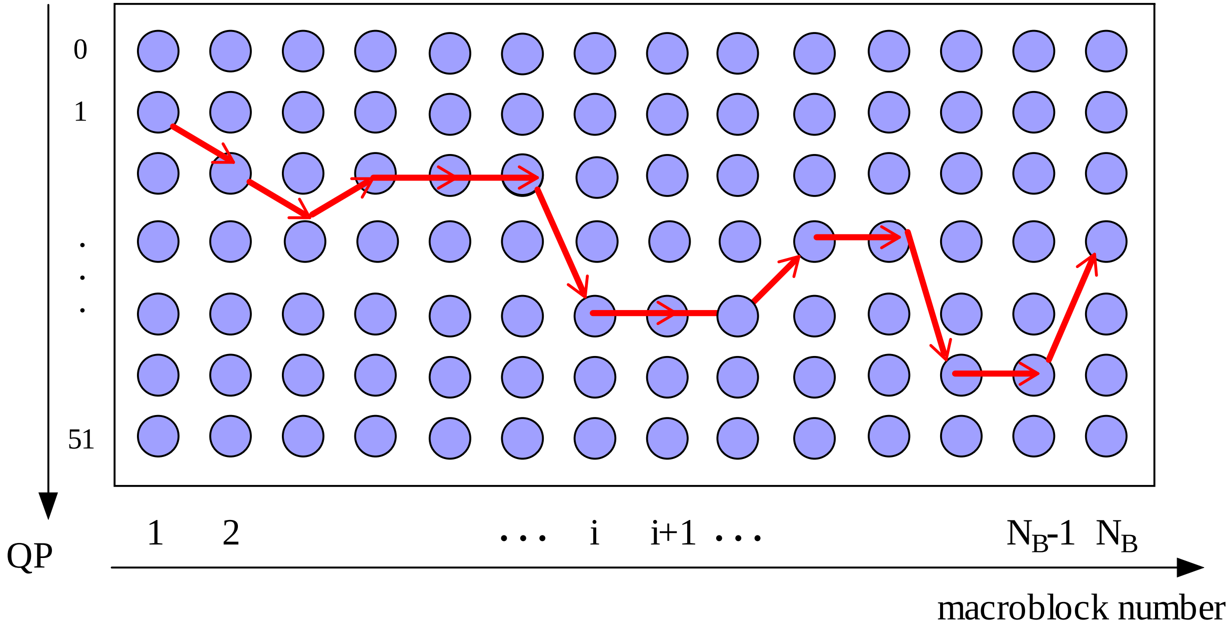
- Suggested solution:
  Extend each Lagrangian iteration with a constrained dynamic programming stage.
  - At a given λ, find an optimal $\{QP_i^*\}$ path by solving a dynamic programming problem:

$$\min_{\{QP_i\}} J \quad subject \quad to \quad |\Delta QP| \leq 2$$

    where $J = D + \lambda(R - R_{target})$.
  - Calculate the corresponding rate: $\sum_{i=1}^{N_B} r_i(QP_i^*)$.
  - Change λ accordingly.

Dynamic programming

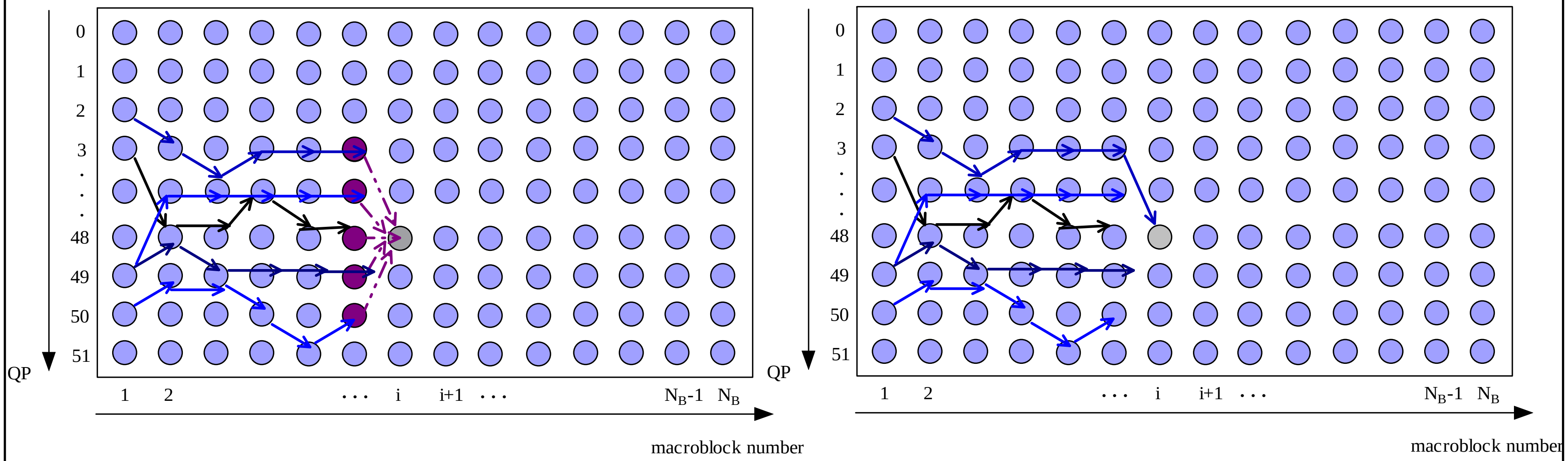
- The dynamic programming algorithm is defined over the set of states $\{(QP, i)\}$



- Each state (QP, i) has a r-d cost: $j_i(QP) = d_i(QP) + \lambda r_i(QP)$.
- The total frame cost along a path is: $J = \sum_{i=1}^{N_B} j_i(QP)$.

Dynamic programming stage update

- The value at state (QP, i) is $V_i(QP) = V_{i-1}(QP_{Prev}) + j_i(QP)$.

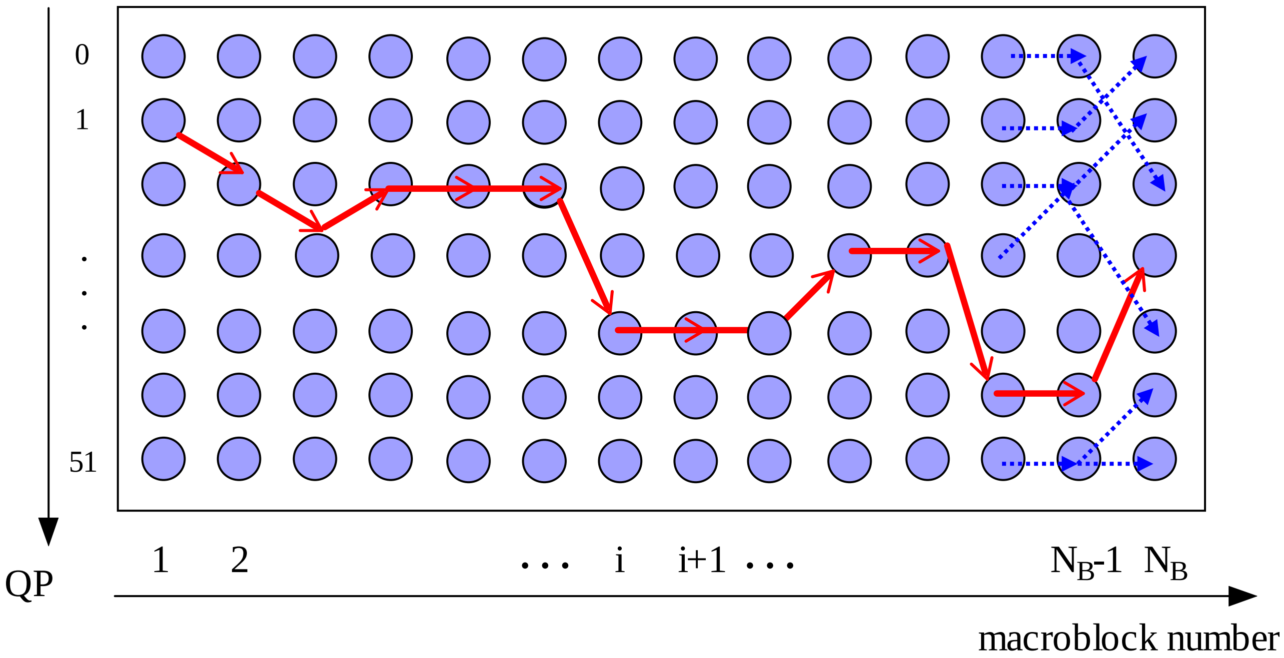


- We choose among the allowed previous states at block (# i-1) the optimal transition from $(QP^*_{Prev}, i-1)$.

Dynamic programming – last stage

- At the last stage ($i=N_B$), the optimal path over the entire frame is:

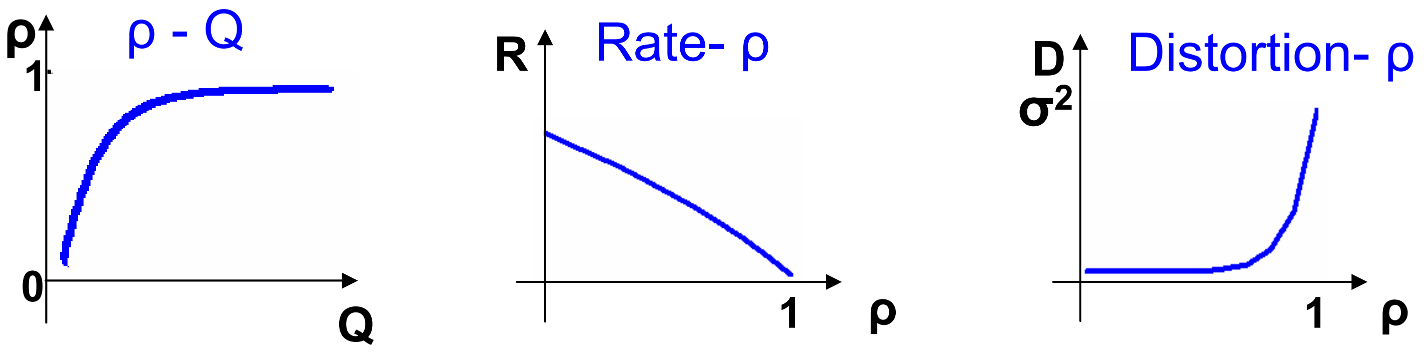
$$BestPathEnd = \operatorname*{argmin}_{QP} V_{N_B}(QP)$$



- Trace back to obtain $\{QP_i^*\}$.

Rate-distortion modeling

- The optimization requires evaluating $d_i(QP), r_i(QP)$ for all the $\{(QP,i)\}$ states → high computational complexity
- Model-based evaluation greatly reduces the computational load.
- Use r-d models in the ρ domain (fraction of zeroed coefficients):
  - More robust than models in the step-size domain
  - Defined over a finite support $0 \leq \rho \leq 1$



- Previous work (He and Mitra, 2002): linear $rate - \rho$ and exponential-linear $distortion - \rho$ models at the frame level.

Rate - ρ models for H.264

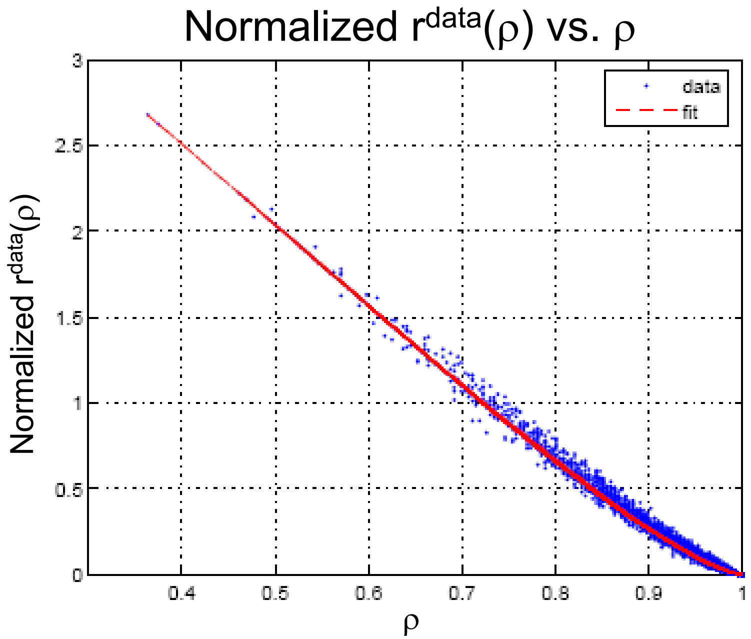
- H.264 context adaptive entropy coding with VLC tables (CAVLC):
  - Encodes 4 types of syntax elements:
    - Run
    - Level
    - Two additional counts that mainly describe zeroed coefficients distribution
  - Switches between VLC tables (context adaptive)
- The $rate - \rho$ relation at a block level in H.264 isn't linear.
- The suggested rate model decomposes the rate into two additive components:
  - "Data" (run-level)
  - "Overhead" (additional counts)

  such that

$$r(\rho) = r^{data}(\rho) + r^{overhead}(\rho)$$

Rate - ρ models for H.264 (cont'd)

- "Data" $rate - \rho$ relation: $r^{data}(\rho) = \theta \cdot ln(1 + (1 - \rho)^{\eta})$
  where $\theta \geq 0$ is the scale parameter and $\eta \geq 1$ is the shape parameter.



$\theta = 6.2$

$\eta = 1.36$

- "Overhead" $rate - \rho$ relation – uses probability based average code length tables

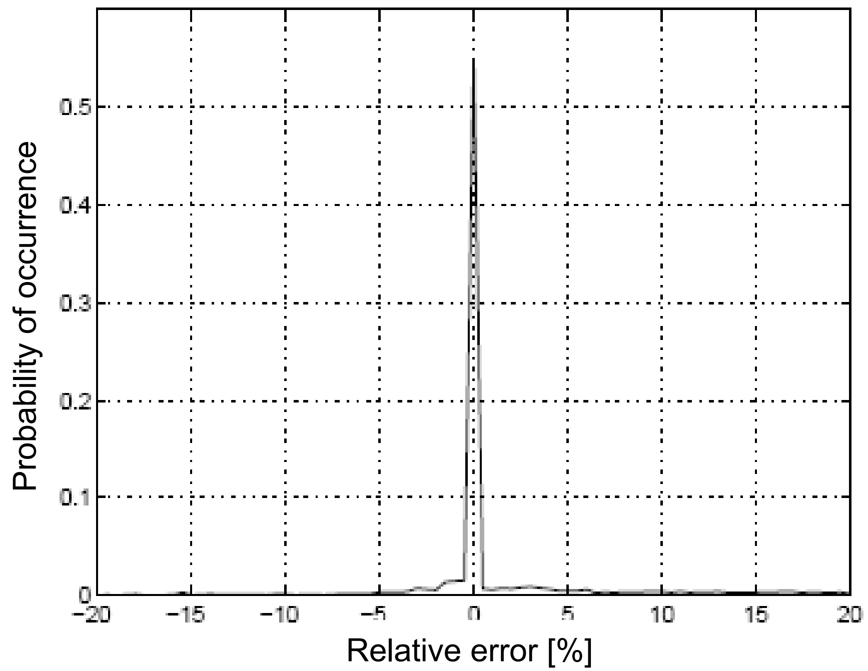
Distortion - ρ model

- The exponential-linear model isn't accurate enough at the block level
- The suggested exponential-quadratic model

$$d(\rho) = \sigma^2 \cdot exp\{\alpha_1 \cdot (1-\rho)^2 + \alpha_2 \cdot (1-\rho)\}$$

better suits the empirical data.

Distortion- ρ model relative error distribution

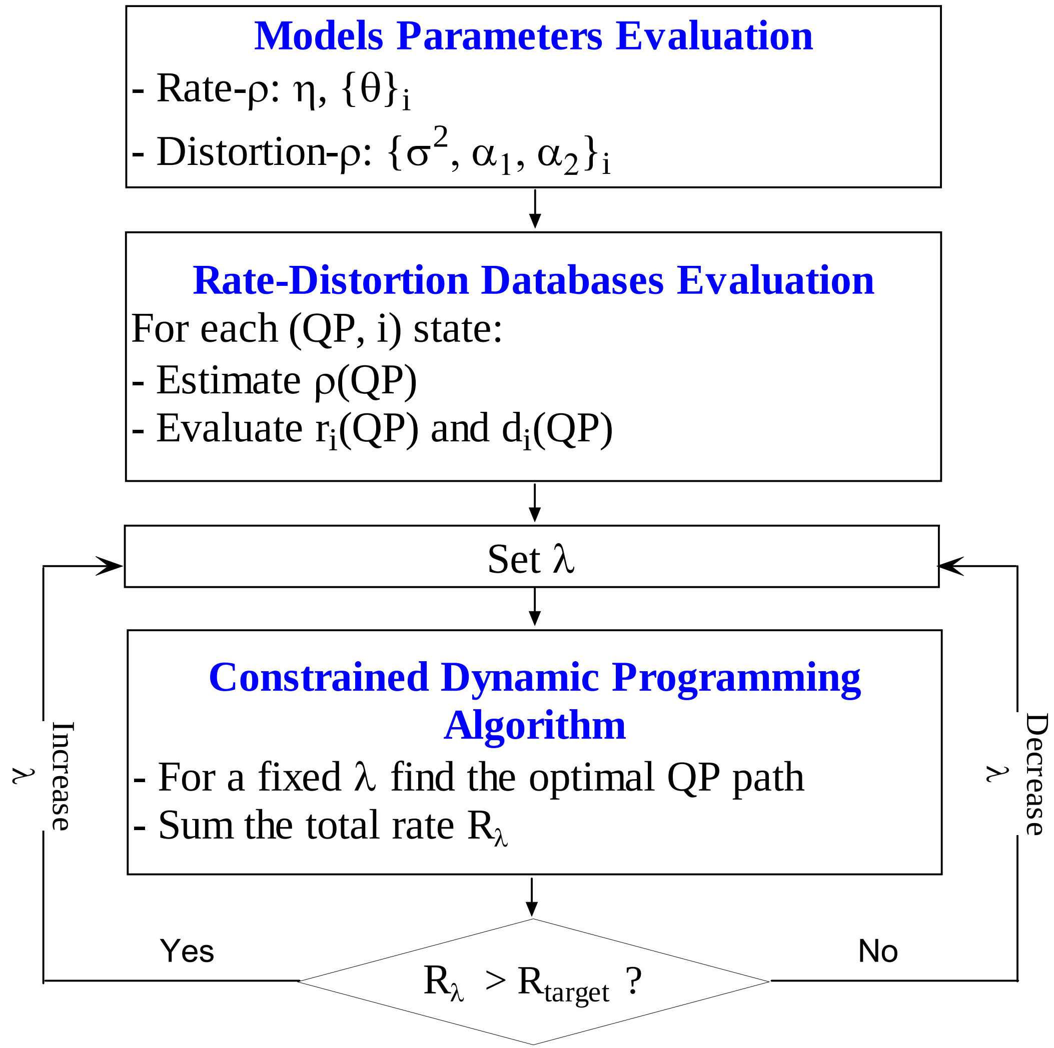


Parameters estimation:

σ^2 - from input data

α_1, α_2 - from two requantization steps coarse evaluation

Proposed algorithm summary

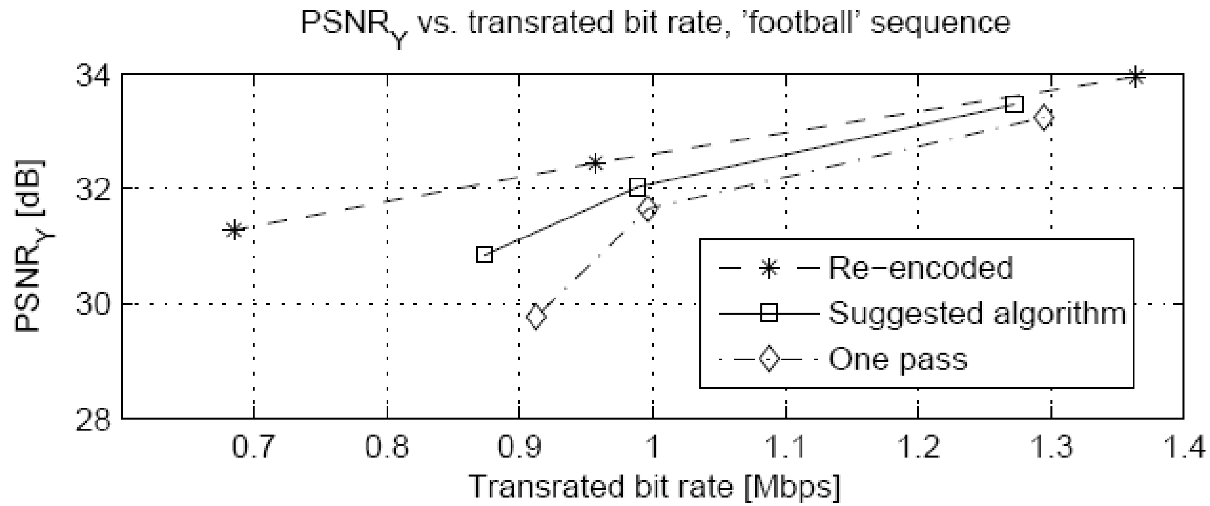


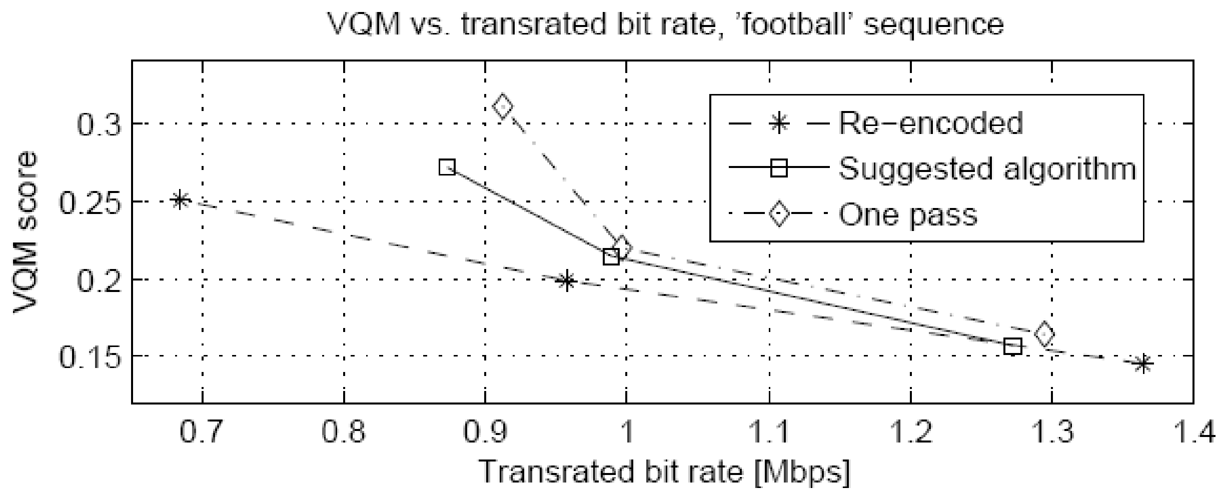
Results

- The proposed algorithm supports only inter coded frames.
  Test description:
  - Intra frame – reencoded.
  - Inter frame – comparison between three schemes:
    - Reencoding
    - Proposed algorithm
    - One-pass requantization
- Computational complexity:
  Our model-based optimization reduces the run time by a factor of 4, as compared to full exhaustive optimization (based on full rate and distortion evaluation).
- Quality comparison:
  - PSNR – objective measure
  - VQM – subjective measure

Quality comparison

Input bit rate: 2 [Mbps]





Conclusion

- We propose a model-based optimal requantization algorithm for H.264 inter coded frames.
- In comparison with a simple one-pass requantization, our algorithm achieves better performance both objectively (PSNR gain of up to 1[dB]) and subjectively.
- We developed rate models suitable for H.264 requantization.
- Model-based optimization reduces the computational complexity by a factor of 4, as compared to full exhaustive optimization.
- Our algorithm has lower computational complexity than reencoding.