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
- $\rho$ 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

\textbf{doesn’t} perform full decoding and full encoding.
Transrating via optimal transform coefficients requantization

\[
\min_{\{QP_i\}} \quad D, \quad \text{subject to} \quad R \leq R_{\text{target}}
\]

where

\[ D = \sum_{i=1}^{N_B} d_i(QP_i) \text{ and } R = \sum_{i=1}^{N_B} r_i(QP_i) \]

- number of blocks in the frame
- quantization parameter for the i-th block
- requantization distortion
- requantization rate

- Common **Lagrangian optimization** approach:
  - Merge rate and distortion using Lagrangian parameter, \( \lambda > 0 \)
    \[ J = D + \lambda(R - R_{\text{target}}) \]  
    (Assunção and Ghanbari, 1997)
  - Decompose cost into sum of independent costs for each block
Optimal requantization in H.264

- \( \Delta QP \) 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 \( \lambda \), find an optimal \( \{QP_i^*\} \) path by solving a dynamic programming problem:
    \[
    \min_{\{QP_i\}} J \quad \text{subject to} \quad |\Delta QP| \leq 2
    \]
    where \( J = D + \lambda (R - R_{\text{target}}) \).
  - Calculate the corresponding rate: \( \sum_{i=1}^{NB} r_i(QP_i^*) \).
  - Change \( \lambda \) 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: \( J = \sum_{i=1}^{NB} j_i(QP) \).
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 \((QP^*_i, i - 1)\).
Dynamic programming – last stage

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

\[
BestPathEnd = \underset{QP}{\arg\min} \ V_{N_B}(QP)
\]

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

---

Optimal requantization in H.264 (cont’d)
Rate-distortion modeling

- The optimization requires evaluating $d_i(QP), r_i(QP)$ for all the $\{(QP, i)\}$ states $\rightarrow$ high computational complexity
- Model-based evaluation greatly reduces the computational load.
- Use r-d models in the $\rho$ 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 - $\rho$ 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 - $\rho$ 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.

![Normalized $r_{data}(\rho)$ vs. $\rho$](image)

- $\theta = 6.2$
- $\eta = 1.36$

- **“Overhead” rate – $\rho$ relation** – uses probability based average code length tables
Distortion - $\rho$ 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.

Parameters estimation:
- $\sigma^2$ - from input data
- $\alpha_1, \alpha_2$ - from two requantization steps
  coarse evaluation

Distortion- $\rho$ model relative error distribution
Proposed algorithm summary

Models Parameters Evaluation
- Rate-\( \rho \): \( \eta, \{\theta\}_i \)
- Distortion-\( \rho \): \( \{\sigma^2, \alpha_1, \alpha_2\}_i \)

Rate-Distortion Databases Evaluation
For each (QP, i) state:
- Estimate \( \rho(QP) \)
- Evaluate \( r_i(QP) \) and \( d_i(QP) \)

Set \( \lambda \)

Constrained Dynamic Programming Algorithm
- For a fixed \( \lambda \) find the optimal QP path
- Sum the total rate \( R_\lambda \)

\( R_\lambda > R_{\text{target}} ? \)

Yes

Increase \( \lambda \)

No

Decrease \( \lambda \)

Nov. 16th, 2006
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

Results (cont’d)

Input bit rate: 2 [Mbps]

\( \text{PSNR}_Y \) vs. transrated bit rate, ‘football’ sequence

\( \text{VQM} \) vs. transrated bit rate, ‘football’ sequence

Nov. 16th, 2006

THE 24TH IEEE CONVENTION OF ELECTRICAL AND ELECTRONICS ENGINEERS IN ISRAEL
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.