H.264 Video Quality Optimizations

Guy Cote
April 2008

gcote@mobilygen.com
http://www.mobilygen.com
Outline

- Introduction
- Background
- Optimal Configuration of Video Encoding
- Downhill Simplex Search Method
- Video Quality Metric (VQM)
- Example 1: MB Level Adaptive Quantization
- Example 2: Quantization Matrix
- Conclusions and Future work
Introduction

- Video codecs such as MPEG-4 AVC / ITU-T H.264 have many encoder parameters to control.
- A systematic framework to optimize encoder parameters is indispensable for optimal video encoding.
- Choice of video quality metric is critical.
- Determining optimal video encoding parameters is formulated as an optimization problem: maximize expected video encoding quality under constraints such as video quality, target bitrate, computation, memory bandwidth, etc.
Background
Motion Compensated Video Coding

FTQ: Forward Transform Quantization
IQT: Inverse Quantization Transform
EC: Entropy Coding
MC: Motion Compensation
ME/IE: Motion Estimation/Intra Estimation
• H.264 offers 2-3X bit rate reduction for the same quality compared to MPEG-2
• The added performance comes with higher encoder complexity

Normalized Codec Computing Requirements
H.264/AVC Encoder Parameters

- **System design parameters**
  Level and profile, Group Of Pictures (GOP) structure
  Entropy coding mode: CAVLC / CABAC, Number of reference frames, Interlace coding support, etc.

- **Encoder internal parameters**
  Motion search range and algorithm, intra / inter encoding partition size and mode decision,
  Intra prediction mode decision, PAFF and MBAFF decisions, MB level quantization parameters, etc.

- **User-controlled external parameters**
  Deblocking filter strength, quantization scaling matrices, quantization rounding offsets, Lagrangian multipliers for RD optimization, various thresholds to bias mode decision, etc.
Formulated as maximizing the expected video encoding performance under a set of constraints

\[ p_{opt} = \arg \max_p \sum_i \kappa_i V(s_i, D(E(s_i, p))) \]

E / D: video encoder / decoder
V: video quality metric
si: video sequence
ki: weighting factor
p: encoding parameters to be optimized
Constrained by a specific bitrate or, quality level
Video Quality Metric (VQM)

- Based on algorithms for objective measurement of video quality
- Correlates well with Mean Opinion Score (MOS) of subjective evaluation, as evaluated by VQEG
- Developed by NTIA / ITS, adopted as an ANSI and ITU standards
- Video Quality Metric has proven to be a very useful tool for not only for evaluation but also for parameter training
- VQM score: between 0 and 1
  0: no perceived impairment
  1: maximum perceived impairment
- Present results as “VQM-Rate” curves
Downhill Simplex (Nelder-Mead) Search

- Multidimensional search within an initial simplex, \( n \) dimensions correspond to \( n \) encoding parameters
- Downhill Simplex use:
  - simplex reflection
  - simplex expansion along minimization direction
  - simplex contraction along inverse minimization direction
- Efficient numerical search method only involving function evaluations
- Parameter optimizer $O$ control the iterations
- Parameter trial generator $G$ generates one or more trial parameter via simplex transformation
Example 1: MB Level Adaptive Quantization
MB Level Adaptive Quantization

- Set QP for each MB based on the local statistics
- MB level quantization step size is picture level quantization step size for an MB modulated by a factor
- The HVS is less sensitive to high motion areas, high frequency areas or areas where DC value is away from picture DC value, so a larger QP should be used in those areas

\[ Q_{MB} = Q_{PIC} \cdot f_{mot} \cdot f_{hf} \cdot f_{dc} \]
A linear fractional model for three modulation factors

\[ f_{mot} = \frac{M_{MB} + \alpha_{mot} \bar{M}_{PIC}}{\alpha_{mot} M_{MB} + \bar{M}_{PIC}} \]

MB level QP is parameterized by three numbers

\[ f_{mot}, f_{hf}, f_{dc} \]

Trial based training framework to obtain optimal configuration for the three parameters
A linear fractional model for all modulation factors

\[
\begin{align*}
  f_{mot} &= \frac{M_{MB} + \alpha_{mot} \overline{M}_{PIC}}{\alpha_{mot} M_{MB} + \overline{M}_{PIC}} \\
  f_{hf} &= \frac{H_{MB} + \alpha_{hf} \overline{H}_{PIC}}{\alpha_{hf} H_{MB} + \overline{H}_{PIC}} \\
  f_{dc} &= \frac{D_{MB} + \alpha_{dc} \overline{D}_{PIC}}{\alpha_{dc} D_{MB} + \overline{D}_{PIC}}
\end{align*}
\]

Logarithmic quantization step size to QP

\[
qp_{MB} = qp_{PIC} + \text{round}(6 \log_2(f_{mot} f_{hf} f_{dc}))
\]
Simulation Results

Bitrate vs VQM score curves

- Chromakey (NoAdaptiveQP)
- Chromakey (AdaptiveQP)
- Driving (NoAdaptiveQP)
- Driving (AdaptiveQP)
- OpeningCeremony (NoAdaptiveQP)
- OpeningCeremony (AdaptiveQP)
- WhaleShow (NoAdaptiveQP)
- WhaleShow (AdaptiveQP)

33% Bitrate Reduction

VQM Score vs Bitrate (Mbps) graph.
Example 2: Quantization matrix (QM) in H.264 / AVC
• QM allows different weighting to be applied according to the sensibility of the Human Visual System (HVS) to a coefficient’s corresponding frequency.

• PSNR is not a good metric to measure visual quality improvements for optimizing quantization matrices.

• Quality metric which considers frequency response of HVS is important – VQM.
Matrix \( [Q_{ij}] \) is a Quantization scaling matrix
Defined to improve visual quality
Up to eight QMs in AVC/H.264
  - Intra: 4x4Y, 4x4Cb, 4x4Cr, 8x8Y
  - Inter: 4x4Y, 4x4Cb, 4x4Cr, 8x8Y
Specified in sequence and / or picture header
Standard defines 4 default matrices
QM Model is necessary to reduce search space
## Default QM

### Inter4x4

\[ Q_4 = \begin{bmatrix} 10 & 14 & 20 & 24 \\ 14 & 20 & 24 & 27 \\ 20 & 24 & 27 & 30 \\ 24 & 27 & 30 & 34 \end{bmatrix} \]

### Inter8x8

Q&M Design

[Watson] use the visibility of quantization error
\[ q(x,y) = 2^*T(x,y) \]
where \( T(x,y) \) is the smallest coefficient that yields a visible signal

[Wu et al] maximize the ratio of distortion decrease to bitrate increase

[Westen et al] minimize a Lagrangian cost that involves bitrate and perceptually weighted distortion

[Lee] QM Modeling for 8x8 blocks:
\[ q(x, y) = a\left(\sqrt{x^2 + y^2} - \frac{7}{\sqrt{2}}\right) + b\sin\left(\frac{\pi}{7\sqrt{2}}\sqrt{x^2 + y^2}\right) + c \]
A Symmetric Quadratic QM Model

\[ q(x, y) = a(x^2 + y^2) + bxy + c(x + y) + d \]

Symmetric \[ q(x, y) = q(y, x) \]

Hankel* when \[ b = 2a \]

Parameters
- \( a, b \) characterize the convexity
- \( c \) characterizes the slope
- \( d \) represents DC stepsize

Four parameters sufficient

*Hankel: constant positive sloping skew-diagonals matrix
Modeling Default QM

Inter4x4

\[(a, b, c, d) = (-0.1875, -0.5500, 5.3125, 9.6750)\]

\[
P_4 = \begin{bmatrix}
10 & 15 & 20 & 24 \\
15 & 19 & 24 & 27 \\
20 & 24 & 27 & 31 \\
24 & 27 & 31 & 33 \\
\end{bmatrix}
\]

Inter8x8

\[(a, b, c, d) = (-0.0290, -0.0563, 2.0712, 10.7535)\]

\[
P_8 = \begin{bmatrix}
11 & 13 & 15 & 17 & 19 & 20 & 22 & 24 \\
13 & 15 & 17 & 19 & 20 & 22 & 24 & 25 \\
15 & 17 & 19 & 20 & 22 & 24 & 25 & 27 \\
17 & 19 & 20 & 22 & 24 & 25 & 27 & 29 \\
19 & 20 & 22 & 24 & 25 & 27 & 29 & 30 \\
20 & 22 & 24 & 25 & 27 & 29 & 30 & 31 \\
22 & 24 & 25 & 27 & 29 & 30 & 31 & 33 \\
24 & 25 & 27 & 29 & 30 & 31 & 33 & 34 \\
\end{bmatrix}
\]

\[\| Q_4 - P_4 \|_E = 0.0256 \]

\[\| Q_8 - P_8 \|_E = 0.0248 \]
QM Design for Perceptual Coding

- QM design for a bitrate range via training

- Train QM for parameters \((a, b, c, d)\) that maximize perceptual video quality measured by Video Quality Metric (VQM)

- Trial based multidimensional numerical search: Downhill Simplex Search
VQM Based QM Training

Bitrate → Sequence → H.264 Encode → Bitstream → H.264 Decode → Decoded sequence → Compute VQM → VQM score → QM optimizer → Optimal QM

QM → Trial QM Generator → H.264 Encode

Mobilygen Proprietary and Confidential
Average VQM (aVQM) Score Over a Bitrate Range

- Model a VQM-rate curve based on four samples
- Third order model
- Average over a given bitrate interval
- Similar to AVSNR

\[ s = a_0 + a_1 r + a_2 r^2 + a_3 r^3 \]

\[
\begin{bmatrix}
1 & r_0 & r_0^2 & r_0^3 \\
1 & r_1 & r_1^2 & r_1^3 \\
1 & r_2 & r_2^2 & r_2^3 \\
1 & r_3 & r_3^2 & r_3^3
\end{bmatrix} \begin{bmatrix}
a_0 \\
a_1 \\
a_2 \\
a_3
\end{bmatrix} = \begin{bmatrix}
s_0 \\
s_1 \\
s_2 \\
s_3
\end{bmatrix}
\]

\[
\bar{s} = \frac{1}{r_h - r_l} \int_{r_l}^{r_h} \left( a_0 + a_1 r + a_2 r^2 + a_3 r^3 \right) dr
\]

Mobilygen Proprietary and Confidential
## Simulation Results

<table>
<thead>
<tr>
<th>Sequence (ARIB 480P)</th>
<th>aVQM Score</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default QM</td>
<td>Trained QM</td>
<td>Difference</td>
<td></td>
</tr>
<tr>
<td>Chromakey</td>
<td>0.4341</td>
<td>0.4197</td>
<td>0.0144</td>
<td></td>
</tr>
<tr>
<td>Driving</td>
<td>0.2465</td>
<td>0.2328</td>
<td>0.0137</td>
<td></td>
</tr>
<tr>
<td>Opening Ceremony</td>
<td>0.2193</td>
<td>0.2061</td>
<td>0.0132</td>
<td></td>
</tr>
<tr>
<td>Whale Show</td>
<td>0.4031</td>
<td>0.3891</td>
<td>0.0140</td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.0138</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Trained QM

**Inter4x4**

\[
T_4 = \begin{bmatrix}
11 & 18 & 29 & 44 \\
18 & 27 & 40 & 56 \\
29 & 40 & 54 & 71 \\
44 & 56 & 71 & 90
\end{bmatrix}
\]

**Inter8x8**

\[
T_8 = \begin{bmatrix}
11 & 14 & 18 & 23 & 29 & 36 & 44 & 53 \\
14 & 18 & 22 & 28 & 34 & 41 & 49 & 59 \\
18 & 22 & 27 & 33 & 40 & 47 & 56 & 65 \\
23 & 28 & 33 & 39 & 46 & 54 & 63 & 73 \\
29 & 34 & 40 & 46 & 54 & 62 & 71 & 82 \\
36 & 41 & 47 & 54 & 62 & 71 & 80 & 91 \\
44 & 49 & 56 & 63 & 71 & 80 & 90 & 101 \\
53 & 59 & 65 & 73 & 82 & 91 & 101 & 113
\end{bmatrix}
\]

*4x4 blocks and 8x8 blocks share the same QM model*
Conclusions

- Introduced a systematic approach of configuring video encoding parameters for optimal video encoding using direct simplex search method
- Search method involves direct encoding of various sequences with different encoding parameter settings
- Applied successfully to find a set of optimal parameter values for MB level QP adaptation in H.264 / AVC, achieving 20-30% bitrate reduction
- Applied successfully to find optimal quantization scaling matrix, achieving 5-8 % bitrate reduction with similar perceptual video quality
Further Reading

- This material is available from Proceedings of SPIE - Volume 6822, January 2008:
  - Huipin Zhang, Guy Cote, “Modeling quantization matrices for perceptual image / video encoding”
  - Huipin Zhang, Guy Cote, “Determining optimal configuration of video encoding parameters using numerical search algorithms”

- VQM references:
  - ANSI T1.801.03-2003, American National Standard for Telecommunications - Digital Transport of One-Way Video Signals - Parameters for Objective Performance Assessment
Future Work

- Explore direct correlation of quantization matrix model parameters to the HVS
- Additional encoder parameters can be trained with the same method
- Ultimately develop low complexity encoder distortion metric that correlates better to MOS to integrate in an RD optimization encoder framework for internal parameter optimization
Questions?