Extending HEVC by an affine motion model

Picture Coding Symposium 2013

Matthias Narroschke, Robin Swoboda
Outline

- Introduction
- Extending HEVC by an affine motion model
- Experiments and results
- Conclusion
Block diagram of an HEVC encoder

Transform / Quantization → Entropy coding

Inverse transform

Loop filters

Block-based MCP (Translational motion model)

MCP: Motion compensated prediction
Translational motion model

Motion description by two-parameter displacement vector: \[ \begin{pmatrix} \Delta_x \\ \Delta_y \end{pmatrix} \]

Not possible to describe complex motion as rotation, zoom
HEVC block partitioning in case of rotation

- Numerous small blocks to approximate description of rotation

- **Problem**: High data rate due to transmission of
  - Displacement vector for each of the numerous small blocks
  - Large prediction error resulting from inaccurate description of motion
Affine motion model to describe complex motion

Description of complex motion with six-parameter vector:

\[ \Delta x_i = a_1 + a_3 \cdot x_i + a_5 \cdot y_j \]

\[ \Delta y_j = a_2 + a_4 \cdot x_i + a_6 \cdot y_j \]
Motivation

- State of the art
  - HEVC applies block-based MCP with a translational motion model, which cannot describe complex motion
  - Affine motion model can describe complex motion and has been investigated for block-based and global MCP on standards preceding HEVC with block sizes limited to 16x16 samples

- Problems
  - Block-based MCP:
    - For small blocks, the data rate increase due to additional model parameters is often larger than the data rate reduction achieved by more accurate prediction
  - Global MCP:
    - Inaccurate prediction in the case of diverse local motion

- Investigation
  - Block-based MCP with affine motion model based on HEVC, which uses increased block sizes of up to 64x64 samples
Block diagram of the investigated encoder

- Transform / Quantization
- Entropy coding
- Inverse transform
- Loop filters
- Block-based MCP (Translational motion model)
- Block-based MCP (Affine motion model)
Encoder control for each block

- Iterative estimation of affine model parameters by minimizing estimated RD costs (Convergence best for small motion)

- Motion model is selected by minimizing estimated RD costs
Coding

- Estimated parameters of the affine model
  - Uniform quantization with step sizes
    - 1/16 for $a_1, a_2$
    - 1/512 for $a_3, \ldots, a_6$
  - Prediction followed by CABAC of prediction error

- Selected motion model: CABAC
Experiments

- Reference: HEVC (HM7)

- Investigated: HEVC (HM7) + Affine motion model

- JCT-VC test conditions
  - Prediction structures: Low Delay, Random Access
  - 4 different quantization step sizes for each sequence
  - Calculation of average bit rate reduction at the same PSNR according to the method proposed by Bjontegaard
Results of experiments for Low Delay prediction structure

Average bit rate reduction at the same PSNR [%]

Sequences of mainly translational motion
Average: 0%

Sequences of mixed motion, translational and non-translational
Average: 1.5%

Sequences of mainly non-translational motion
Average: 6.3%

Random Access: 3.7% on average
Block partitioning in case of rotation

Motion can be efficiently described with small number of large blocks
### Data rate reduction versus maximum block size

<table>
<thead>
<tr>
<th>Maximum applied block size [Samples x Samples]</th>
<th>Defined in HEVC</th>
<th>Beyond HEVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 x 16</td>
<td>0.1</td>
<td>6.3</td>
</tr>
<tr>
<td>64 x 64</td>
<td></td>
<td>7.6</td>
</tr>
<tr>
<td>128 x 128</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average data rate reduction achieved by affine motion model [%]

Block-size increase benefits the use of an affine motion model
Conclusion

- Investigated an extension of HEVC by an affine motion model

Results

- Average bit rate reduction at the same PSNR for sequences of mainly non-translational motion:
  - Low Delay: 6.3%, Random Access: 3.7%
  - For particular sequences: Up to 24%

- Block size enlargement from 64x64 to 128x128 samples further increases average bit rate reduction from 6.3% to 7.6%

Application

Affine motion model together with further increased block sizes could be one direction towards a future video coding standard