Viewport Tile Selection Experiment:
Using Test Model for Immersive Video (TMIV)

Presenter: Soonbin Lee (soonbinlee@skku.edu)
Jong-Beom Jeong, Soonbin Lee, Sungbin Kim, Il-Woong Ryu, Dongmin Jang, Eun-Seok Ryu

Multimedia Computing Systems Lab. (MCSL)
http://mcsl.skku.edu
Department of Computer Education
Sungkyunkwan University
3DoF+ Overview

- 3DoF+ Overview System
  - the MPEG defined 3DoF+ to support the user’s head movements.
  - requires a high computing power and bandwidth
Viewport Streaming for 3DoF+

- Viewport Streaming for 3DoF+
  - When a user’s head movement are transferred from the client, selects the tile sets that belong to the user’s viewport and their bitstreams.
OMAF (Omnidirectional MediA Format)

- OMAF Coordinate System
  - Global coordinate system is defined in OMAF, OMAF is part 2 of MPEG-I
  - Viewport tile decision method for a single 360 video, which is compatible with OMAF
Viewport Tile Selection (VTS)

- Viewport Tile Selection (VTS)
  - Single-pass encoding/rendering with MCTS-encoder and modified TMIV
  - Saving bandwidth with selective tile-based streaming
  - Fewer decoder instance
Experimental Results

- Experimental results for CTC
  - We conducted experiment with CTC defined in MPEG
  - Every evaluation frames is 97

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Class</th>
<th>Resolution</th>
<th>View FoV</th>
<th>Global FoV</th>
<th>No. of views</th>
<th>Evaluation frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClassroomVideo</td>
<td>CG1-A</td>
<td>4096 × 2048</td>
<td>360 × 180</td>
<td>360 × 180</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>TechnicolorMuseum</td>
<td>CG1-B</td>
<td>2048 × 2048</td>
<td>180 × 180</td>
<td>360 × 180</td>
<td>24</td>
<td>100</td>
</tr>
<tr>
<td>TechnicolorHijack</td>
<td>CG1-C</td>
<td>4096 × 4096</td>
<td>180 × 180</td>
<td>180 × 180</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>NokaChess</td>
<td>CG1-N</td>
<td>2048 × 2048</td>
<td>180 × 180</td>
<td>360 × 180</td>
<td>10</td>
<td>60</td>
</tr>
</tbody>
</table>
Experimental Results

- Experimental results for CTC
  - For the objective evaluation, PSNR and IV-PSNR* is used.
  - VTS method showed the better results on BD-rate and number of decoders.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Class</th>
<th>No. of decoders required</th>
<th>High BD-rate Y-PSNR(%)</th>
<th>Low BD-rate Y-PSNR(%)</th>
<th>High BD-rate IV-PSNR(%)</th>
<th>Low BD-rate IV-PSNR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClassroomVideo</td>
<td>CG1-A</td>
<td>4</td>
<td>-23.5</td>
<td>-21.3</td>
<td>-31.1</td>
<td>-24.8</td>
</tr>
<tr>
<td>TechnicolorHijack</td>
<td>CG1-C</td>
<td>6</td>
<td>2.8</td>
<td>3.5</td>
<td>2.8</td>
<td>2.7</td>
</tr>
<tr>
<td>NokaChess</td>
<td>CG1-N</td>
<td>8</td>
<td>-7.3</td>
<td>-4.4</td>
<td>-8.1</td>
<td>-5.0</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>6.5</td>
<td>-13.2</td>
<td>-11.6</td>
<td>-15.1</td>
<td>-12.7</td>
</tr>
</tbody>
</table>
Experimental Results (-Cont’d)

ClassroomVideo

TechnicolorMuseum

Proposal Number (TBD)
Experimental Results (-Cont’d)

- **Technicolor Hijack**:
  - **IV-PSNR (dB)** vs. **Bitrate (Mbps)**
  - Graphs show performance comparison between Anchor (MIV mode), VTS (MIV mode), Anchor (MIV view mode), and VTS (MIV view mode).

- **Nokia Chess**:
  - Similar graph configuration as Technicolor Hijack, comparing performance metrics.

Proposal Number (TBD)
Appendix (-Cont’d)

- CG1-A, position at v0, \((x, y, z, \text{yaw}, \text{pitch}, \text{roll}) = (0, 0, 0, 0, 0, 0)\)
- Pose trace Ap02, \((x, y, z, \text{yaw}, \text{pitch}, \text{roll}) = (0,0,0,1.85\times10^{-05},-2.99\times10^{-06},3.20\times10^{-05})\)
Appendix (-Cont’d)

- CG1-A, position at v0, (x, y, z, yaw, pitch, roll) = (0, 0, 0, 0, 0, 0)
- Pose trace Ap01, (x, y, z, yaw, pitch, roll) = (0.205, 0.09, 0, -20, -5, 0)
Appendix (-Cont’d)

- CG1-N, position at v1, (x, y, z, yaw, pitch, roll) = (-0.20000000298023224, -0.5, 1.0, -0.0, 0.0, 0.0)
- Pose trace Np01, (x, y, z, yaw, pitch, roll) = (0.25, 0.0, 0.0, 5.0888476044965E-14, -2.414835874568E-06, 2.414835874568E-06)
Appendix (-Cont’d)

- CG1-N, position at v1, (x, y, z, yaw, pitch, roll) = (-0.20000000298023224, -0.5, 1.0, -0.0, 0.0, 0.0)
- Pose trace Np02, (x, y, z, yaw, pitch, roll) = (-0.25, 0.0, 0.0, -20.200002725009400, 29.99999741973150, -8.79232615657642E-09)