Tile Extractor Optimization for Low-latency Viewport-dependent 360 Video Tiled Streaming

Presenter: Eun-Seok Ryu (esryu@skku.edu)
Jong-Beom Jeong, Soonbin Lee, Inae Kim, Eun-Seok Ryu

Multimedia Computing Systems Lab. (MCSL)
http://mcsl.skku.edu
Department of Computer Education
Sungkyunkwan University
Subclause 5.2.1 of the *IEEE-SA Standards Board Bylaws* states, "While participating in IEEE standards development activities, all participants...shall act in accordance with all applicable laws (nation-based and international), the IEEE Code of Ethics, and with IEEE Standards policies and procedures."

The contributor acknowledges and accepts that this contribution is subject to


Tile Extractor Optimization for Low-latency Viewport-dependent 360 Video Tiled Streaming

Date: 2020-10-19

Author(s): Jong-Beom Jeong, Soonbin Lee, Il-Woong Ryu, Sungbin Kim, Inae Kim, Eun-Seok Ryu

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Phone [optional]</th>
<th>Email [optional]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jong-Beom Jeong</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:uof4949@skku.edu">uof4949@skku.edu</a></td>
</tr>
<tr>
<td>Soonbin Lee</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:soonbinlee@skku.edu">soonbinlee@skku.edu</a></td>
</tr>
<tr>
<td>Inae Kim</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:inaelk@skku.edu">inaelk@skku.edu</a></td>
</tr>
<tr>
<td>Eun-Seok Ryu</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:esryu@skku.edu">esryu@skku.edu</a></td>
</tr>
</tbody>
</table>
360-degree Video Tiled Streaming

- High resolution and framerate for high-quality 360-degree video streaming
  - high bandwidth
- Tiled streaming based on MCTS for viewport-dependent streaming
  - well-known selective streaming method

### Requirements for high quality VR

Source: Technicolor, Oct. 2016 (m39532, MPEG 116th Meeting)

<table>
<thead>
<tr>
<th>Requirement</th>
<th>details</th>
</tr>
</thead>
<tbody>
<tr>
<td>pixels/degree</td>
<td>40 pix/deg</td>
</tr>
<tr>
<td>video resolution</td>
<td>11520x6480</td>
</tr>
<tr>
<td>framerate</td>
<td>90 fps</td>
</tr>
</tbody>
</table>

360-degree Video Tiled Streaming: Challenges

- In HEVC test model (HM) 16.20, a tile extraction SW is included
  - single tile extraction is available
- Based on HM 16.20, number of tile bitstreams are generated
- Generally, a VR player has one or few number of decoders
  - single tile extraction causes decoding resource issues

Example of viewport-dependent 360-degree video streaming
Multiple-tile Extraction for Low-latency Streaming

- Server-driven approach to reduce the latency: multiple-tile extraction, base layer simulcasting
- Extracts multiple tiles and generates single bitstream ➢ two decoders for high-quality low-latency tiled streaming

Low-latency 360-degree video streaming using multiple-tile extraction
Experimental Results: BD-rate Saving

- Used four 4K 360-degree test sequences
- Three tiling scenarios (2 × 4, 3 × 6, 6 × 12) were used
- In tiled streaming, low-quality base layer (e.g. QP=42) was transmitted
- 16.98% of BD-rate saving for Y-PSNR was shown
- 6 × 12 tiling showed the best BD-rate
  - in traditional tiled streaming, generates many bitstreams

<table>
<thead>
<tr>
<th>Tiling</th>
<th>Y-PSNR</th>
<th>VMAF</th>
<th>MS-SSIM</th>
<th>IV-PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 × 4</td>
<td>-6.05%</td>
<td>2.95%</td>
<td>-1.73%</td>
<td>-5.14%</td>
</tr>
<tr>
<td>3 × 6</td>
<td>-19.93%</td>
<td>-8.36%</td>
<td>-12.90%</td>
<td>-17.86%</td>
</tr>
<tr>
<td>6 × 12</td>
<td>-24.97%</td>
<td>-13.96%</td>
<td>18.77%</td>
<td>24.32%</td>
</tr>
<tr>
<td>Average</td>
<td>-16.98%</td>
<td>-6.45%</td>
<td>-11.13%</td>
<td>-15.77%</td>
</tr>
</tbody>
</table>

BD-rate savings of the tiled streaming compared to non-tiled streaming

RD-curve of the tiled streaming, non-tiled streaming
Experimental Results: Decoding Resource Saving

- Compared single-tile extractor (STE) and multiple-tile extractor (MTE) in terms of: decoding memory use and delay
- The client has a Intel i7-7700k CPU (4 cores, 8 threads), 16 GB memory, GTX 1080 Ti
- $6 \times 12$ tiling consumed 52.99 GB of memory and 43.04 seconds using STE
  - MTE required 10.00 GB and 7.08 seconds

Performance comparison between a single tile extractor and multiple tile extractor in terms of (a) decoding memory use, (b) decoding delay
Conclusion

• Motivation
  • 360-degree video streaming requires high bandwidth: tiled streaming can be used
  • Single tile extraction requires number of decoders
    ➢ increases latency which influences quality of experience (QoE)

• Proposed Methods and Insights
  • Multiple-tile extraction for generating single bitstream
  • Showed decoding resource savings compared to the single tile extraction
    ➢ 66.16%, 69.79% decoding memory and decoding delay savings

• Future Work
  • Experiments on high-resolution video (e.g. 8K / 16K) will be conducted