Latency for Viewport-adaptive 360-degree Video Streaming Toward Immersive Experience

Presenter: Jong-Beom Jeong (uof4949@skku.edu)
Tuan Thanh Le, Jong-Beom Jeong, Soonbin Lee, Il-Woong Ryu, Sungbin Kim, Inae Kim, Eun-Seok Ryu

Multimedia Computing Systems Lab. (MCSL)
http://mcsl.skku.edu
Department of Computer Education
Sungkyunkwan University
Compliance with IEEE Standards Policies and Procedures

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:


Latency for Viewport-adaptive 360-degree Video Streaming Toward Immersive Experience

**Date:** 2020-07-09

**Author(s):** Tuan Thanh Le, 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>Tuan Thanh Le</td>
<td>Gachon Univ.</td>
<td></td>
<td></td>
</tr>
<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></td>
</tr>
<tr>
<td>Il-Woong Ryu</td>
<td>Gachon Univ.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sungbin Kim</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inae Kim</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td></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>
Challenges

- **High Bandwidth Requirement of VR**
- Requires 40 pix/deg, 12K resolution for High quality VR
- To avoid the sickness, 90 fps and 20 ms MTP are required
- Immersive video contains texture (color) and depth (geometry) \( \times 2 \)
- Also, immersive video has high quality (nearly 4K) multiple views \( \times N \)
  -> Requires high bandwidth

<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>Resolution</td>
<td>11520x6480 (12K)</td>
</tr>
<tr>
<td>Framerate</td>
<td>90 fps</td>
</tr>
<tr>
<td>Motion-to-photon-latency</td>
<td>20 ms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Resolution</th>
<th>No. of views</th>
<th>Frame count</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClassroomVideo</td>
<td>4096x2048</td>
<td>15</td>
<td>120</td>
</tr>
<tr>
<td>TechnicolorMuseum</td>
<td>2048x2048</td>
<td>24</td>
<td>300</td>
</tr>
<tr>
<td>TechnicolorHijack</td>
<td>4096x4096</td>
<td>10</td>
<td>300</td>
</tr>
<tr>
<td>TechnicolorPainter</td>
<td>2048x1088</td>
<td>16</td>
<td>300</td>
</tr>
<tr>
<td>IntelKermit</td>
<td>1920x1080</td>
<td>13</td>
<td>300</td>
</tr>
</tbody>
</table>

Requirements for high quality VR  
Characteristics of immersive video

Viewport Adaptive Streaming

- **Motion Constrained Tile Sets (MCTS)** refer to the encoder for time and space movement reference for independent tile transfer within the current location tile.
- Extract and composite specific tiles from the bitstream with MCTS to form an adaptive environment at the time of the user.
- Reduce bandwidth when sending only tiles that correspond to a user’s area of interest.

Viewport-driven tiled streaming system based on MCTS.

Ensure independence between tiles.

Tiles
VR Tiled Streaming Latency

- We define latency as the total time between movement of the user’s head and the updated image being displayed on the screen.
- It includes the times for sensor response, fusion, rendering, image transmission, and display response.

\[ Total_{\text{latency}} = \Delta t_1 + \Delta t_2 + \Delta t_3 + \Delta t_4 + \Delta t_5 \]
VR Tiled Streaming System Architecture

Streaming Server

Selected Tiles

Decoding, mapping

Tile selection, chunk request

DASH/RTSP/RTP Client

Prediction Model

Adaptive Network Bandwidth Model

VR rendering

PC - Client

Head’s movement

viewport

viewport

3079-20-0029-02-0003 Latency for Viewport-adaptive 360-degree Video Streaming Toward Immersive Experience
Opportunities for Research

- Optimizing size of tiles
  - Adaptive Bandwidth
  - Reduce latency (Δt_3, Δt_4 and Δt_5)

- Head’s movement prediction
  - According to eyes-coordinates function and a prediction model
  - Decide the next tiles of new tiled stream
  - Reduce latency (Δt_1)

- High quality and low-quality streams
  - Instead of requesting a new chunk, client uses low-quality tiles at exact same location
  - Reduce latency (Δt_1, Δt_3, Δt_4, and Δt_5)