On Inter-view Redundancy-aware Cluster Based Coding Structure Decision on MV-HEVC for MPEG Immersive Video

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


On Inter-view Redundancy-aware Cluster Based Coding Structure Decision on MV-HEVC for MPEG Immersive Video

Date: 2021-10-04

Author(s): Jong-Beom Jeong, Changhee Han, Soonbin Lee, Inae Kim, Junho Park, 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>Changhee Han</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:chillycorn@skku.edu">chillycorn@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>Junho Park</td>
<td>Sungkyunkwan Univ.</td>
<td></td>
<td><a href="mailto:wnsgh2933@skku.edu">wnsgh2933@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>
MPEG Immersive Video (MIV)

- Challenges of multi-view videos processing to provide immersion:
  - high bandwidth, high computational complexity, high latency
- MIV for immersive video pre- and post-processing
  - Test model for immersive video (TMIV) as a reference software
  - MIV mode: removes redundancy from the source views (+HEVC, VVC) → efficient on low bandwidth
  - MIV view mode: selects subset of views (+MV-HEVC) → efficient on high bandwidth, advantageous for multi-view color consistency (low sickness)

MIV pre-processing for immersive video (version 8.0)
**Background**

- **MV-HEVC** as an extension of HEVC to:
  - Efficient compression of multi-view videos by exploiting inter-view redundancy
  - Highly efficient for structured (rectangular) camera coordination
  - Inter-view coding structure needs to be defined
  - How to decide the optimal coding structure?

**Example of MV-HEVC coding structure**
Cluster-based Pruning in MIV

- Redundancies are removed by pruning process
- View optimizer: divides basic views (BVs) and additional views (AVs)
  - Redundancies are removed from AVs
- Pruning cluster generates cluster graph to:
  - Reduce the computational complexity
  - Remove the redundancies more efficiently

Example of pruning cluster generation by TMIV (version 8.0)
Redundancy-aware Coding Structure Decision

- Method 1: view index based ordering
  - Decides coding structure by view index (ascending order)
- Method 2: pruning based ordering
  - Coding structure mode decision by pruning order (no group)
- Method 3: cluster based ordering
  - Uses cluster-based approach for 1) increased coding efficiency, 2) parallel decoding → low sickness and low latency

Coding structure decision modes.
(a) view index based ordering, (b) pruning based ordering, (c) cluster based ordering
Experimental Setup

• MV-HEVC test model (HTM) v13.0, TMIV v3.0
  ➢ TMIV v8.0 was used to decide the pruning graph
  ➢ TMIV v3.0 is better when using MV-HEVC
• Two MIV test sequences were used
  ➢ Painter(2048 \times 1088) – 16 views, Frog(1920 \times 1080) – 13 views

View representations and camera coordinations.
Upper: Painter, Lower: Frog.
Experimental Results: Coding Efficiency

- Y-PSNR BD-rate gains compared to the anchor (view index based)
- In average, the cluster-based method showed the best bandwidth saving

<table>
<thead>
<tr>
<th>Sequences</th>
<th>View index based</th>
<th>Pruning based</th>
<th>Cluster based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Painter</td>
<td>0.00%</td>
<td>-0.78%</td>
<td>-8.33%</td>
</tr>
<tr>
<td>Frog</td>
<td>0.00%</td>
<td>-3.36%</td>
<td>0.56%</td>
</tr>
<tr>
<td>Average</td>
<td>0.00%</td>
<td>-2.07%</td>
<td>-3.88%</td>
</tr>
</tbody>
</table>

Y-PSNR BD-rates

Rate-distortion curves of Painter (left), Frog (right)
Experimental Results: Decoding runtime

- Decoding runtime savings compared to the anchor (view index based)
- If parallel decoding is available, cluster-based method can halve the runtime

<table>
<thead>
<tr>
<th>Sequences</th>
<th>View index based</th>
<th>Pruning based</th>
<th>Cluster based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Painter</td>
<td>100.00%</td>
<td>96.53%</td>
<td>96.81%</td>
</tr>
<tr>
<td>Frog</td>
<td>100.00%</td>
<td>96.06%</td>
<td>101.41%</td>
</tr>
<tr>
<td>Average</td>
<td><strong>100.00%</strong></td>
<td><strong>96.30%</strong></td>
<td><strong>99.11%</strong></td>
</tr>
</tbody>
</table>

Decoding runtime curves
Conclusion

• Motivation
  • High BW & computing time for multi-view videos streaming: MIV & MV-HEVC can be used
  • View clustering and ordering is needed to:
    ➢ Increased coding efficiency (better quality) and decreased latency

• Proposed Methods and Insights
  • Coding structure decision considering: 1) pruning order and 2) clustering by MIV
  • BD-rate, decoding time savings compared to the anchor
  • $2 \times$ faster decoding runtime when using parallel processing to the cluster-based method

• Conclusion and Future Work
  • Inter-view redundancy-aware cluster based coding structure decision for immersive video leads to the decreases of bitrate and latency
  • Experiments on parallel processing / selective group streaming needs to be conducted