Spatio-Temporal Activity based Tiling for Panorama Streaming

Y. Sanchez, R. Skupin, C. Hellge, and T. Schierl
Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute (HHI)

Outline

- Motivation
- Problems to be solved
- Related work
- Spatial-temporal activity metrics
- Validations
- Conclusion
Motivation

- In panorama streaming, users can navigate the high-resolution videos with an arbitrary Region-of-Interest (RoI)
- Transmitting the whole video is unfeasible
- **Tile-based panoramic streaming** overcomes the mentioned drawbacks
To be solved...

- Tile based panoramic streaming allows users to receive a set of tiles that match their RoI instead of the whole video.
- How to derive from the video content the optimal tile size in a low complexity manner.
Related work

- Brute force approach [1]
- High computational complexity
- Pixel overhead & bitrate per pixel (bpp)

\[ BR(t_w, t_h) = Roi_{size} \times \eta(t_w, t_h) \times \varphi(t_w, t_h) \]  \hspace{1cm} (1)

with \( \eta \) being:

\[ \eta(t_w, t_h) = \frac{(r_w+t_w-1)(r_h+t_h-1)}{Roi_{size}} \]  \hspace{1cm} (2)

Spatial-temporal activity metrics

- Based on aforementioned method,

\[
(t_{w}^{opt}, t_{h}^{opt}) = \arg \min_{(t_{w}, t_{h})} \eta(t_{w}, t_{h}) \times (1 + BD(t_{w}, t_{h}))
\]

- Pixel overhead is easy to be computed and does not depend on the content

- Bjøntegaard-Delta bitrate (BD-rate) measurement method
  - For the same PSNR, and
  - Negative values tell how much lower the bitrate is reduced (coding efficiency is increased), and positive values tell how much the bitrate is increased (coding efficiency is reduced)
Spatial-temporal activity metrics (cont.)

- For a high number of tiles the BD-rate can vary from around 20% to around 120%.
- The test conditions used by JCT-VC during standardization.
Spatial-temporal activity metrics (cont.)

- Based on [2] and [3], it can derive the equation below,

\[
BD_{model}(t_w, t_h) = \frac{1}{0.82 + 1.5SA^{-0.05}TA^{-0.04}} \times (N_{tiles}(t_w, t_h) - 1)^{0.32SA^{0.33}TA^{0.11}}
\] (10)
Spatial-temporal activity metrics (cont.)

- The presented model for each sequence as well as the actual BD-rate values

\[ BD_{model}(t_w, t_h) = \frac{1}{0.82 + 1.5 \times SA^{-0.05} \times TA^{-0.04} \times (N_{tiles}(t_w, t_h) - 1)^{0.32 + SA^{0.33} \times TA^{0.11}}} \]
Validations

- **Real** -> the actual BD-rate values measured for tiled encodings
- **Model** -> the predict values measured by the proposed model

### Table 4: Optimal size using the real BD-rate and the model in Eq. 10

<table>
<thead>
<tr>
<th></th>
<th>Real</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq1</td>
<td>384x192</td>
<td>384x192</td>
</tr>
<tr>
<td>Seq2</td>
<td>256x128</td>
<td>256x128</td>
</tr>
<tr>
<td>Seq3</td>
<td>256x128</td>
<td>256x128</td>
</tr>
<tr>
<td>Seq4</td>
<td>256x128</td>
<td>256x128</td>
</tr>
<tr>
<td>Seq5</td>
<td>576x192</td>
<td>576x192</td>
</tr>
<tr>
<td>Seq6</td>
<td>384x192</td>
<td>256x128*</td>
</tr>
<tr>
<td>Seq7</td>
<td>384x192</td>
<td>384x192</td>
</tr>
<tr>
<td>Seq8</td>
<td>384x192</td>
<td>384x192</td>
</tr>
<tr>
<td>Seq9</td>
<td>256x128</td>
<td>256x128</td>
</tr>
</tbody>
</table>

*different result when using the model
Validations (cont.)

- The most occurring tile sizes among the 9 sequences is 384x192 and 256x128
- It computes the BD-rate savings in comparison to having a static configuration

![Graph showing BD-rate savings for different sequences]

- Avg. 5.9%
- Avg. 2.5%
Conclusion

- On average gains of 5.9% are achieved as tile size is 256x128
- On average gains of 2.5% are achieved as tile size is 384x192

- They proposed an optimization process, which has a low complexity compared to performing encodings for each tiling variant, in order to minimize the transmitted bitrate of the RoI content
Q & A