Virtual Reality Content Streaming: Viewport-Dependent Projection and Tile-based Techniques

IEEE International Conference on Image Processing (ICIP) 2017

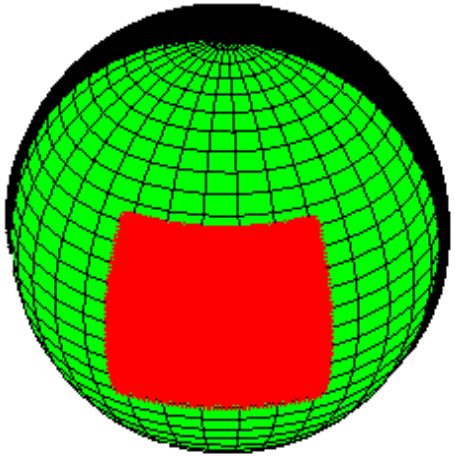
Wednesday, September 20

Alireza Zare, Alireza Aminlou, Miska M. Hannuksela

Nokia Technologies, Tampere, Finalnd

NOKIA

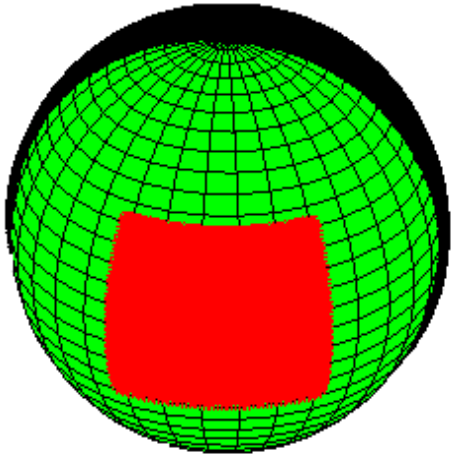
Problem statement



- Head Mountain Displays (HMDs) requirements:
  - High spatial and temporal fidelity contents
  - Strict low-latency
- A limited part of content is displayed.
- Transmitting the entire 360° video sacrifices
  - Network bandwidth
  - decoder capability
- Delivering only the viewport
- But we need a full representation of the spherical video

Solution

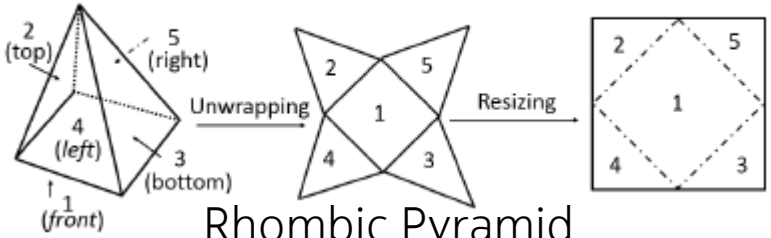
Transmitting the viewport at high quality and the non-viewport part at a lower quality



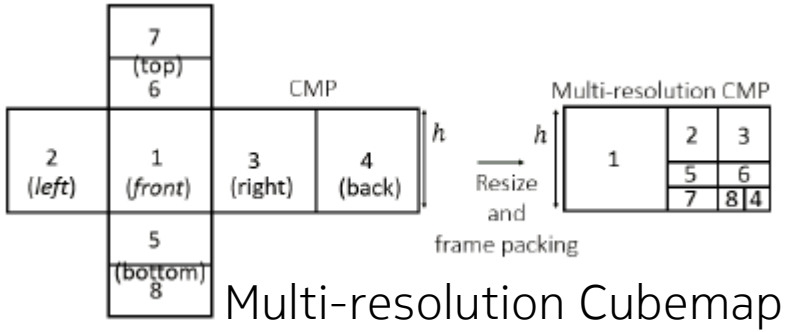
- Viewport-aware adaptation VR streaming techniques
  - Viewport-dependent projection
  - Tile-based technique

Viewport-dependent projection

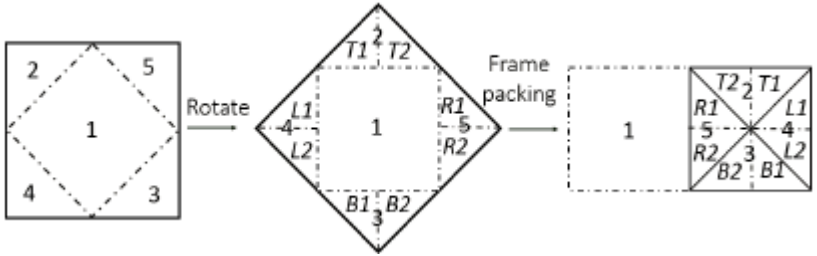
Projecting/mapping 360° video onto multiple viewport representations



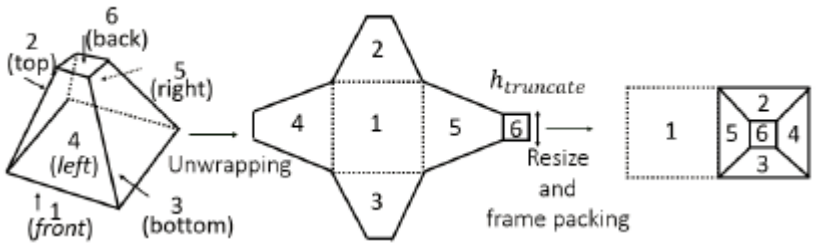
Rhombic Pyramid



Multi-resolution Cubemap



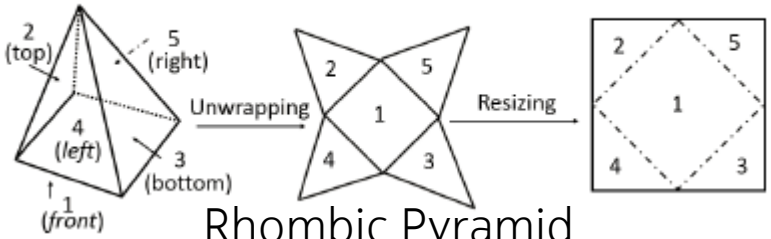
Square Pyramid



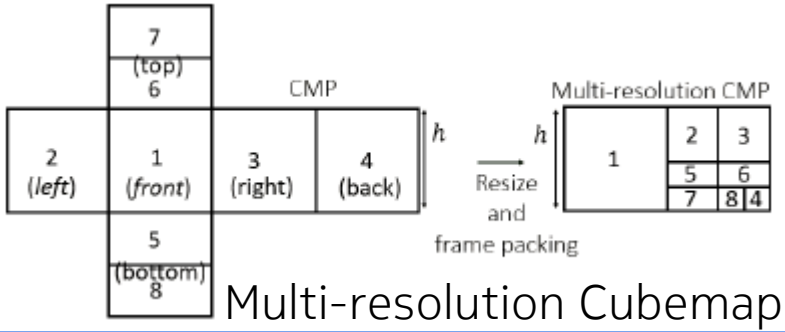
Truncated Square pyramid

Viewport-dependent projection

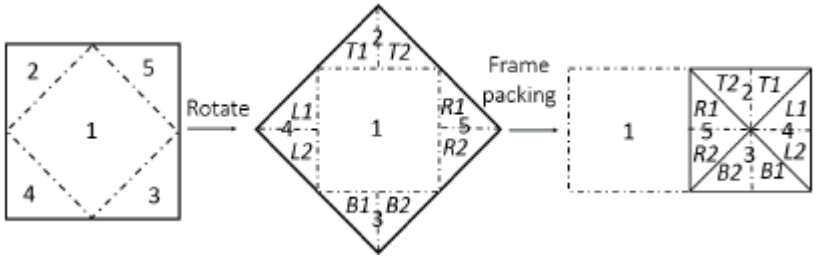
Projecting/mapping 360° video onto multiple viewport representations



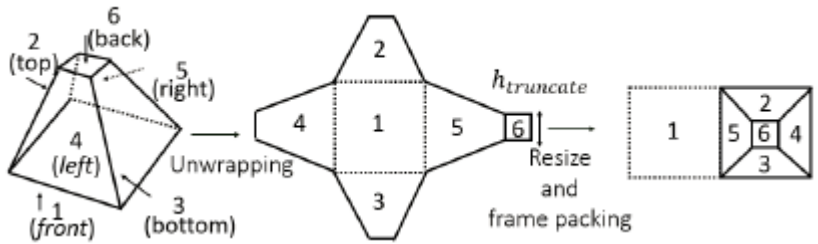
Rhombic Pyramid



Multi-resolution Cubemap



Square Pyramid



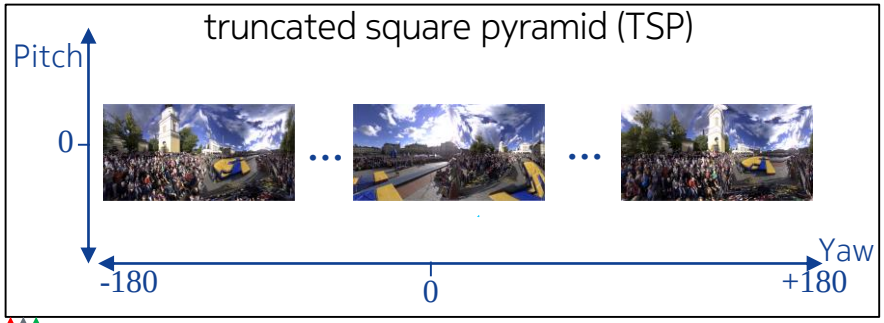
Truncated Square pyramid

NOKIA

Viewport-dependent projection

In each viewport representation:

- Viewport at higher quality
- Non-viewport at lower quality.



Input 360° content

Viewport-based projection/mapping

Encoding and packaging

Orientation feedback

Network

Selected viewport representation

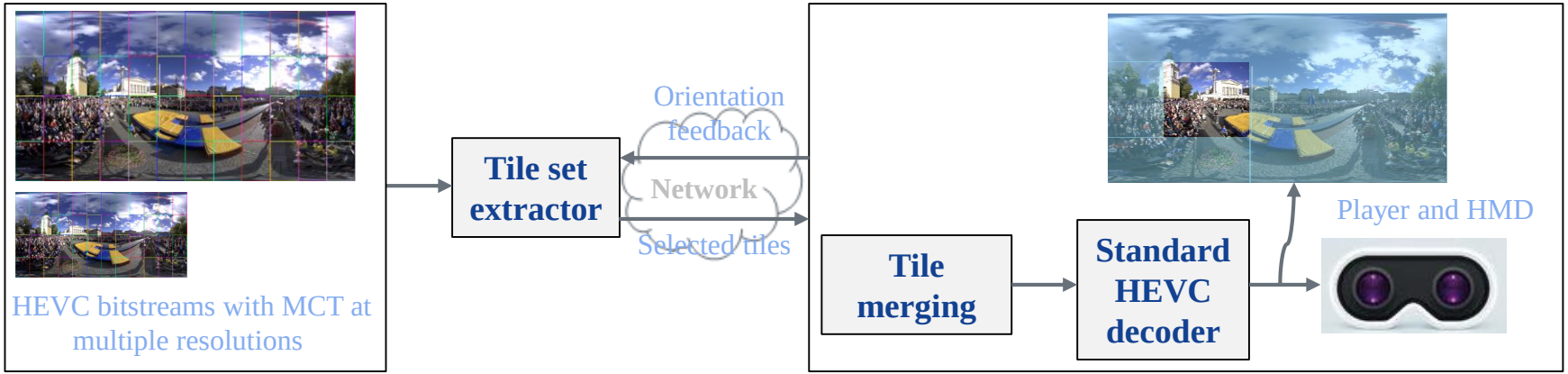
Standard HEVC decoder

Player and HMD

VR viewport-dependent unicast streaming system

Tile-based streaming

- Dividing 360° video to several tiles coded independently in varying quality
- Combining tiles with varying quality to generate a viewport representation



VR tile-based unicast streaming system

NOKIA

Systematic comparison of VR streaming techniques

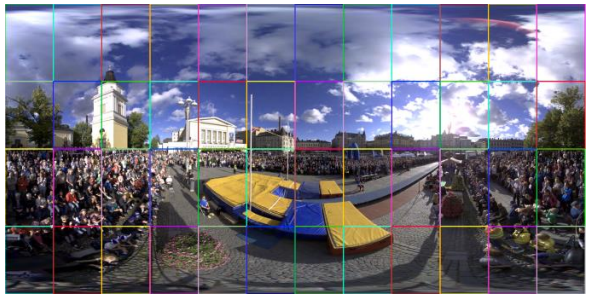
- Defining a framework for fair comparison
  - Projection/mapping
  - Number and distribution of viewport representations
  - FOV of viewport
  - Switching delay
  - Head motion model
- A fair quality assessment methodology

Our contribution

- Defining a framework for fair comparison of two techniques
  - Projection/mapping
  - Number and distribution of viewport representations
  - FOV of viewport
  - Switching delay
  - Head motion speed
- Proposing a quality assessment methodology

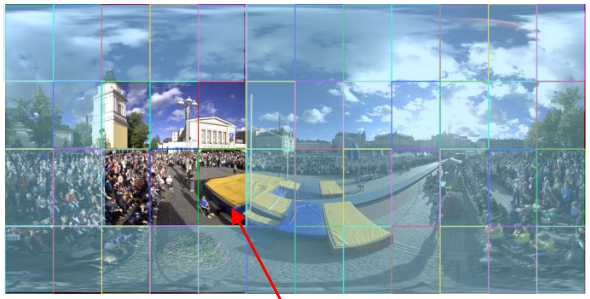
Streaming framework

- Aligning two streaming techniques
- 12x4 tiling
- 3x2 tiles cover 90°x90° FOV

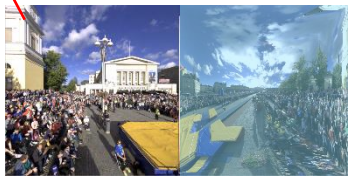


Streaming framework

- Aligning two streaming techniques
- 12x4 tiling
- 3x2 tiles cover 90°x90° FOV

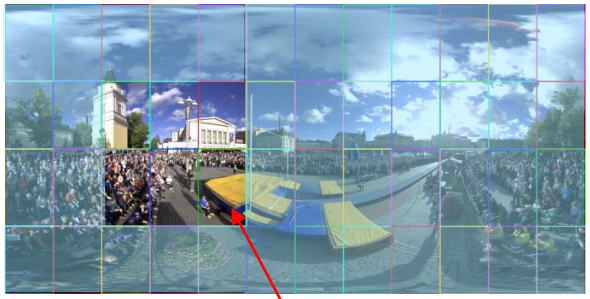


Corresponding 90°x90° FOV viewport in TSP-based method

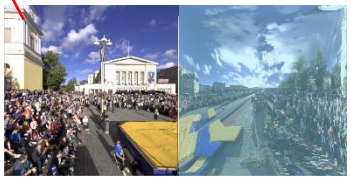


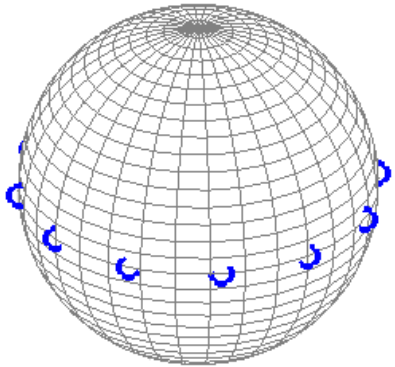
Streaming framework

- Aligning two streaming techniques
- 12x4 tiling
- 3x2 tiles cover 90°x90° FOV
- 12 viewport representations along the equator



Corresponding 90°x90° FOV viewport in TSP-based method

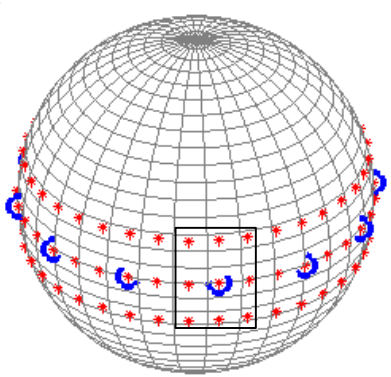




Blue marks: viewport center

Quality assessment methodology

- measuring the quality of experience over a set of discrete quality assessment view (QAV).
- Rendering a cubemap using the closest viewport representation
- To consider head motion: Separating viewport and non-viewport parts



Red marks: center of QAV

Experimental results

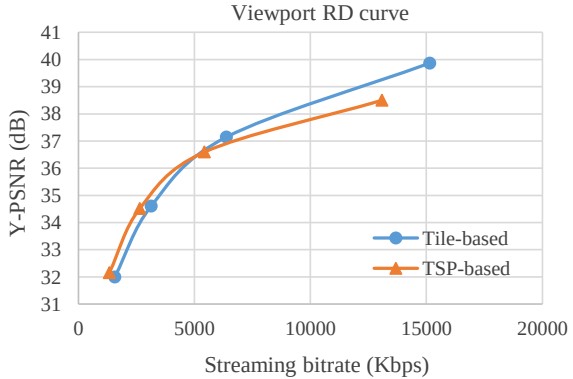
- Standard HEVC encoder and decoder
- Joint Video Exploration Team (JVET) video sequences and 360Lib tool
- The non-viewport tiles coded with +7 higher QP
- Positive values in BD-Rate indicate that TSP-based method outperforms
- Negative values in BD-PSNR indicate that TSP-based method outperforms

Streaming bitrate comparison between tile-based and TSP-based methods.

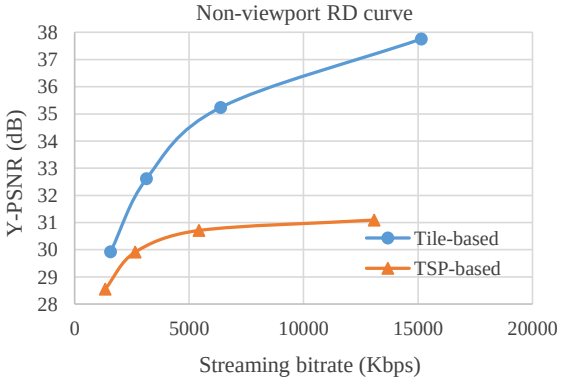
	Viewport		Non-viewport	
Test sequences	BD-Rate (%)	BD-PSNR (dB)	BD-Rate (%)	BD-PSNR (dB)
AerialCity	0.68	0.12	-54.90	3.31
DrivingInCity	4.78	-0.05	-66.60	3.77
DrivingInCountry	-2.07	0.16	-54.15	2.56
PoleVault	5.52	-0.08	-60.02	3.52
Harbor360	13.89	-0.44	-40.82	2.28
KiteFlite360	19.07	-0.77	-27.22	1.65
Skateboard_trick	3.59	-0.09	-16.86	0.63
Train	12.20	-0.43	-14.65	0.83
Average	**7.21**	**-0.20**	**-41.90**	**2.32**

NOKIA

Experimental results



Viewport R-D curve, PoleVault test sequence



Non-viewport R-D curve, PoleVault test sequence

- Considerable low storage requirement in tiling method
- The ratio of 29% with 12 number of viewport representations

NOKIA

Drawbacks of viewport-dependent projection

- Not adapted to the characteristics of the HMDs
- Significant encoding and storage overhead
- Extra pre-processing

Conclusion

A comparison was made between the two recently emerged viewport-adaptive streaming techniques, tile-based and viewport-dependent projection.
A VR quality assessment method was proposed.
Slightly lower streaming performance in tile-based method
Achieving higher performance in tiling method by optimizing non-viewport
Much less preprocessing and encoding time in tiling method
More flexibility to adapt to the characteristics of HMDs in tiling method

Thank you for your attention!

NOKIA OZO