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

- The IEEE Standards copyright policy as stated in the *IEEE-SA Standards Board Bylaws*, section 7, http://standards.ieee.org/develop/policies/bylaws/sect6-7.html#7, and the *IEEE-SA Standards Board Operations Manual*, section 6.1, http://standards.ieee.org/develop/policies/opman/sect6.html
- The IEEE Standards patent policy as stated in the *IEEE-SA Standards Board Bylaws*, section 6, http://standards.ieee.org/guides/bylaws/sect6-7.html#6, and the *IEEE-SA Standards Board Operations Manual*, section 6.3, http://standards.ieee.org/develop/policies/opman/sect6.html

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

Name	Affiliation	Phone [optional]	Email [optional]
Tuan Thanh Le	Gachon Univ.		
Jong-Beom Jeong	Sungkyunkwan Univ.		uof4949@skku.edu
Soonbin Lee	Sungkyunkwan Univ.		
Il-Woong Ryu	Gachon Univ.		
Sungbin Kim	Sungkyunkwan Univ.		
Inae Kim	Sungkyunkwan Univ.		
Eun-Seok Ryu	Sungkyunkwan Univ.		esryu@skku.edu

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) -> (×2)
  - Also, immersive video has high quality (nearly 4K) multiple views -> (×N)
    -> Requires high bandwidth

Requirement	details
Pixels/degree	40 pix/deg
Resolution	11520x6480 (12K)
Framerate	90 fps
Motion-to-photon-latency	20 ms

Requirements for high quality VR

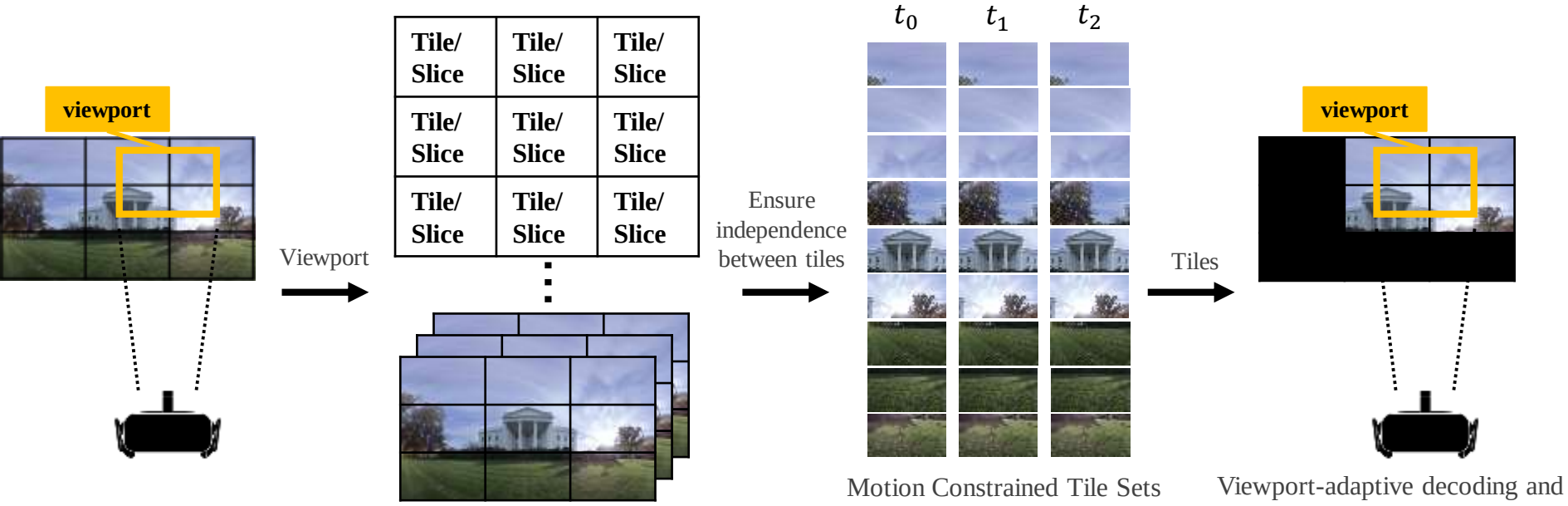
Sequence	Resolution	No. of views	Frame count
ClassroomVideo	4096x2048	15	120
TechnicolorMuseum	2048x2048	24	300
TechnicolorHijack	4096x4096	10	300
TechnicolorPainter	2048x1088	16	300
IntelKermit	1920x1080	13	300

Characteristics of immersive video

Source: Mary-Luc Champel, Thomas Stockhammer, Thierry Fautier, Emmanuel Thomas, and Rob Koenen. "Quality Requirements for VR", document MPEG116/m39532, 116th MPEG meeting of ISO/IEC JTC1/SC29/WG11.

Virtual Reality Streaming Technologies

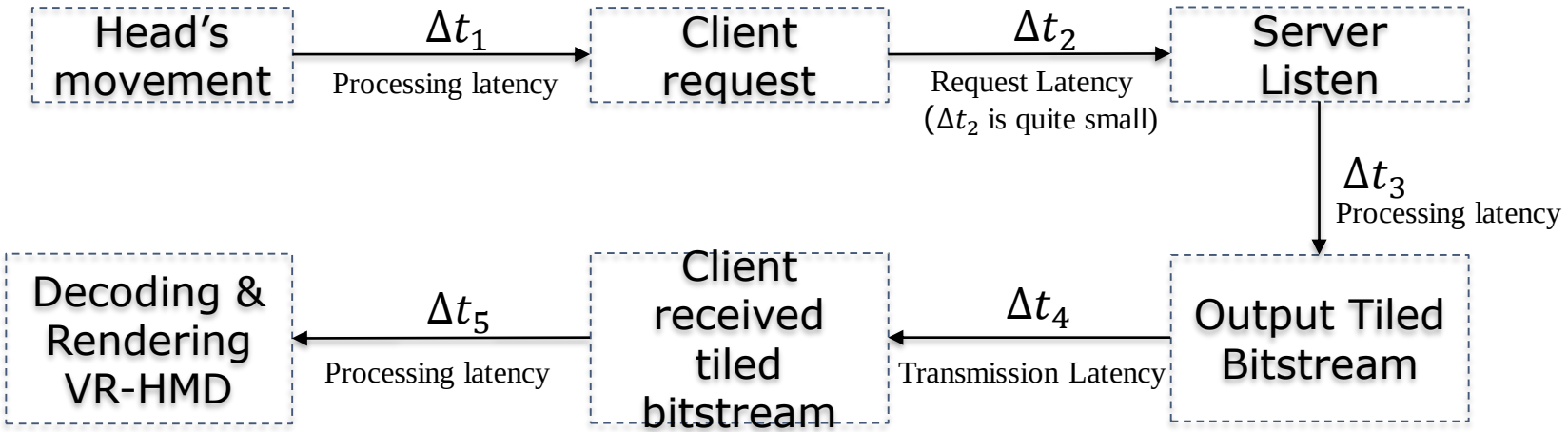
- Viewport Adaptive Streaming
- Motion Constrained Tile Sets (MCTS) refer to the encoder for time and space movement reference for independent tile transfer within the current location tile
- Extract and composite specific tiles from the bitstream with MCTS to form an adaptive environment at the time 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

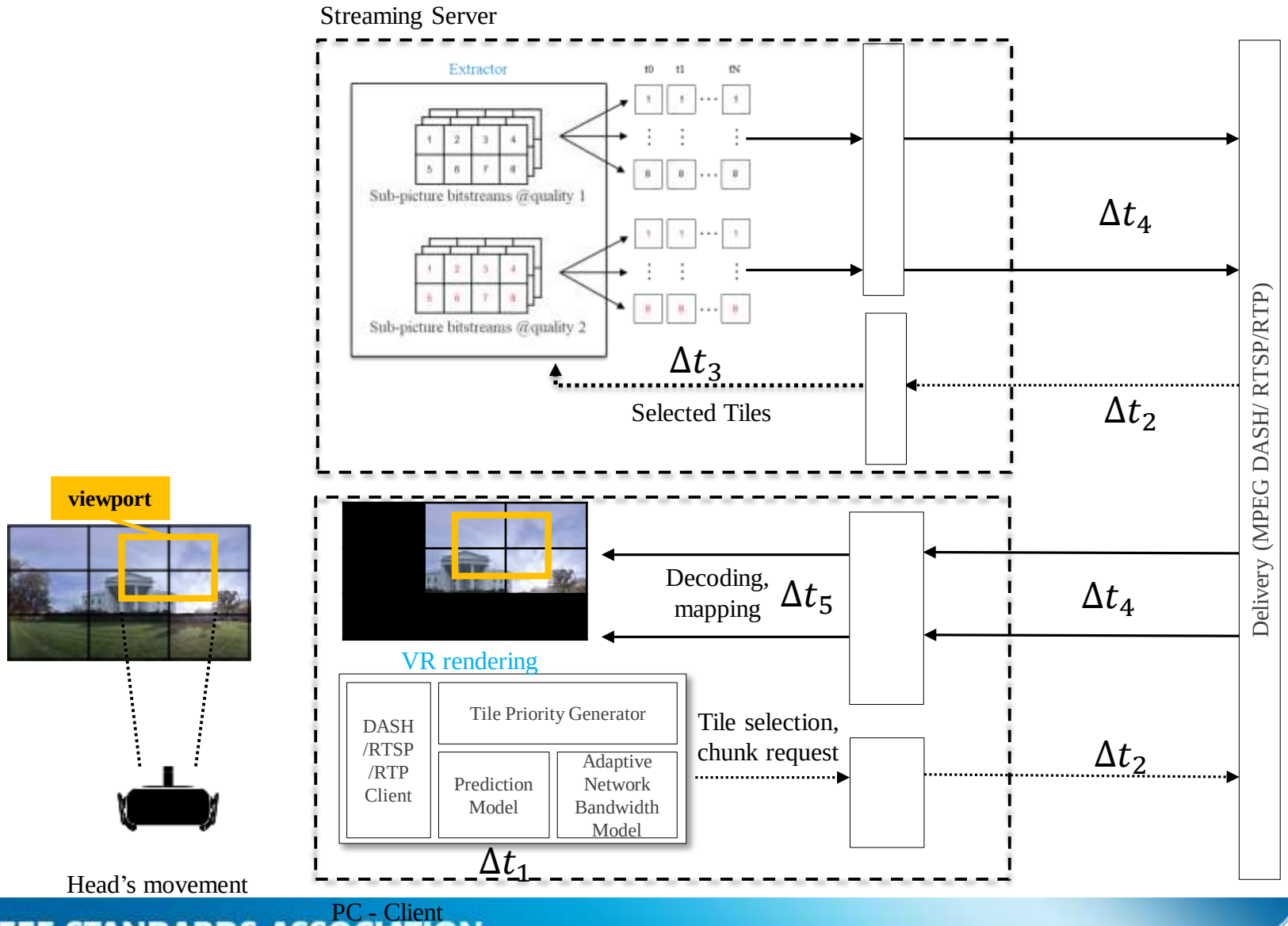
VR Tiled Streaming Latency

- 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

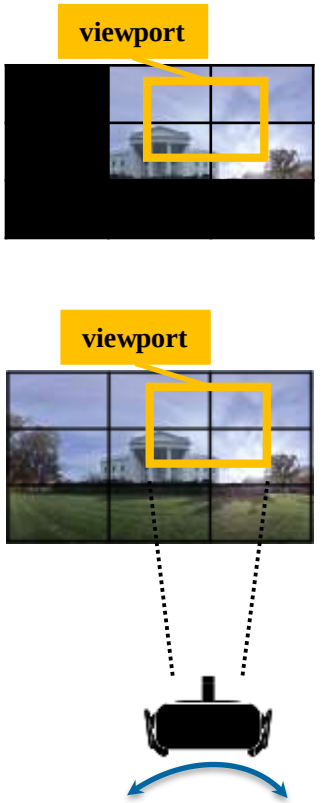


$$Total\_latency = \Delta t_1 + \Delta t_2 + \Delta t_3 + \Delta t_4 + \Delta t_5$$

VR Tiled Streaming System Architecture



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)