



Evaluation of Point Cloud Streaming and Rendering for VR-based Telepresence in the OR

Roland Fischer¹, Andre Mühlenbrock¹, Farin Kulapichitr¹,

Verena Nicole Uslar², Dirk Weyhe², Gabriel Zachmann¹

¹ University of Bremen, Bremen, Germany

² University Hospital for Visceral Surgery, University of Oldenburg, PIUS-Hospital, Oldenburg,
Germany

r.fischer@uni-bremen.de

EuroXR 2022 – Scientific Track

14-16 September, Stuttgart, Germany



Lecture Notes in
Computer Science

LNCS

LNAI

LNBI

- Telemedicine/-presence plays vital role

- Telemedicine/-presence plays vital role
 - Rural emergencies

- Telemedicine/-presence plays vital role
 - Rural emergencies
 - Limiting physical contact

- Telemedicine/-presence plays vital role
 - Rural emergencies
 - Limiting physical contact
 - Mentoring and education

- Telemedicine/-presence plays vital role
 - Rural emergencies
 - Limiting physical contact
 - Mentoring and education
- Mostly video-based ➡ inherently limited



Video systems in the OR [GettyImages]

- Telemedicine/-presence plays vital role
 - Rural emergencies
 - Limiting physical contact
 - Mentoring and education
- Mostly video-based → inherently limited
- VR+RGB-D solutions still with challenges



Video systems in the OR [GettyImages]



VR in Medicine [Arch Virtual]

Motivation

- Ideal telepresence system:
 - Easy to use/set up



Vision [Digital-Operating-Theatre of AMRC]

Motivation

- Ideal telepresence system:
 - Easy to use/set up
 - Multi-user with personalized avatars



Vision [Digital-Operating-Theatre of AMRC]

- Ideal telepresence system:
 - Easy to use/set up
 - Multi-user with personalized avatars
 - High fidelity 3D scenes, real-time



Vision [Digital-Operating-Theatre of AMRC]

- Ideal telepresence system:
 - Easy to use/set up
 - Multi-user with personalized avatars
 - High fidelity 3D scenes, real-time



- High immersion & presence



Vision [Digital-Operating-Theatre of AMRC]

Related Work

- Telepresence:

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]
 - Low performance (10 Hz), high latency (300-400 ms) [Thoravi Kumaravel19, Roth21]

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]
 - Low performance (10 Hz), high latency (300-400 ms) [Thoravi Kumaravel19, Roth21]
 - Multiple PCs [Roth21]

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]
 - Low performance (10 Hz), high latency (300-400 ms) [Thoravi Kumaravel19, Roth21]
 - Multiple PCs [Roth21]
- 3D visualization/rendering:

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]
 - Low performance (10 Hz), high latency (300-400 ms) [Thoravi Kumaravel19, Roth21]
 - Multiple PCs [Roth21]
- 3D visualization/rendering:
 - High quality volumetric reconstruction, slow, complicated [Dou17, Orts03]

- Telepresence:
 - Limited compression, high bandwidth [Antón17, Boehlen20]
 - Single camera/user, no avatars [Antón17, Boehlen20]
 - Low performance (10 Hz), high latency (300-400 ms) [Thoravi Kumaravel19, Roth21]
 - Multiple PCs [Roth21]
- 3D visualization/rendering:
 - High quality volumetric reconstruction, slow, complicated [Dou17, Orts03]
 - Faster point cloud splatting, no continuous surface [Gamelin21, Yu21]

Our Contributions

- A multi-user VR-based telepresence system
 - Easy to set up, single PC

Our Contributions

- A multi-user VR-based telepresence system
 - Easy to set up, single PC
 - Custom splat- and mesh-based point cloud renderers

- A multi-user VR-based telepresence system
 - Easy to set up, single PC
 - Custom splat- and mesh-based point cloud renderers
 - Avatar face reconstruction prototype

- A multi-user VR-based telepresence system
 - Easy to set up, single PC
 - Custom splat- and mesh-based point cloud renderers
 - Avatar face reconstruction prototype
- A modular low-latency multi-camera RGB-D streaming pipeline

- A multi-user VR-based telepresence system
 - Easy to set up, single PC
 - Custom splat- and mesh-based point cloud renderers
 - Avatar face reconstruction prototype
- A modular low-latency multi-camera RGB-D streaming pipeline
 - Includes pre-processing and compression

Our Contributions

- A multi-user VR-based telepresence system
 - Easy to set up, single PC
 - Custom splat- and mesh-based point cloud renderers
 - Avatar face reconstruction prototype
- A modular low-latency multi-camera RGB-D streaming pipeline
 - Includes pre-processing and compression
- Extensive qualitative evaluations including 2 user studies



System Overview

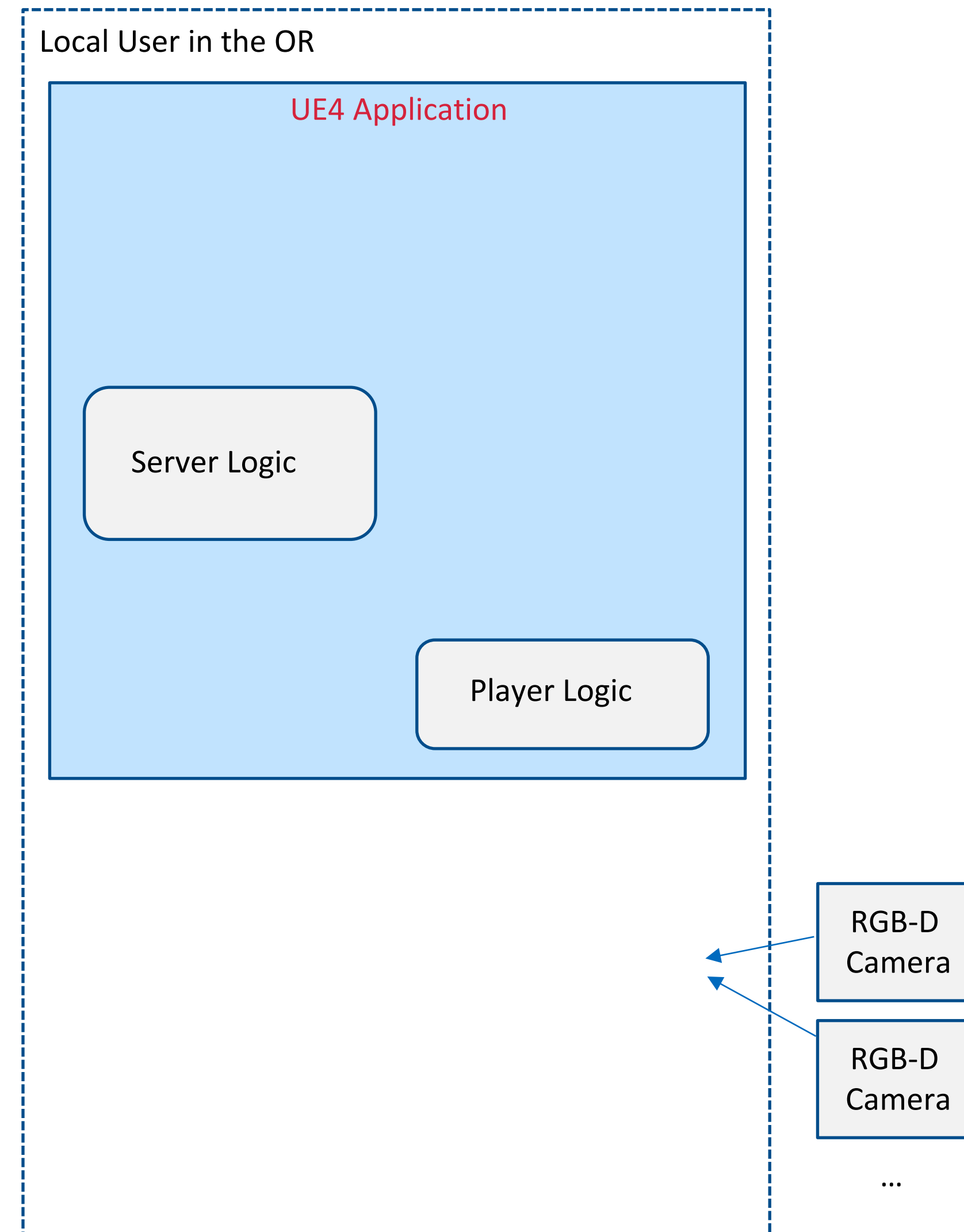


Server in the OR

System Overview



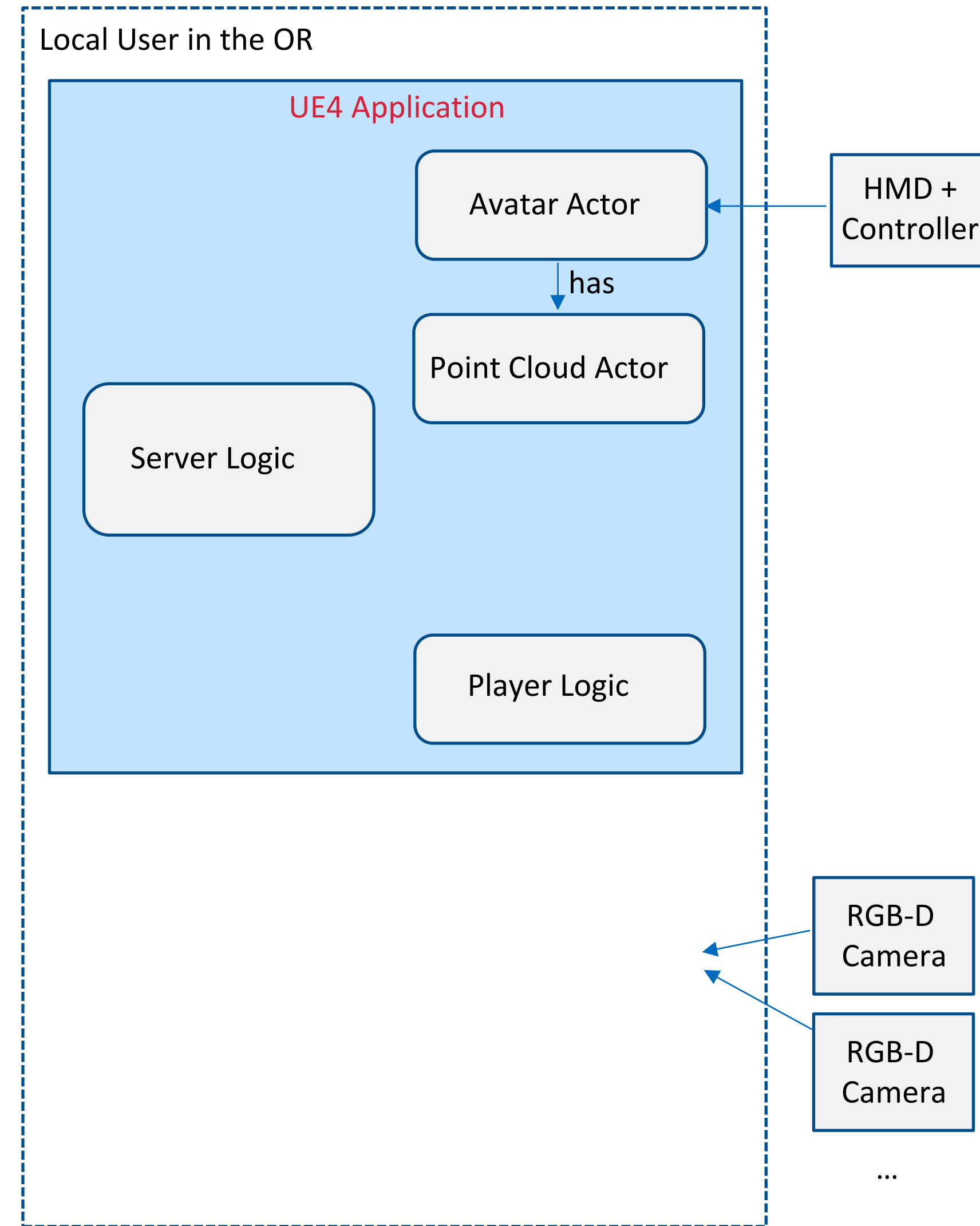
Server in the OR



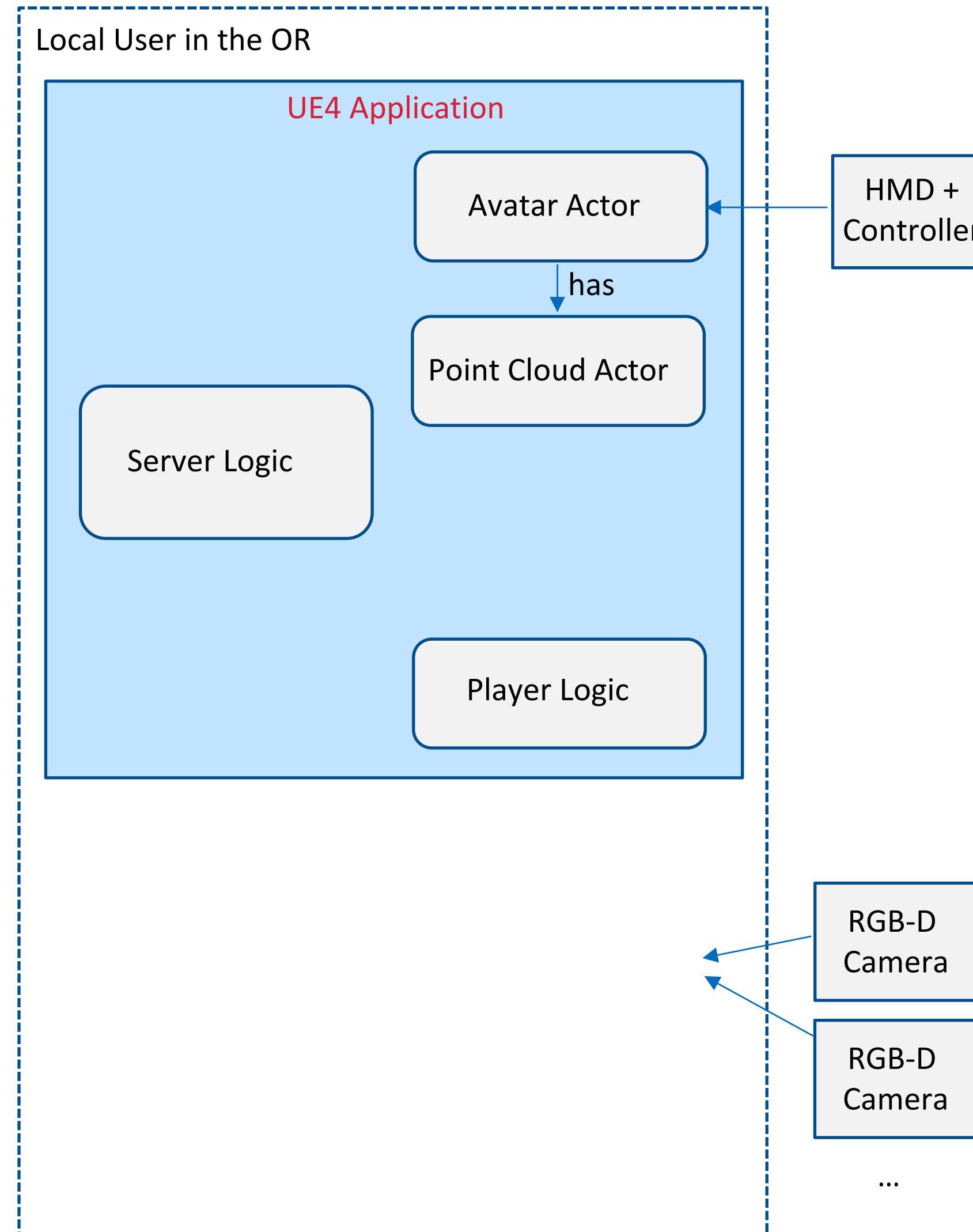
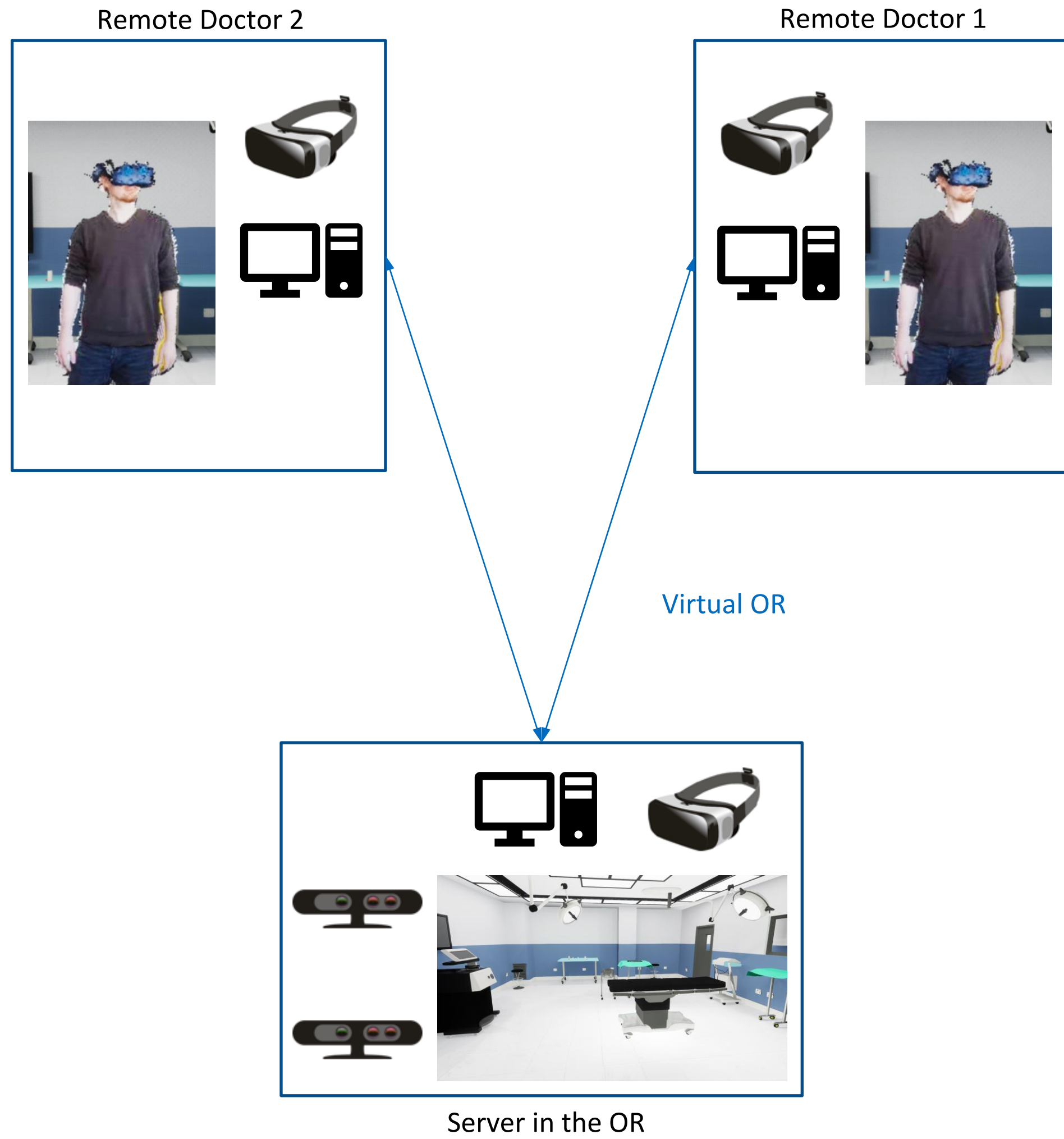
System Overview



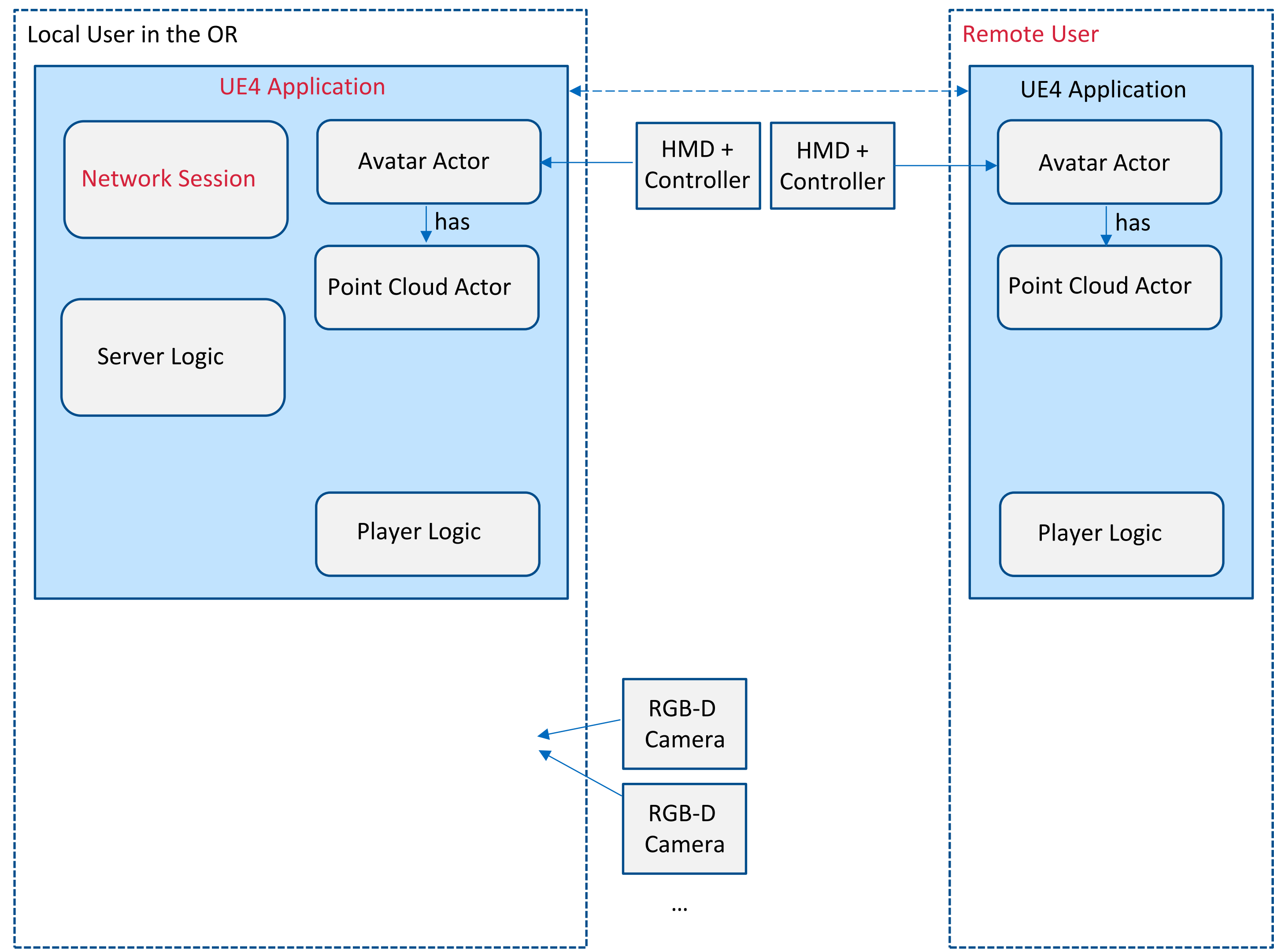
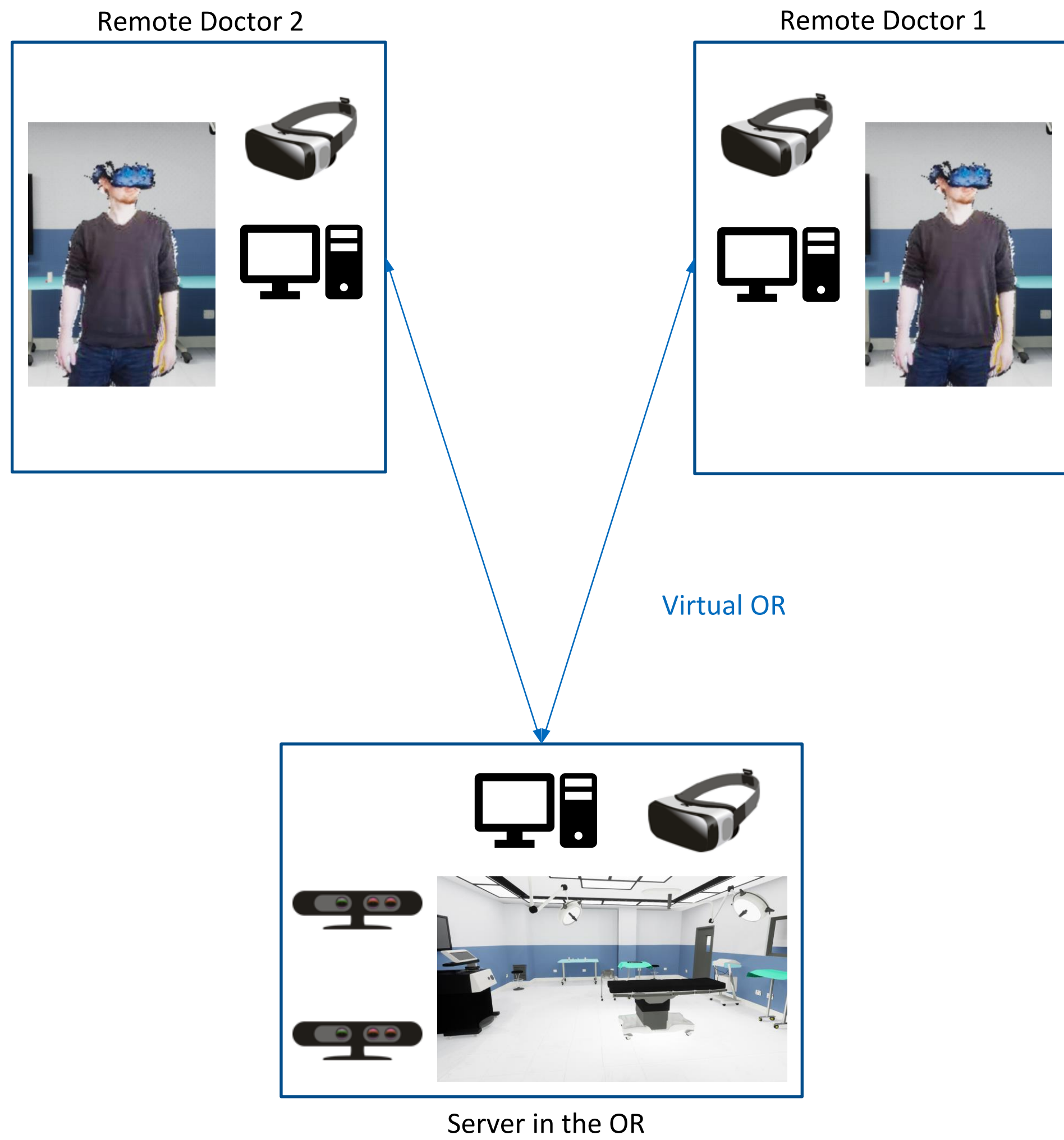
Server in the OR



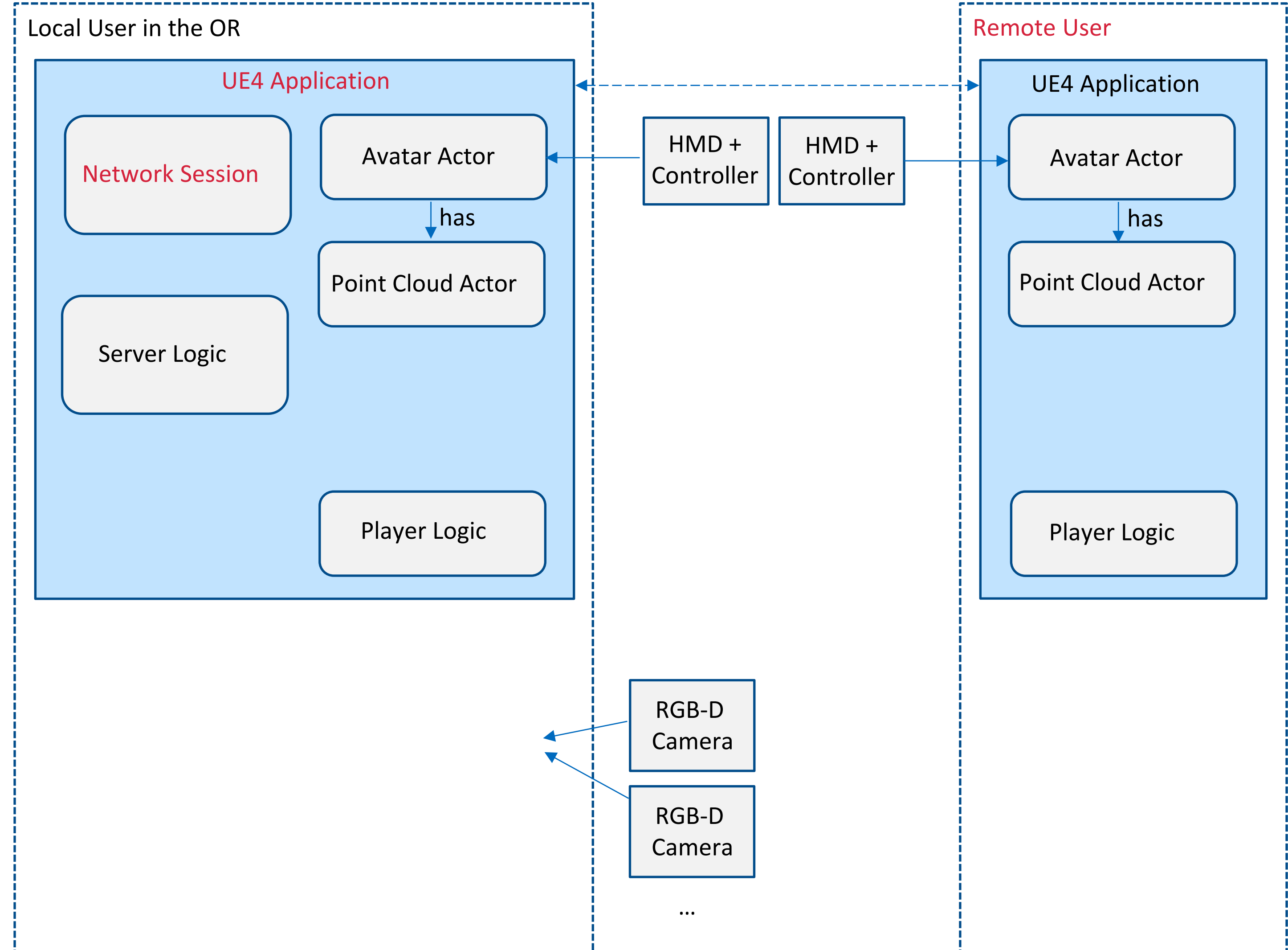
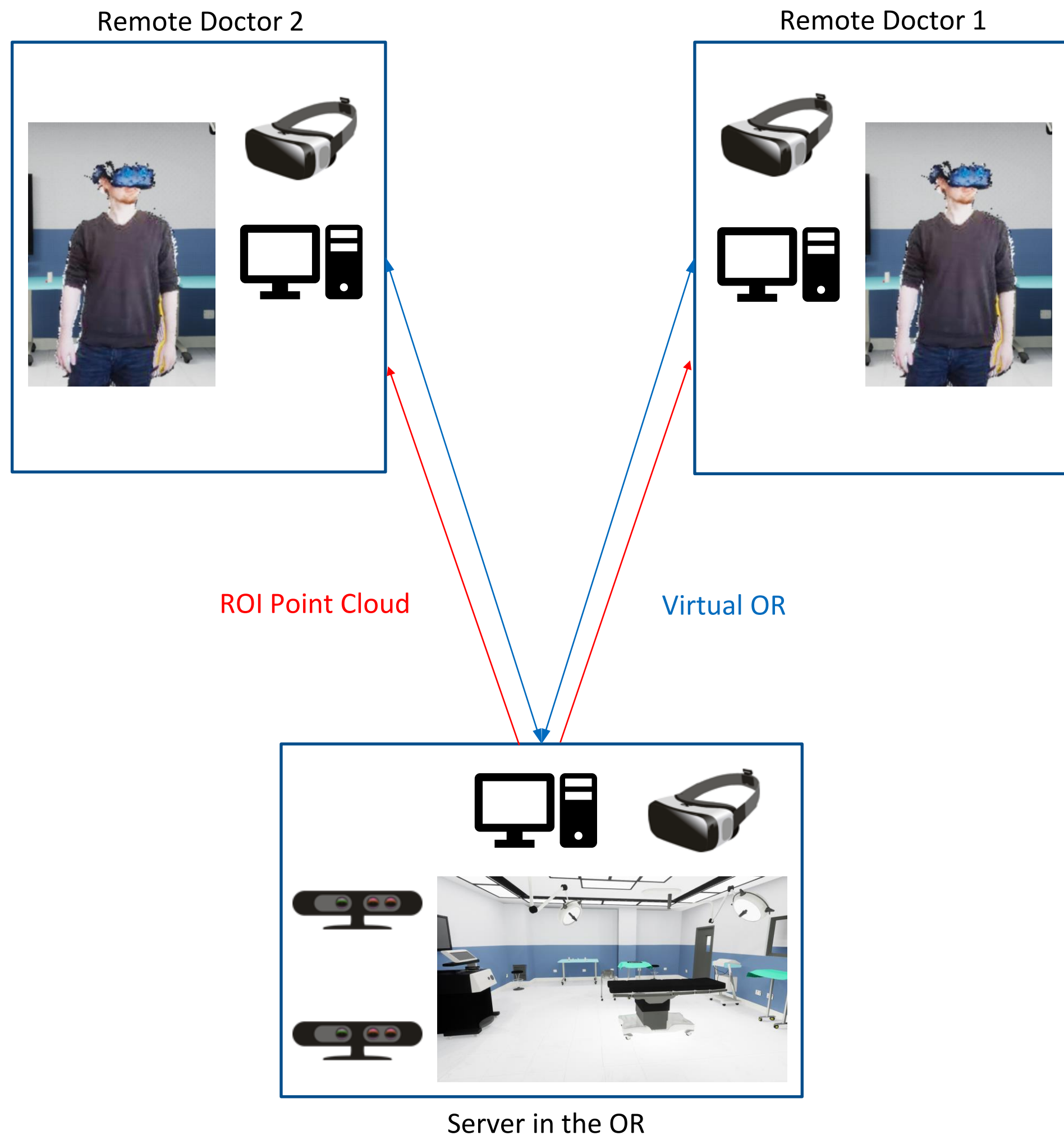
System Overview



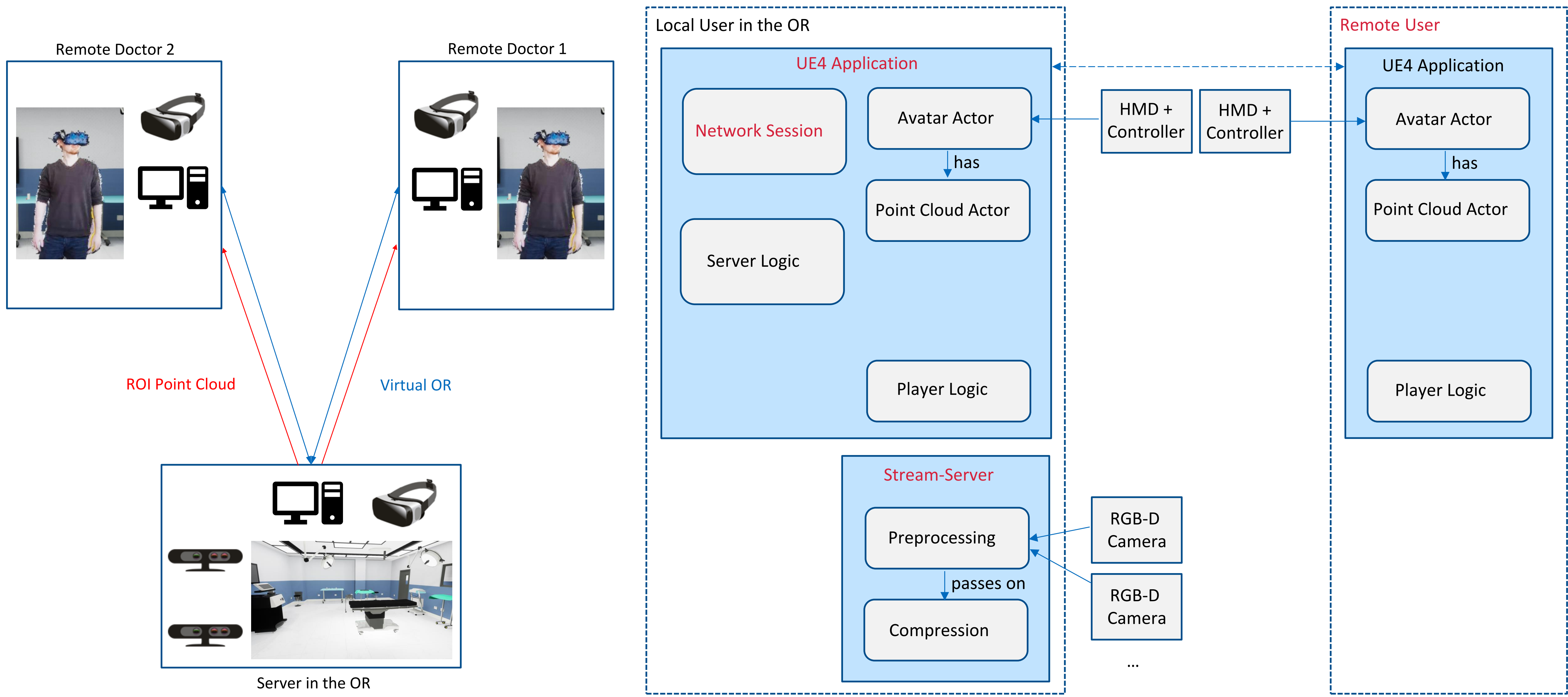
System Overview



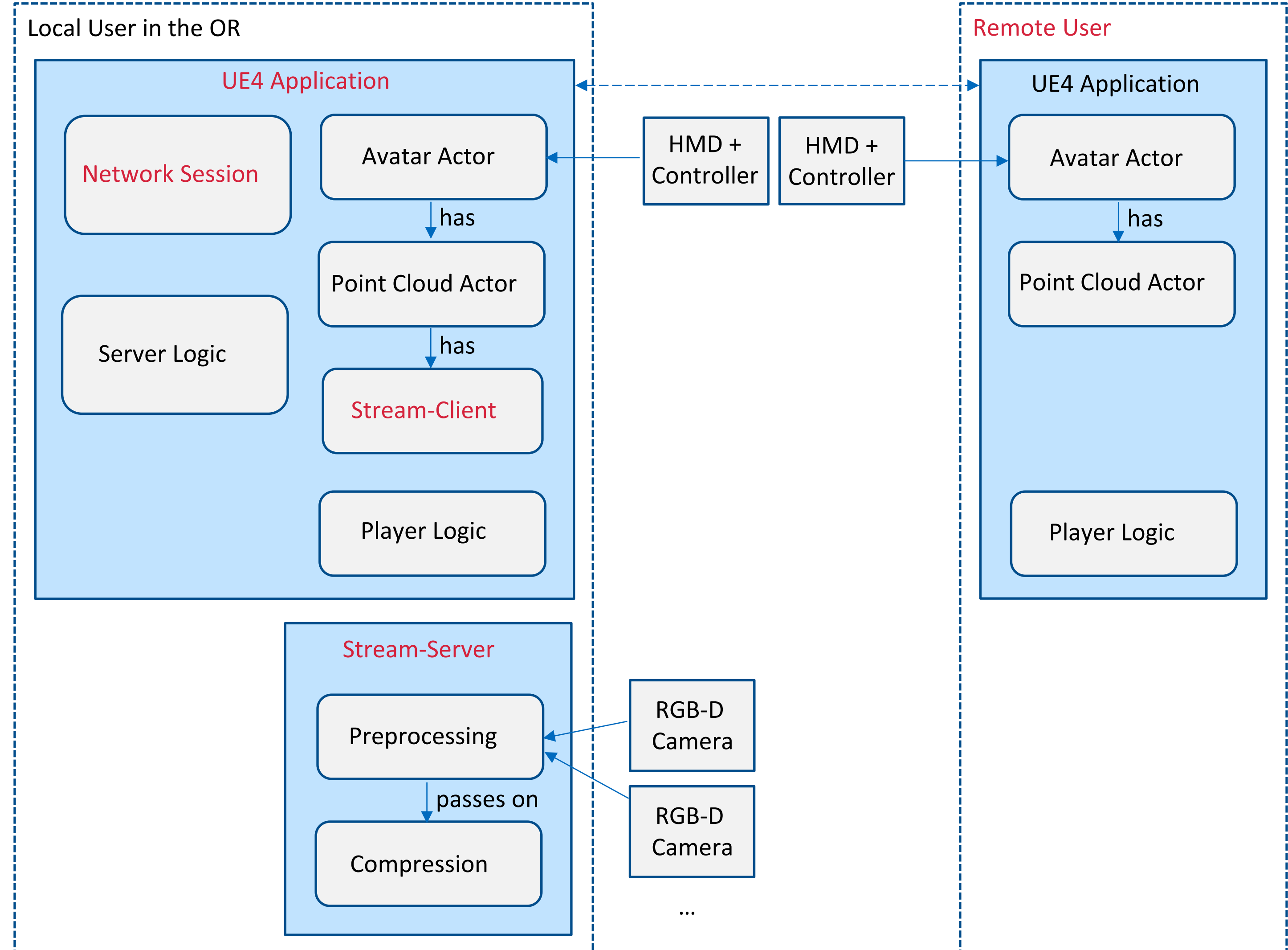
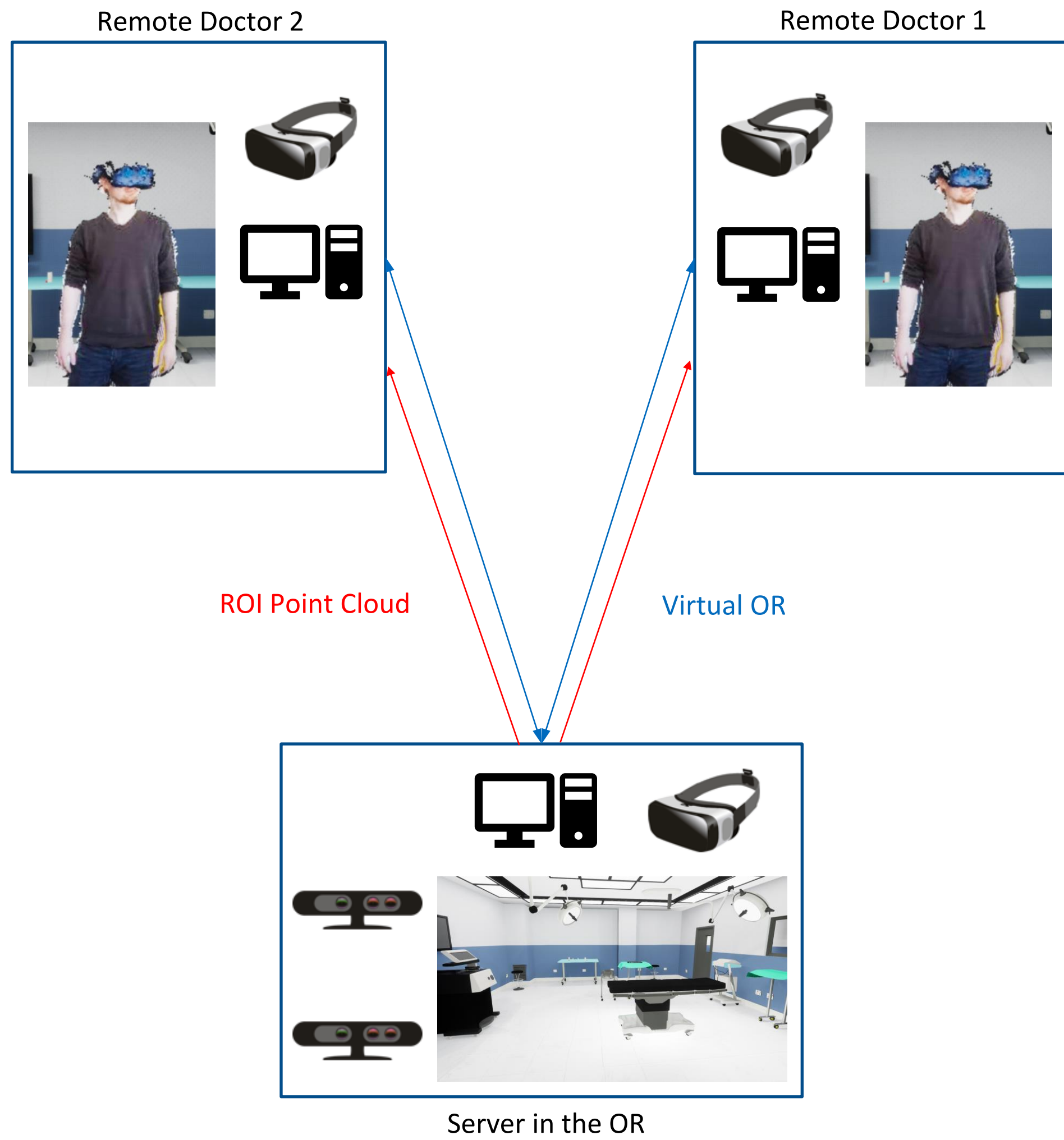
System Overview

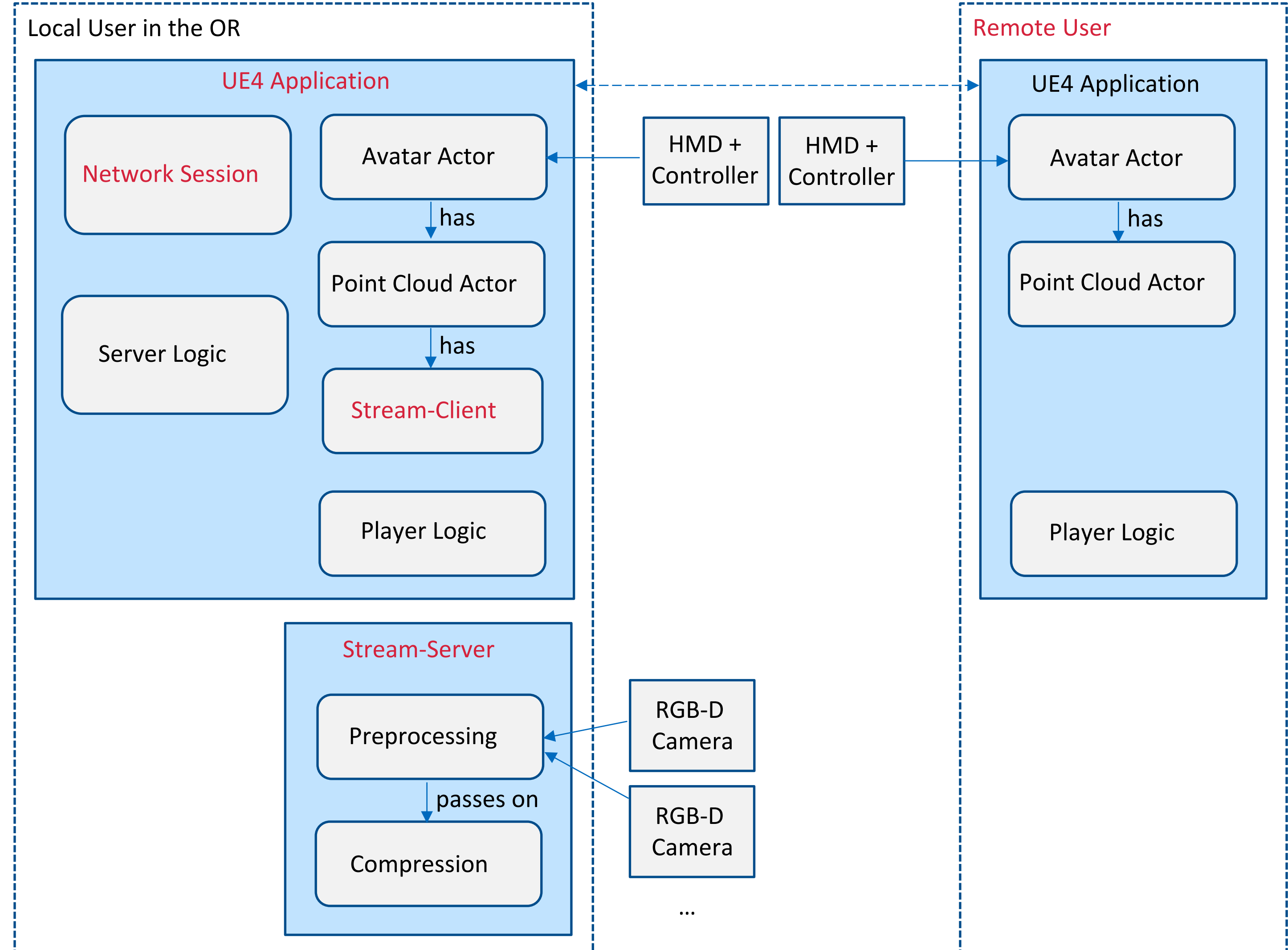
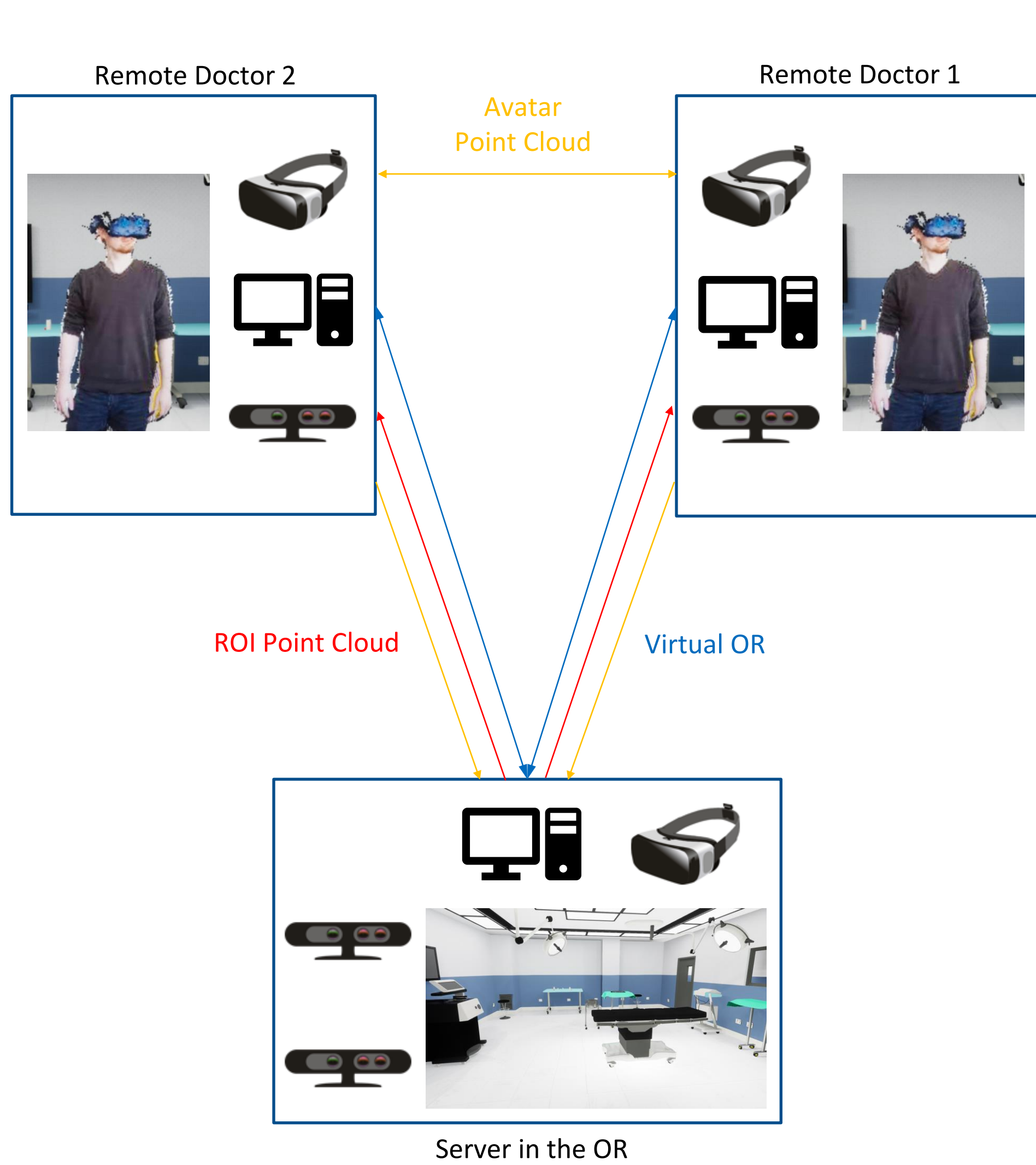


System Overview

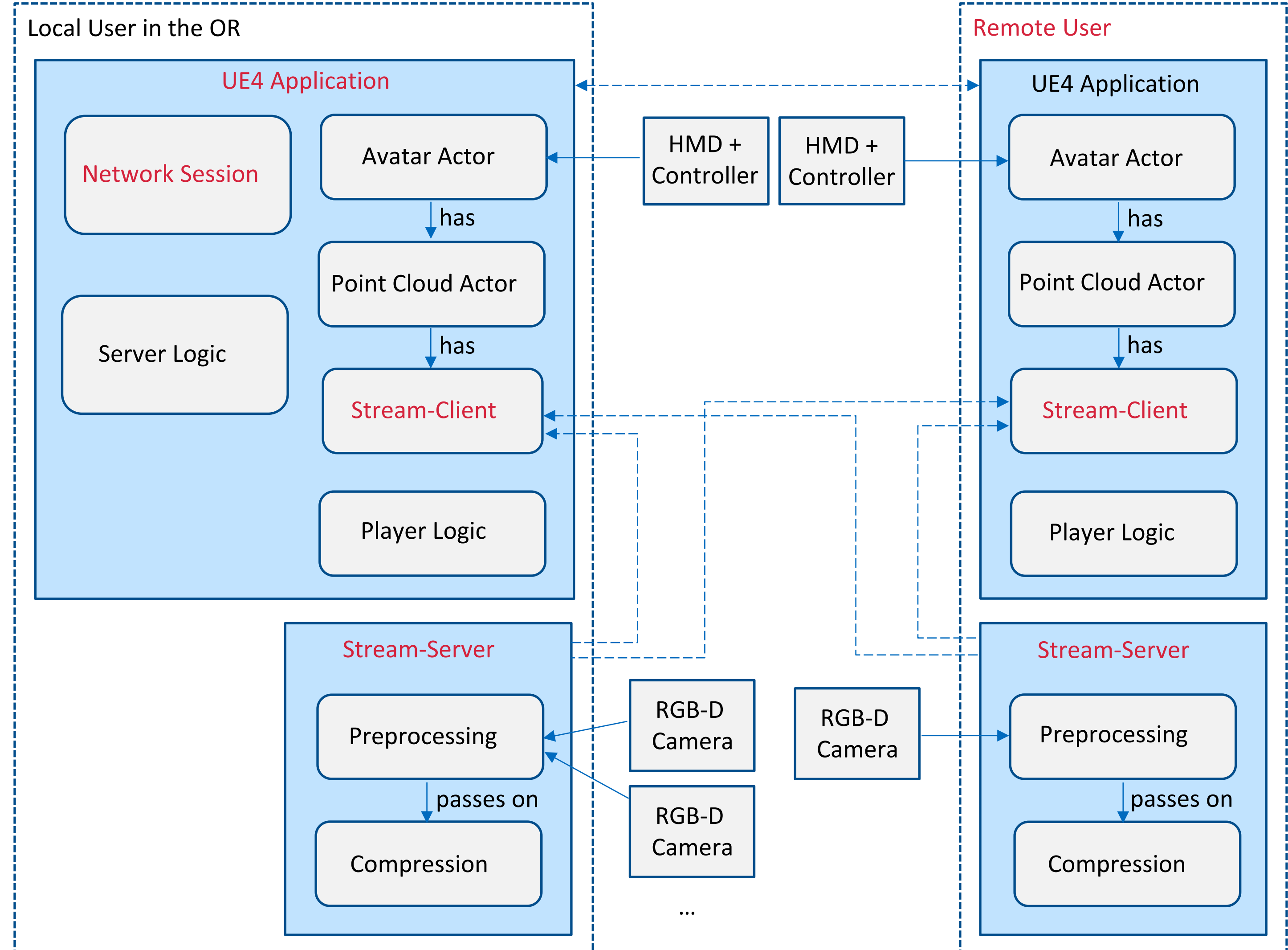
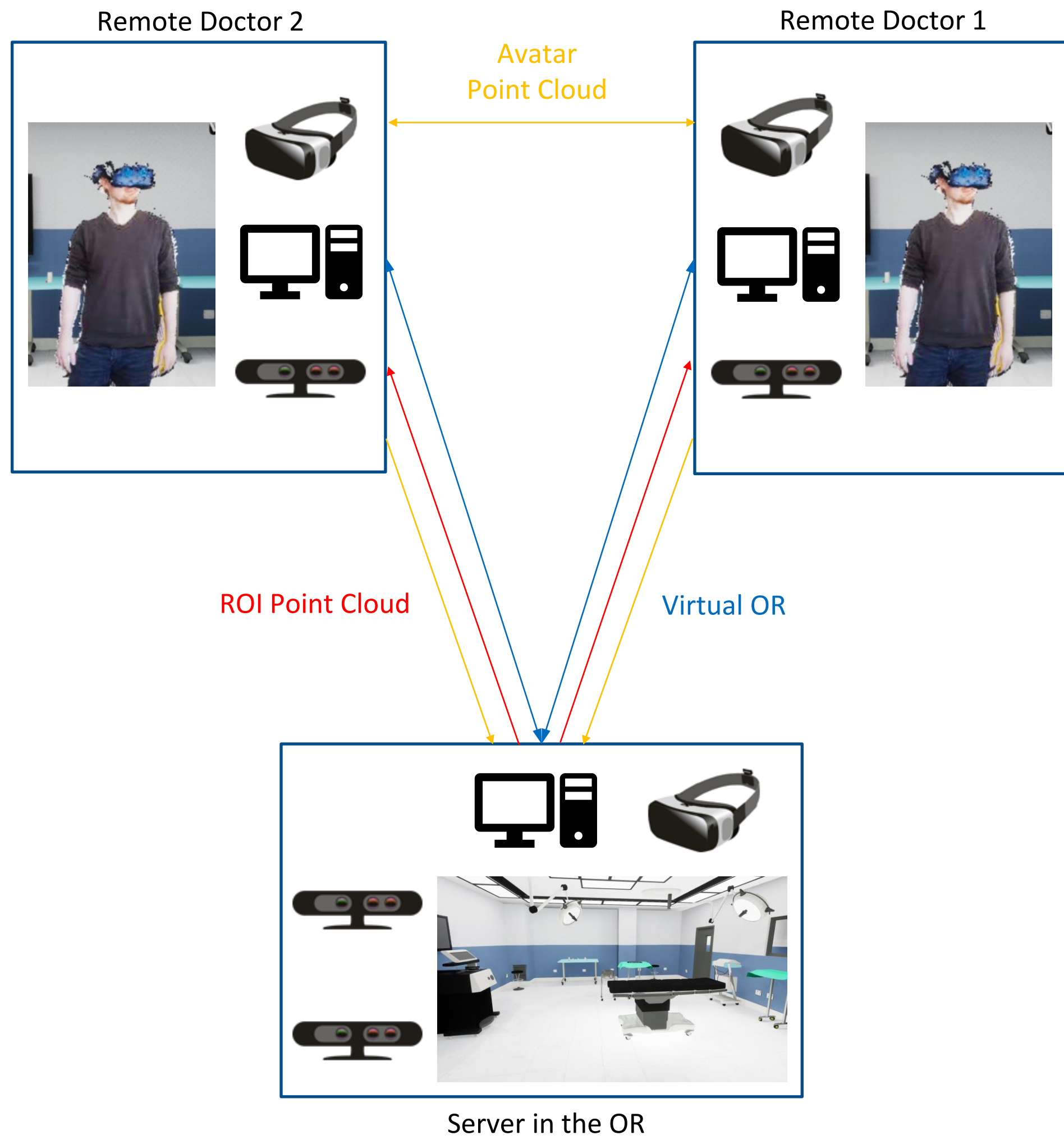


System Overview





System Overview

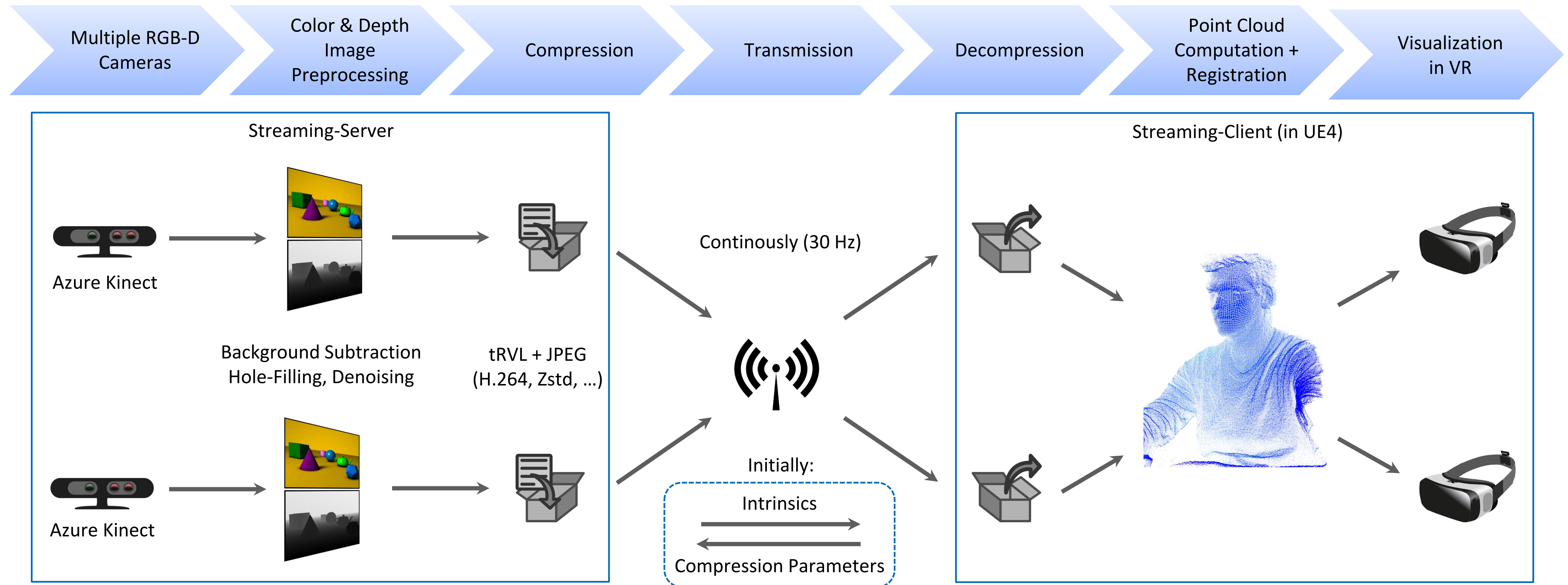


Point Cloud Streaming

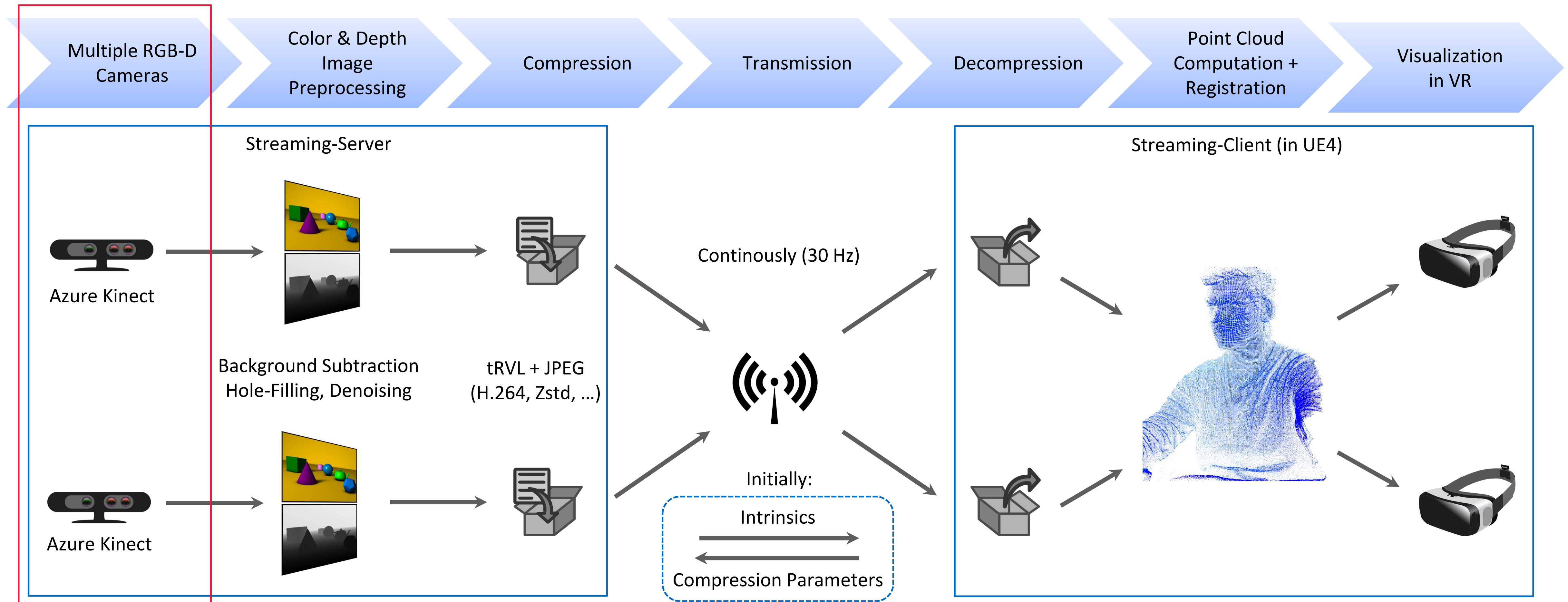
- Based on DynCam [Schröder18]

Point Cloud Streaming

- Based on DynCam [Schröder18]

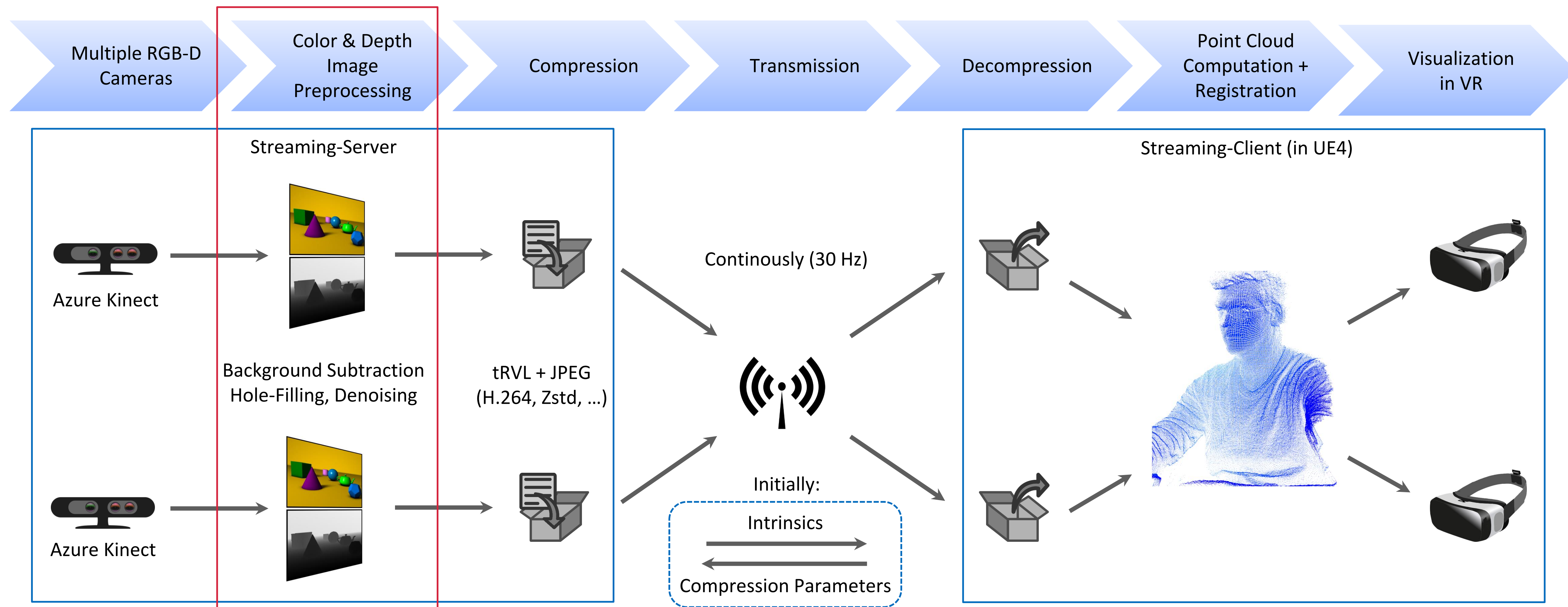


- Based on DynCam [Schröder18]



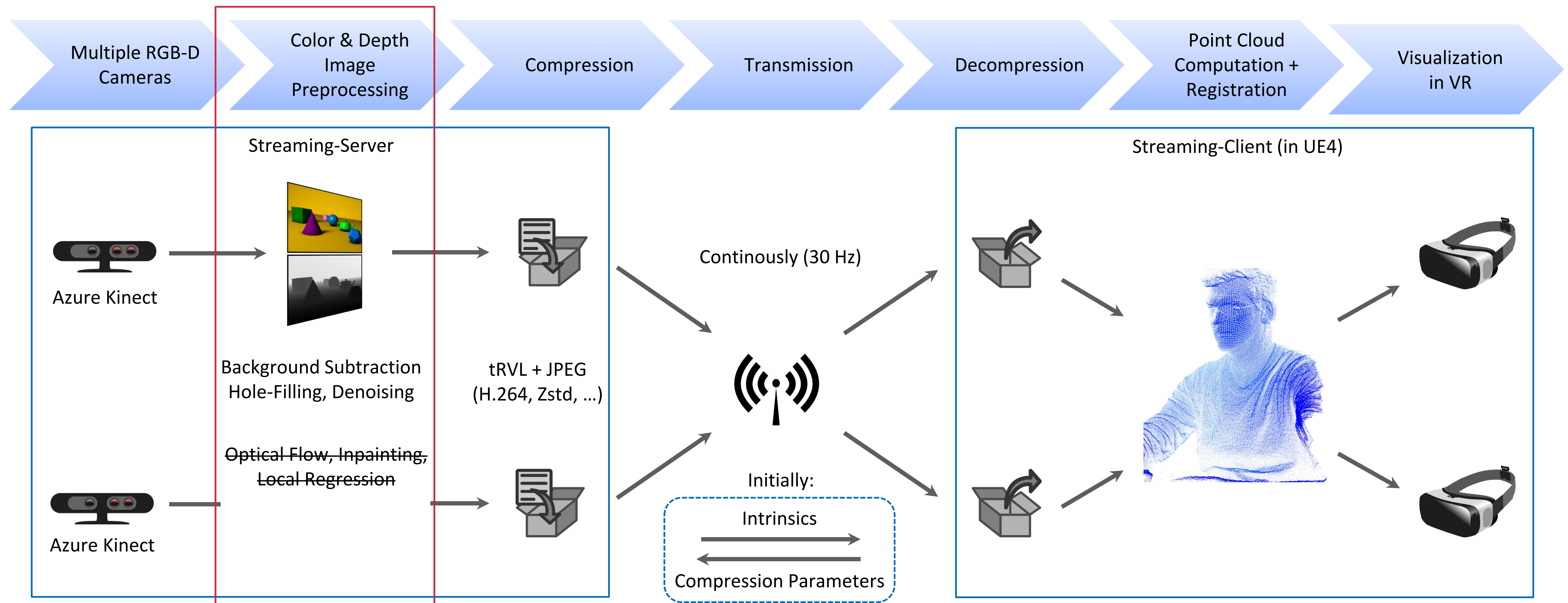
Point Cloud Streaming

- Based on DynCam [Schröder18]



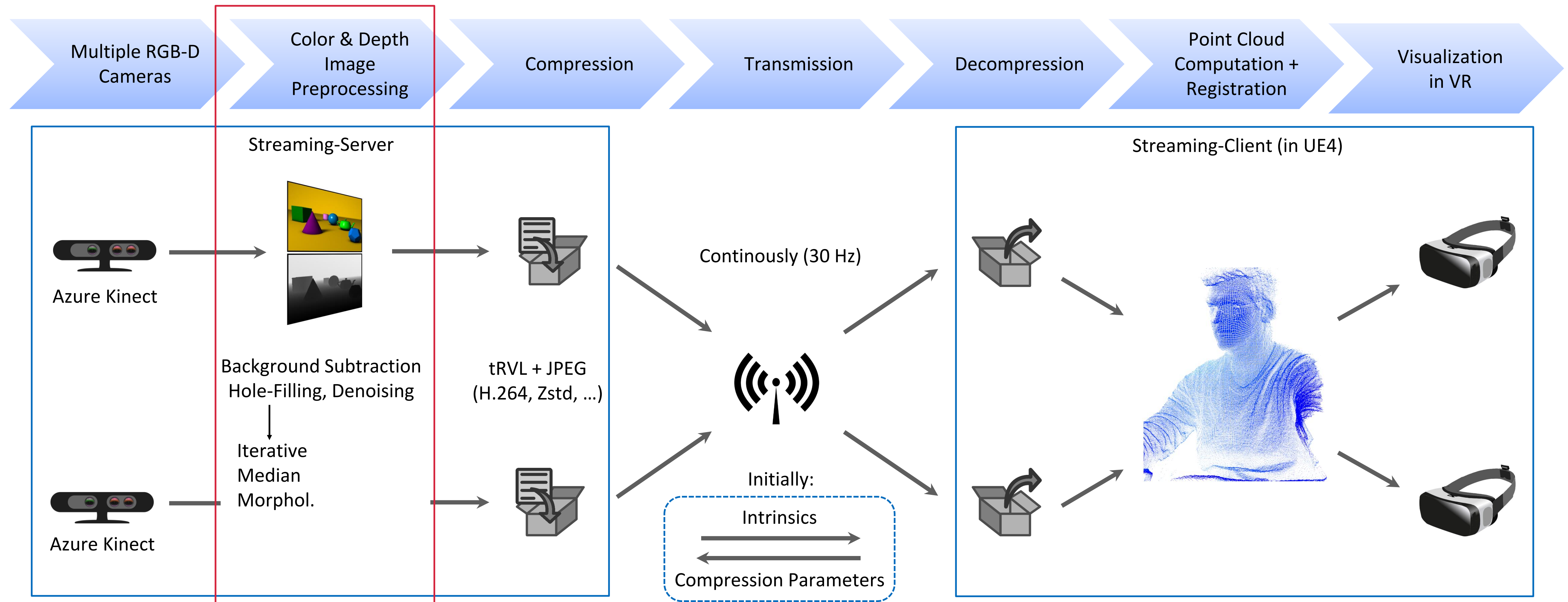
Point Cloud Streaming

- Based on DynCam [Schröder18]



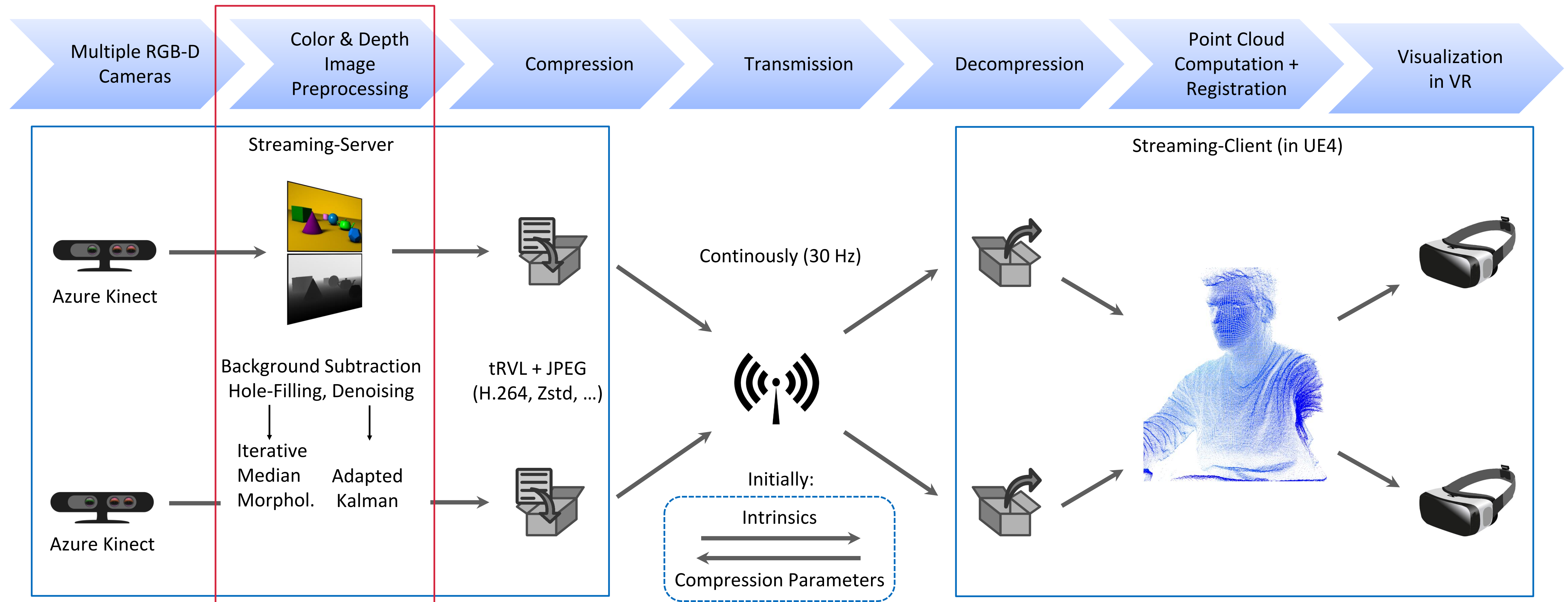
Point Cloud Streaming

- Based on DynCam [Schröder18]



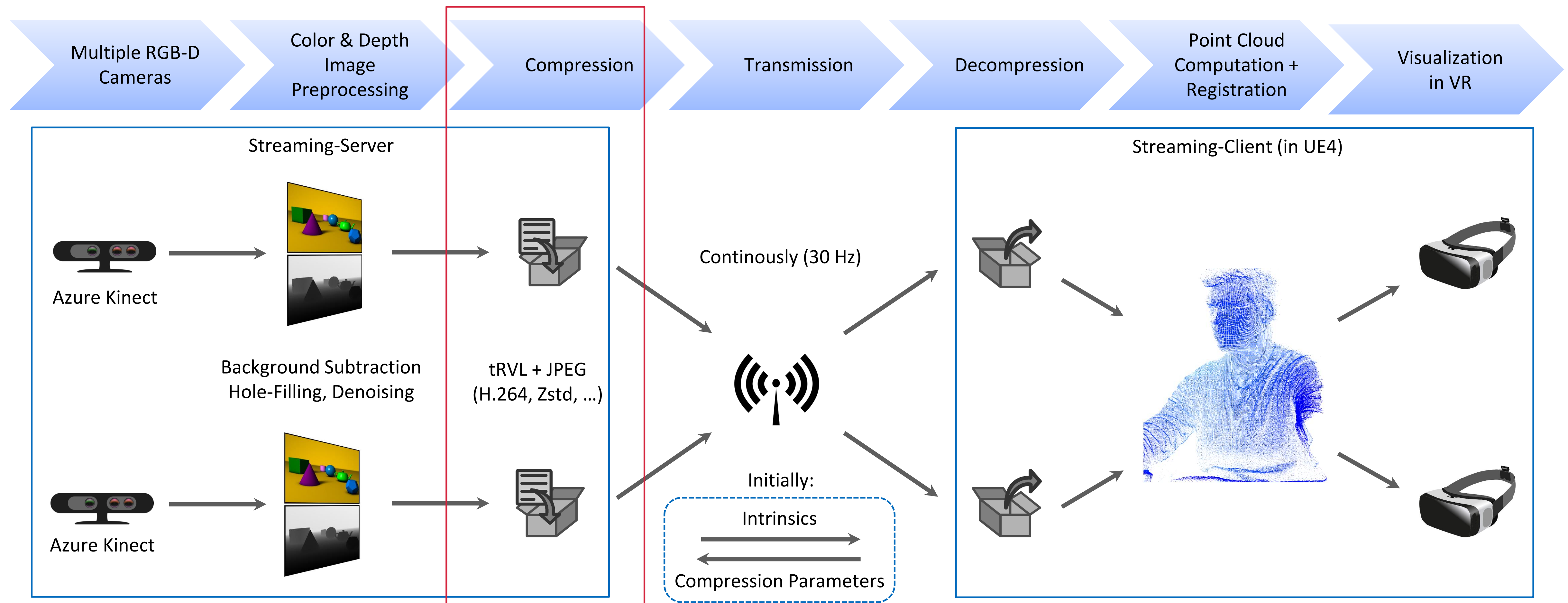
Point Cloud Streaming

- Based on DynCam [Schröder18]



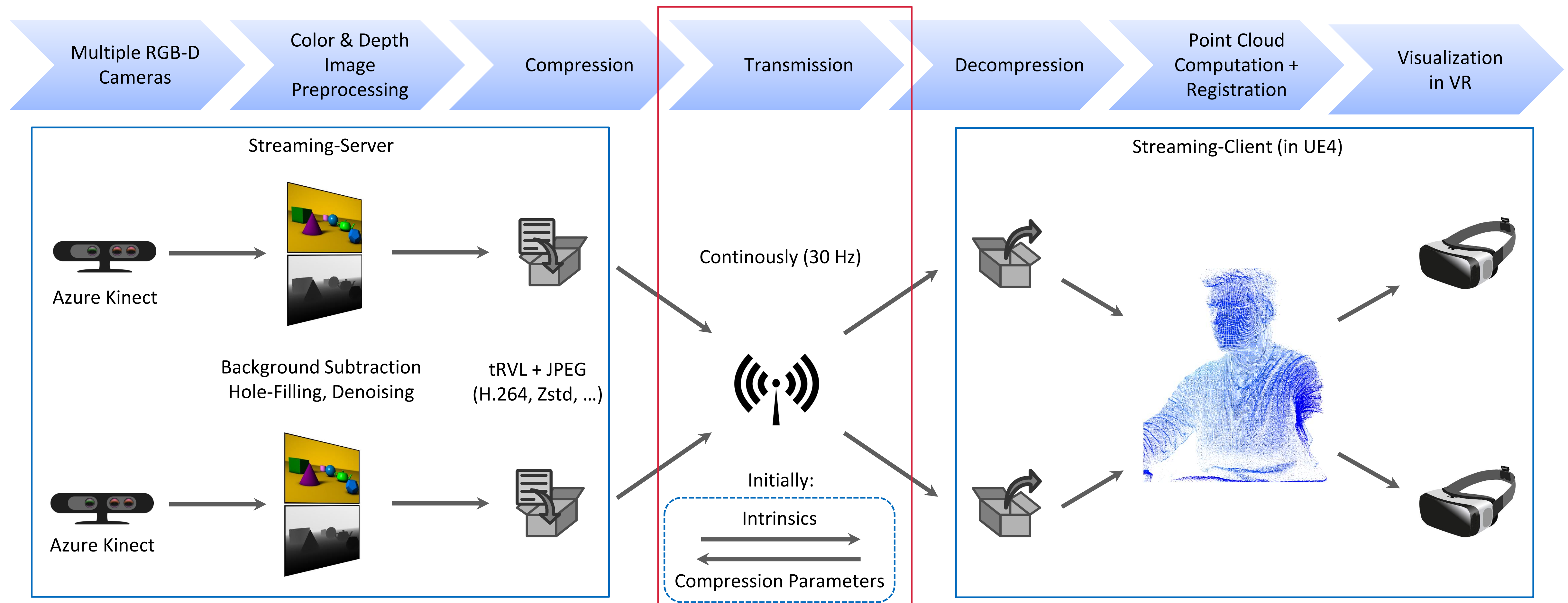
Point Cloud Streaming

- Based on DynCam [Schröder18]



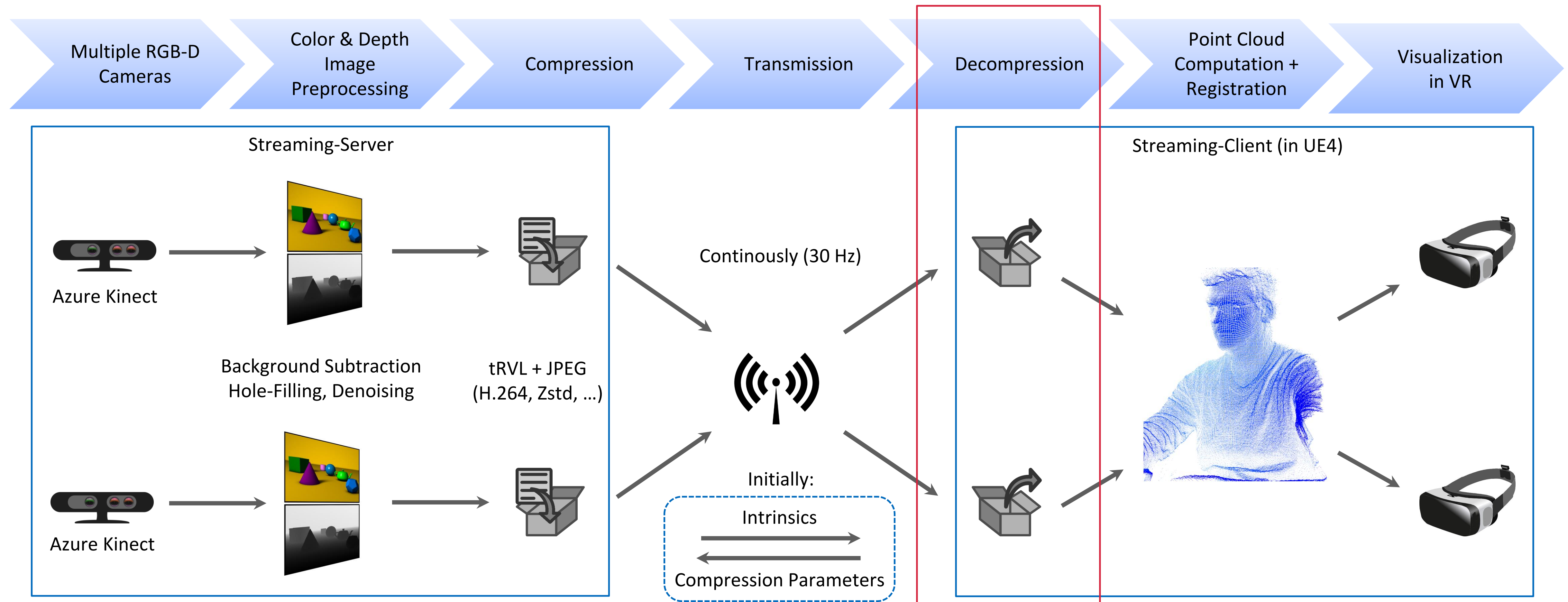
Point Cloud Streaming

- Based on DynCam [Schröder18]



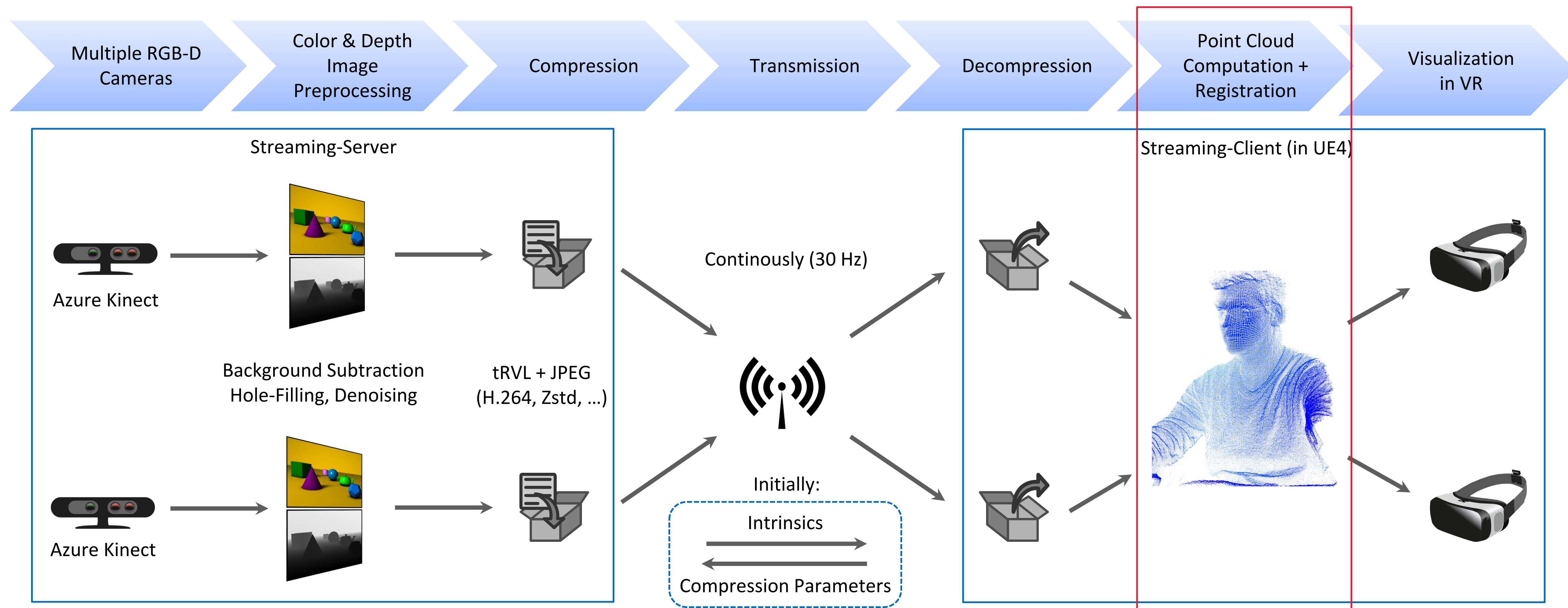
Point Cloud Streaming

- Based on DynCam [Schröder18]



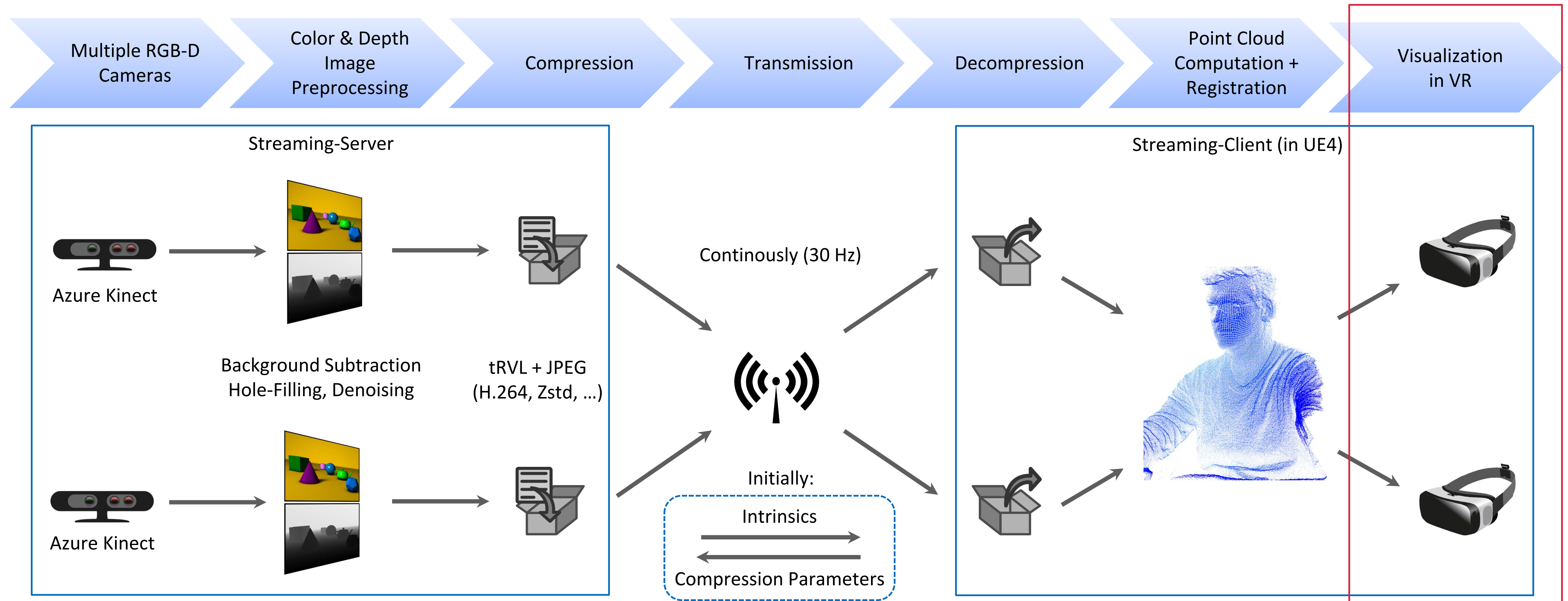
Point Cloud Streaming

- Based on DynCam [Schröder18]



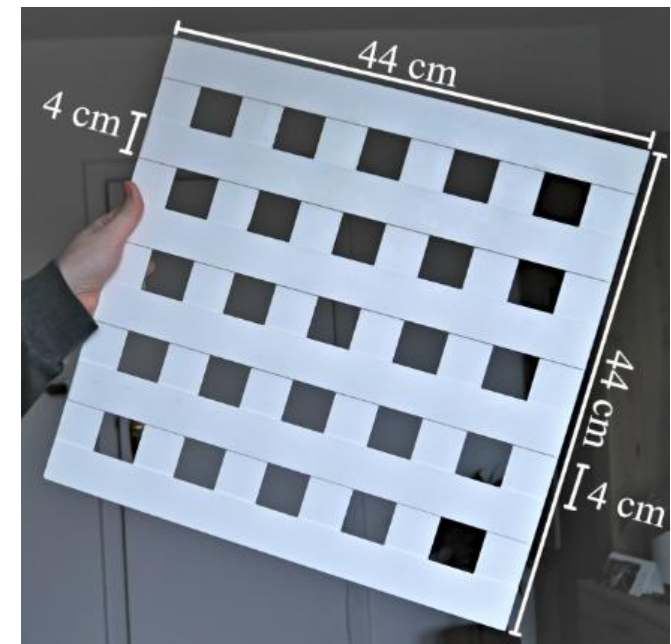
Point Cloud Streaming

- Based on DynCam [Schröder18]



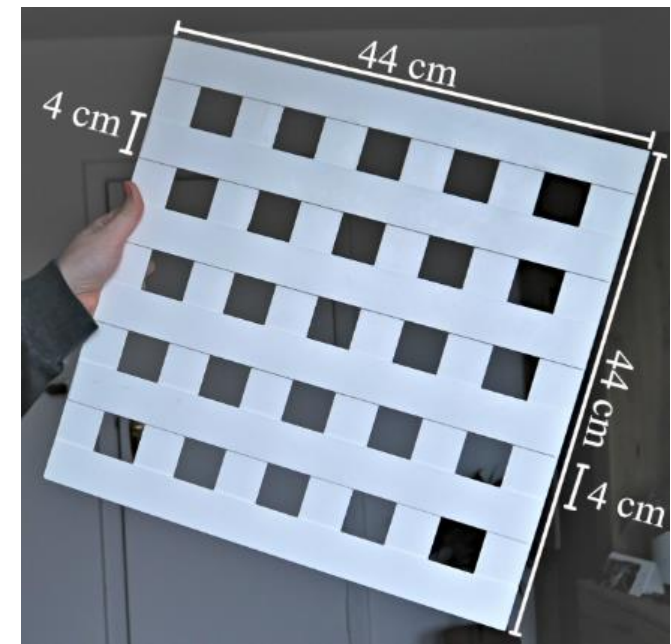
Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]'s method



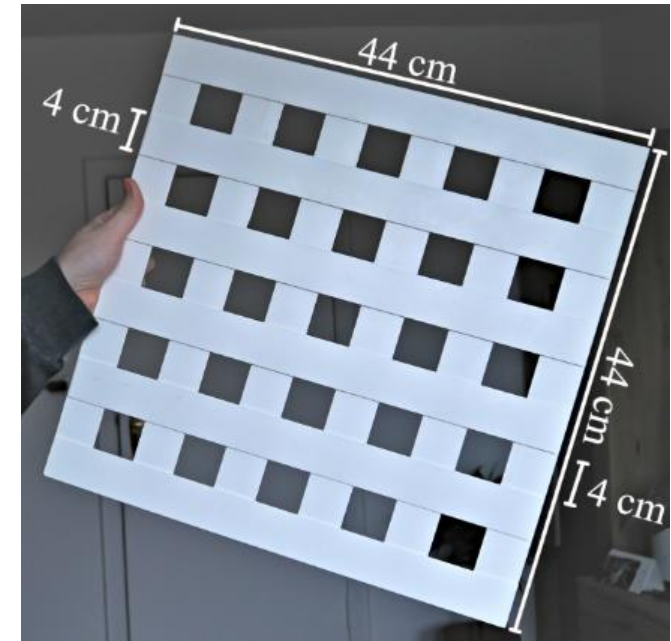
Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing



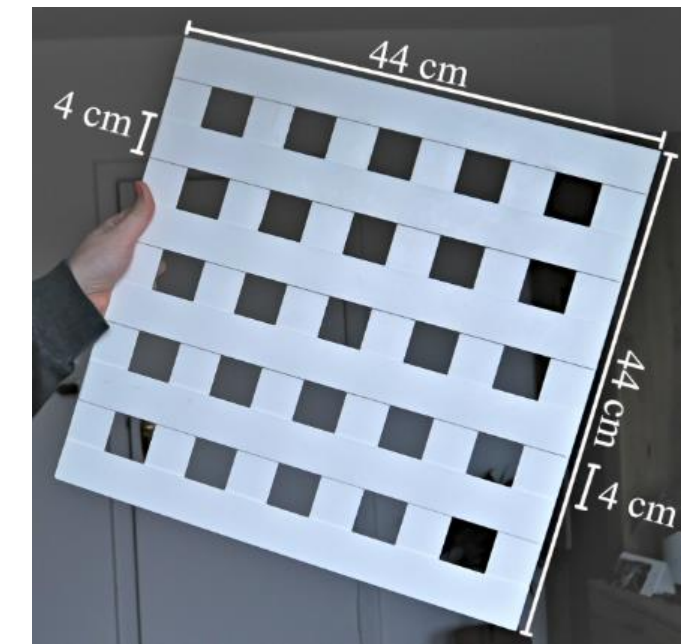
Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]'s method
- Dyncam's rendering and UE4 LiDAR PC plugin not convincing
➔ 2 own solutions



Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions

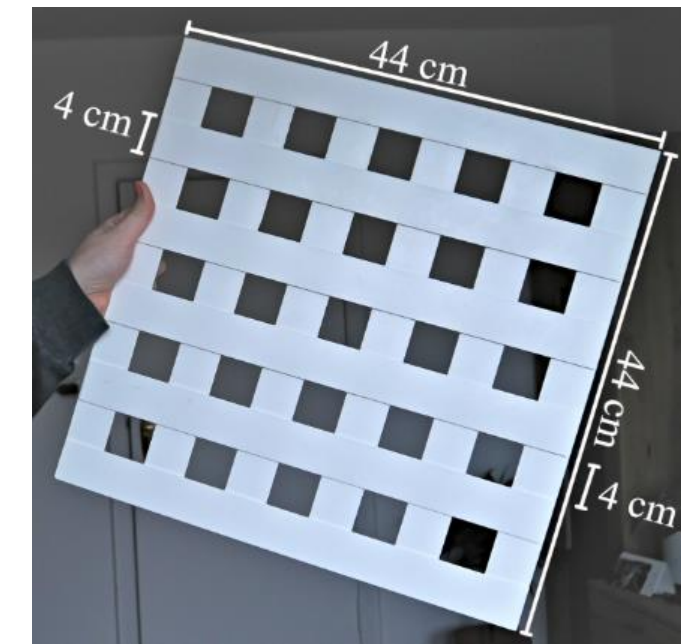


- Splatting:

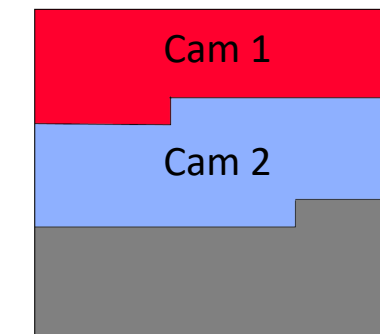


Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions

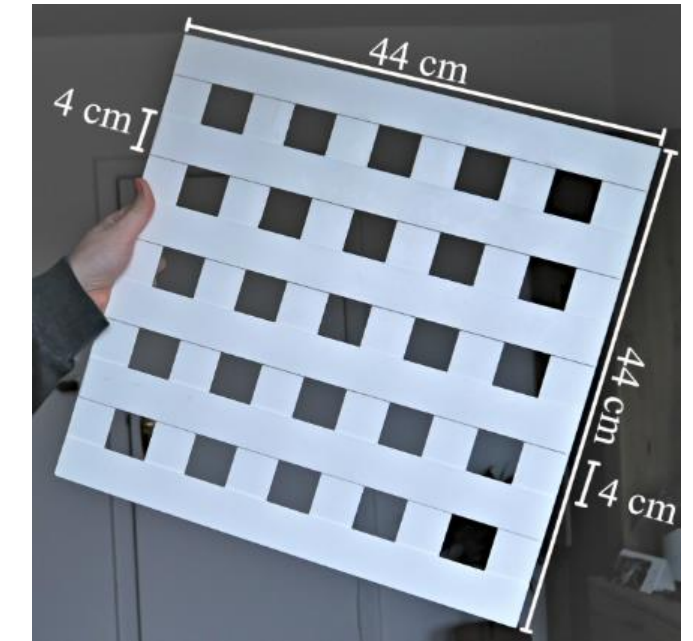


- Splatting:



Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions

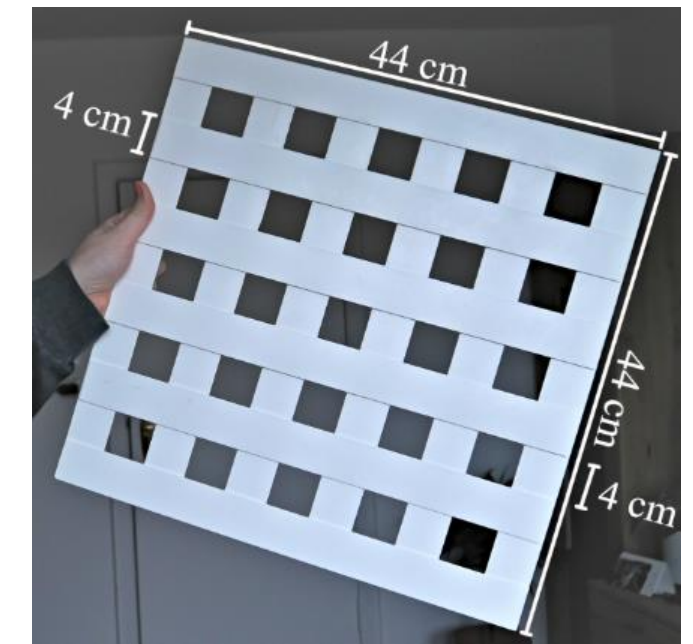


- Splatting:



Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions

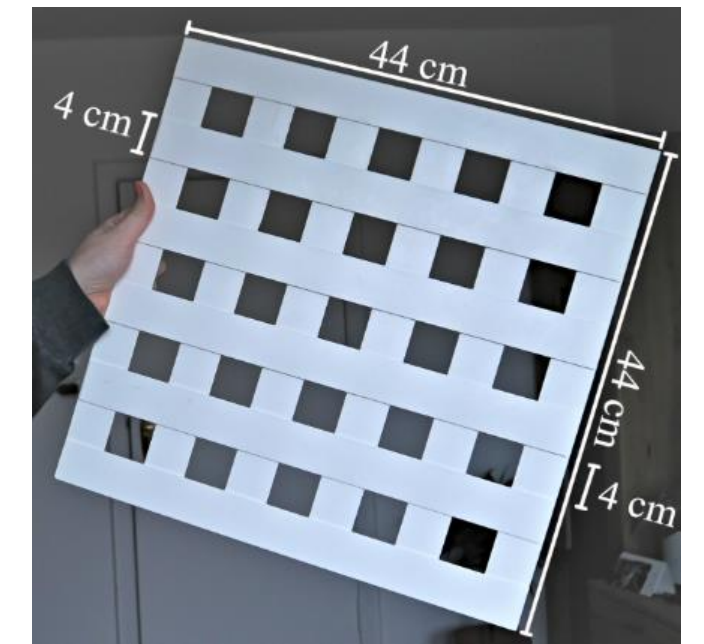


- Splatting:



Point Cloud Registration and Rendering

- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions

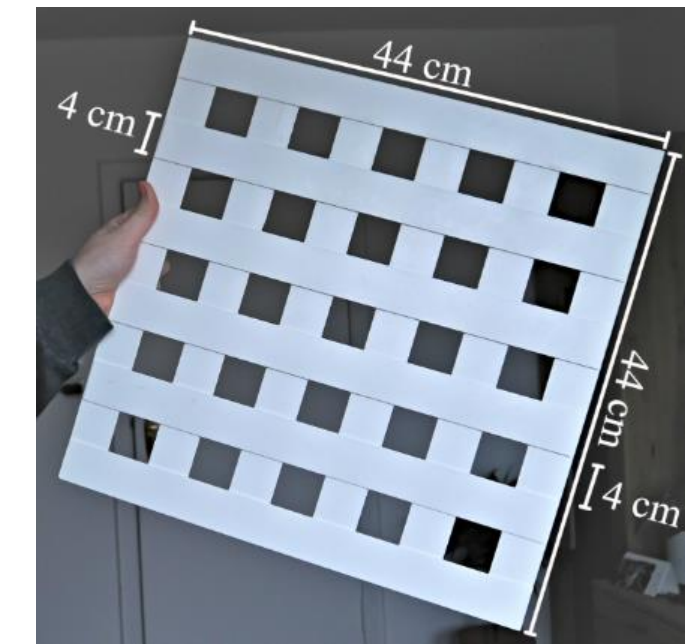


- Splatting:



Point Cloud Registration and Rendering

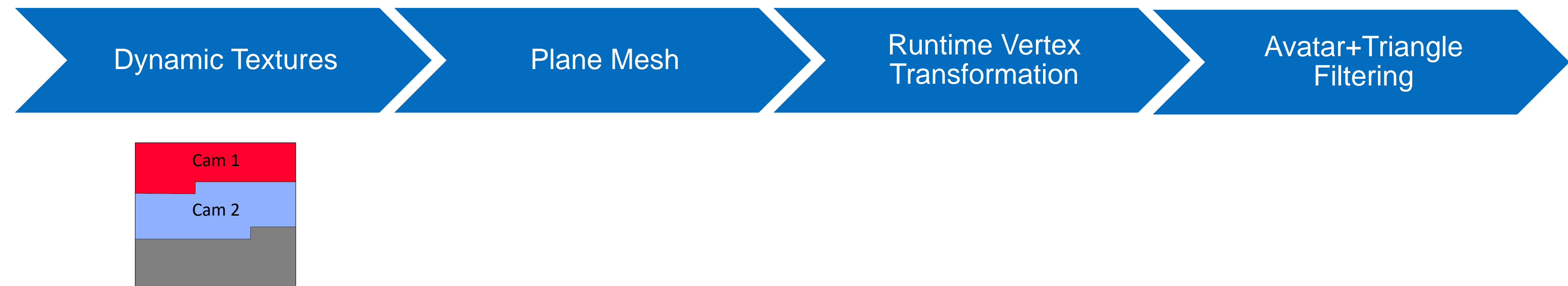
- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions



- Splatting:

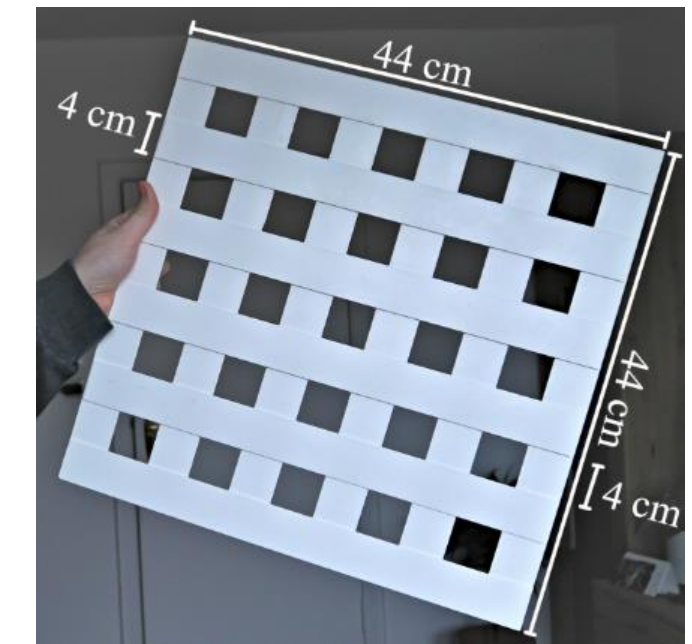


- Mesh:



Point Cloud Registration and Rendering

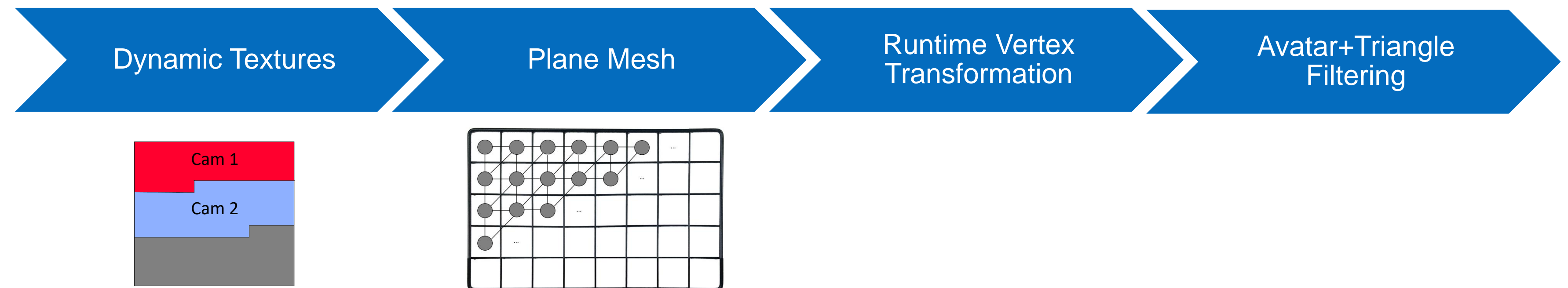
- Registration using [Mühlenbrock21]’s method
- Dyncam’s rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions



- Splatting:

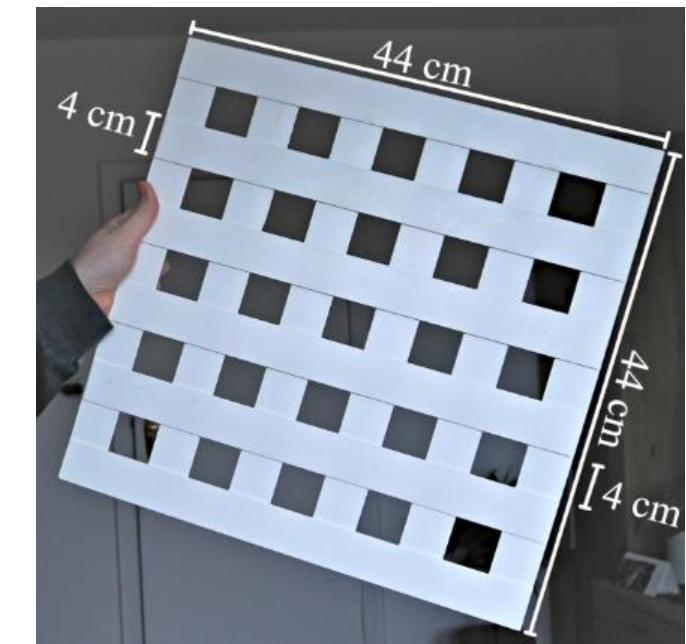


- Mesh:



Point Cloud Registration and Rendering

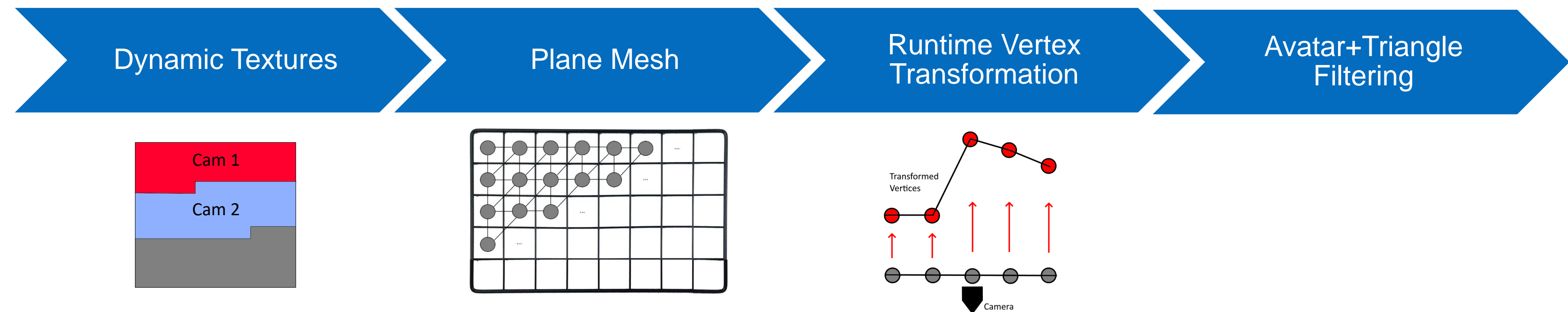
- Registration using [Mühlenbrock21]'s method
- Dyncam's rendering and UE4 LiDAR PC plugin not convincing
➔ 2 own solutions



- Splatting:

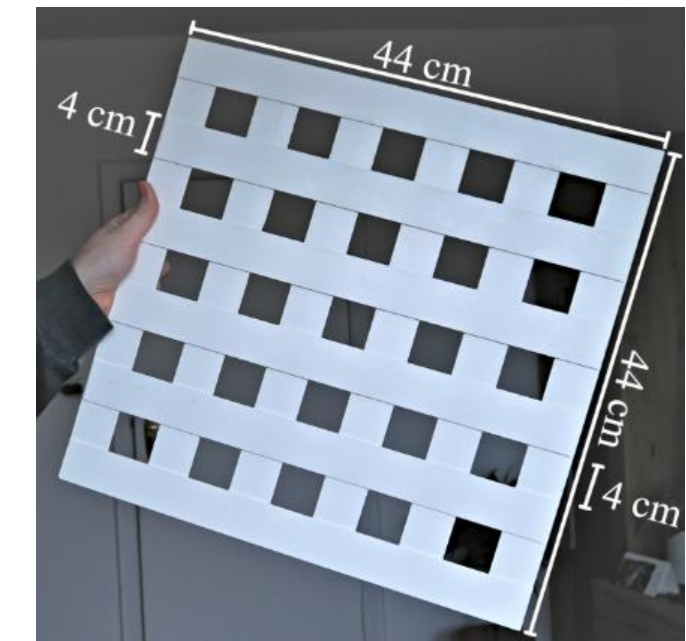


- Mesh:

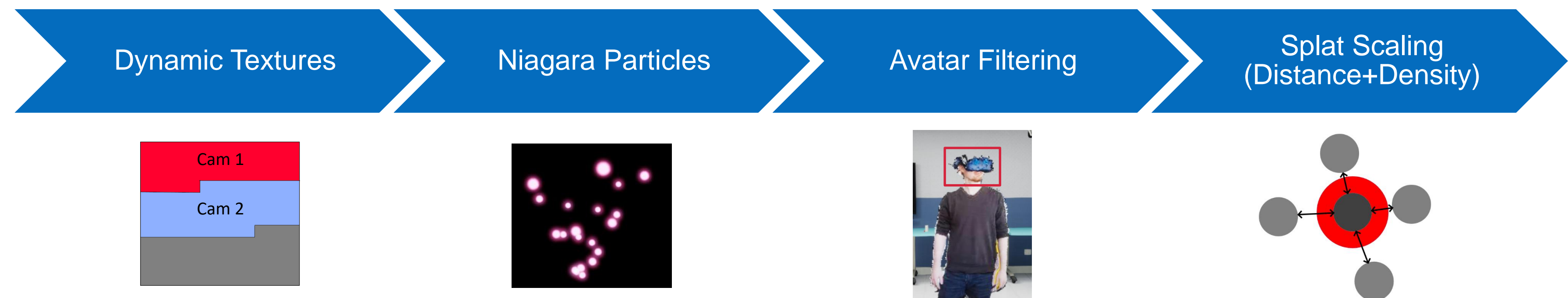


Point Cloud Registration and Rendering

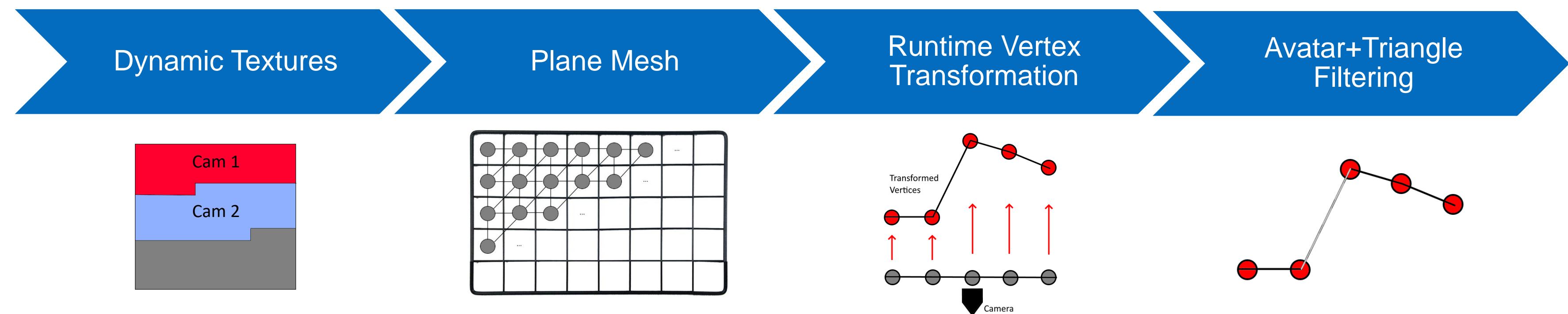
- Registration using [Mühlenbrock21]'s method
- Dyncam's rendering and UE4 LiDAR PC plugin not convincing
 ➔ 2 own solutions



- Splatting:



- Mesh:



Point Cloud Registration and Rendering



Point Cloud



Mesh



Physical Scene

Point Cloud Registration and Rendering



Point Cloud

Holes



Mesh



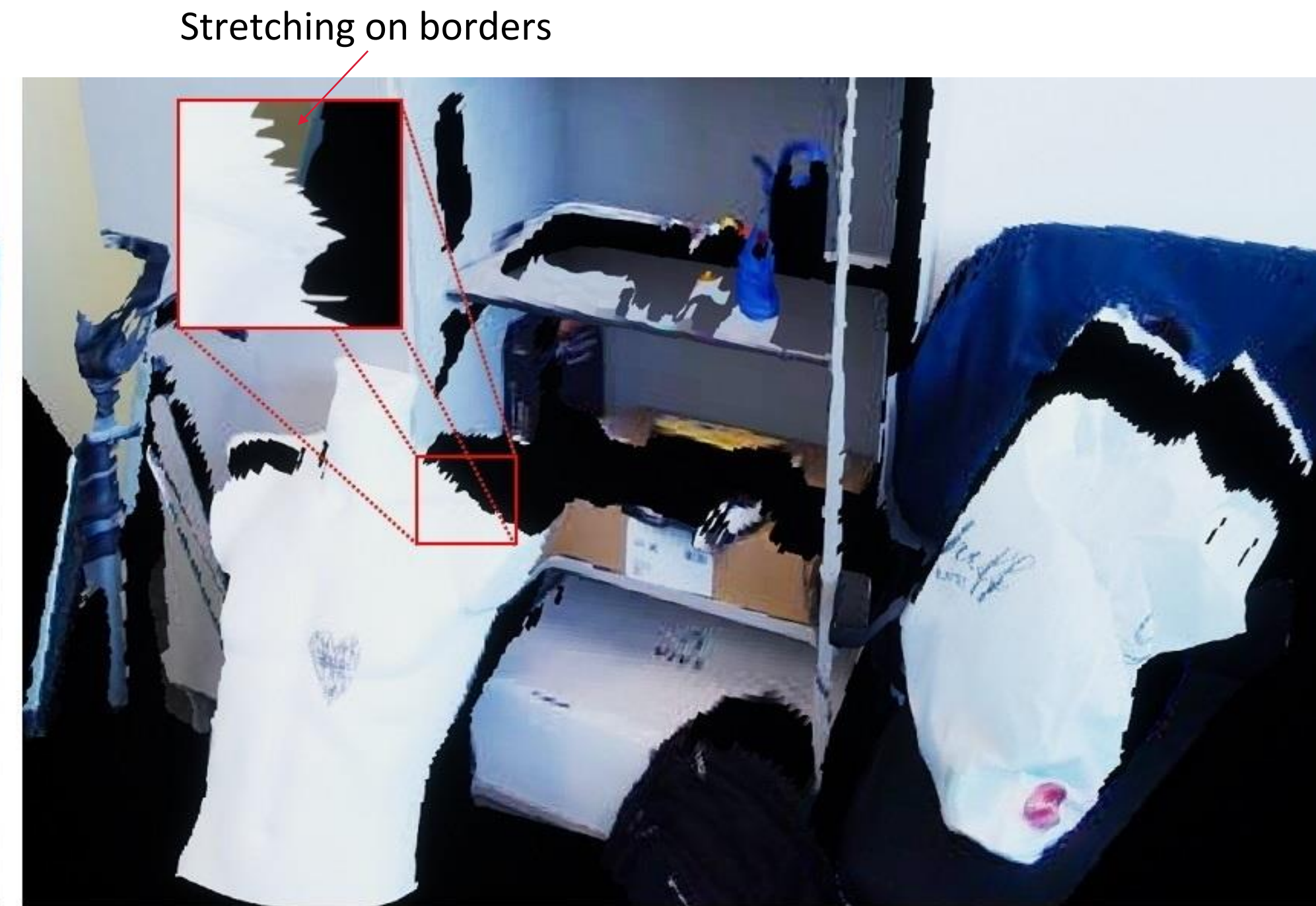
Physical Scene

Point Cloud Registration and Rendering



Point Cloud

Holes

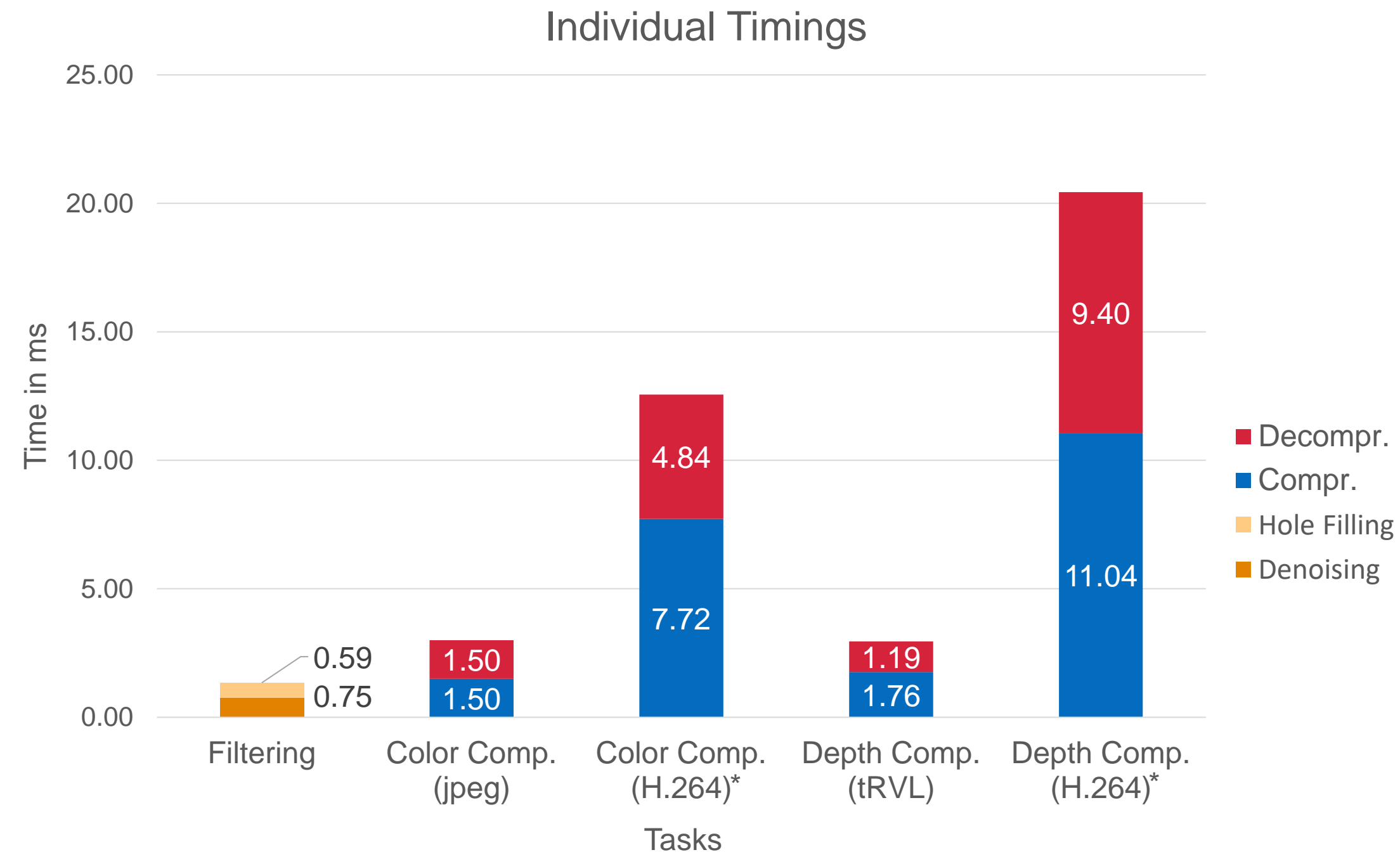


Mesh



Physical Scene

Results: Timings & Compression

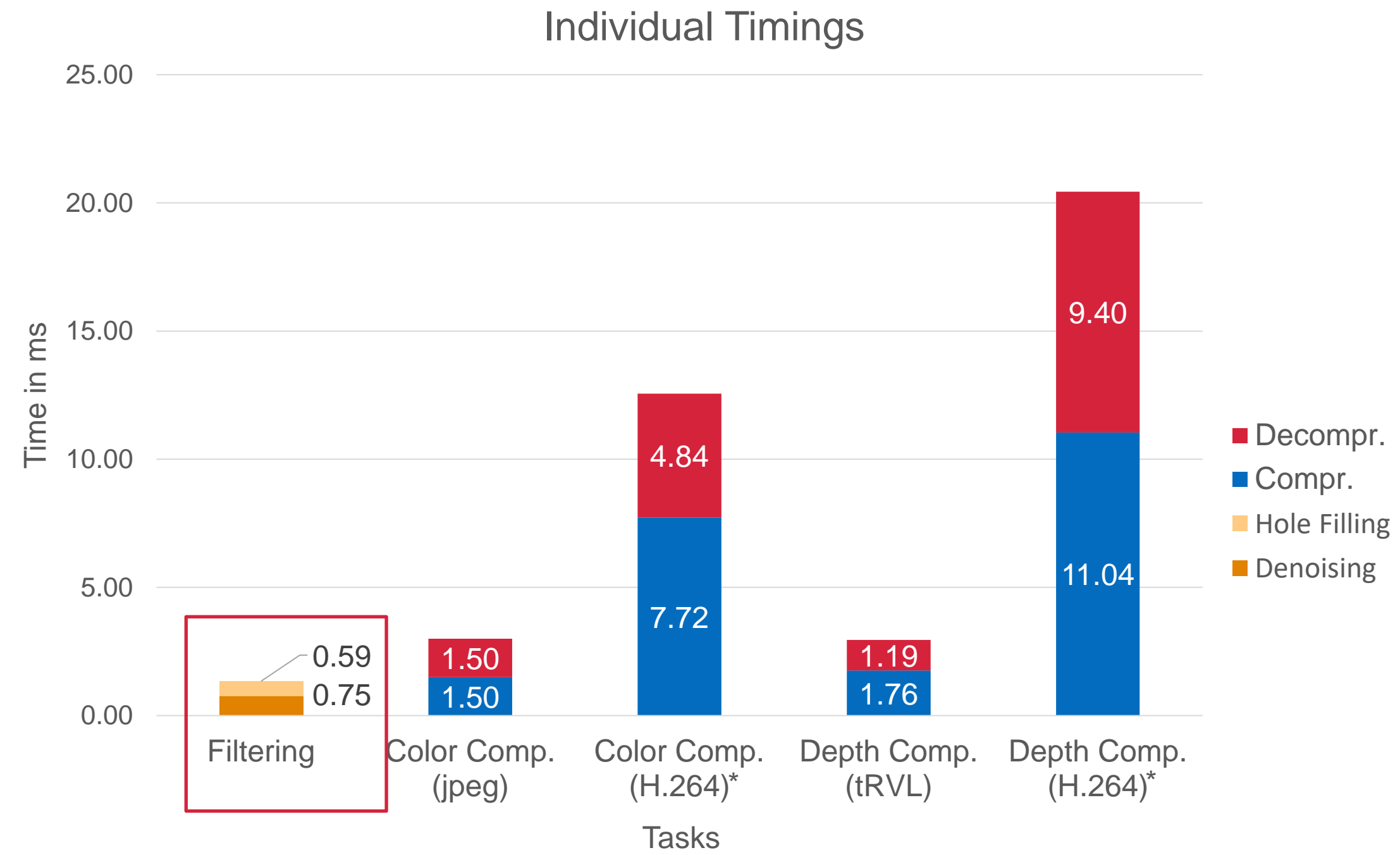


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

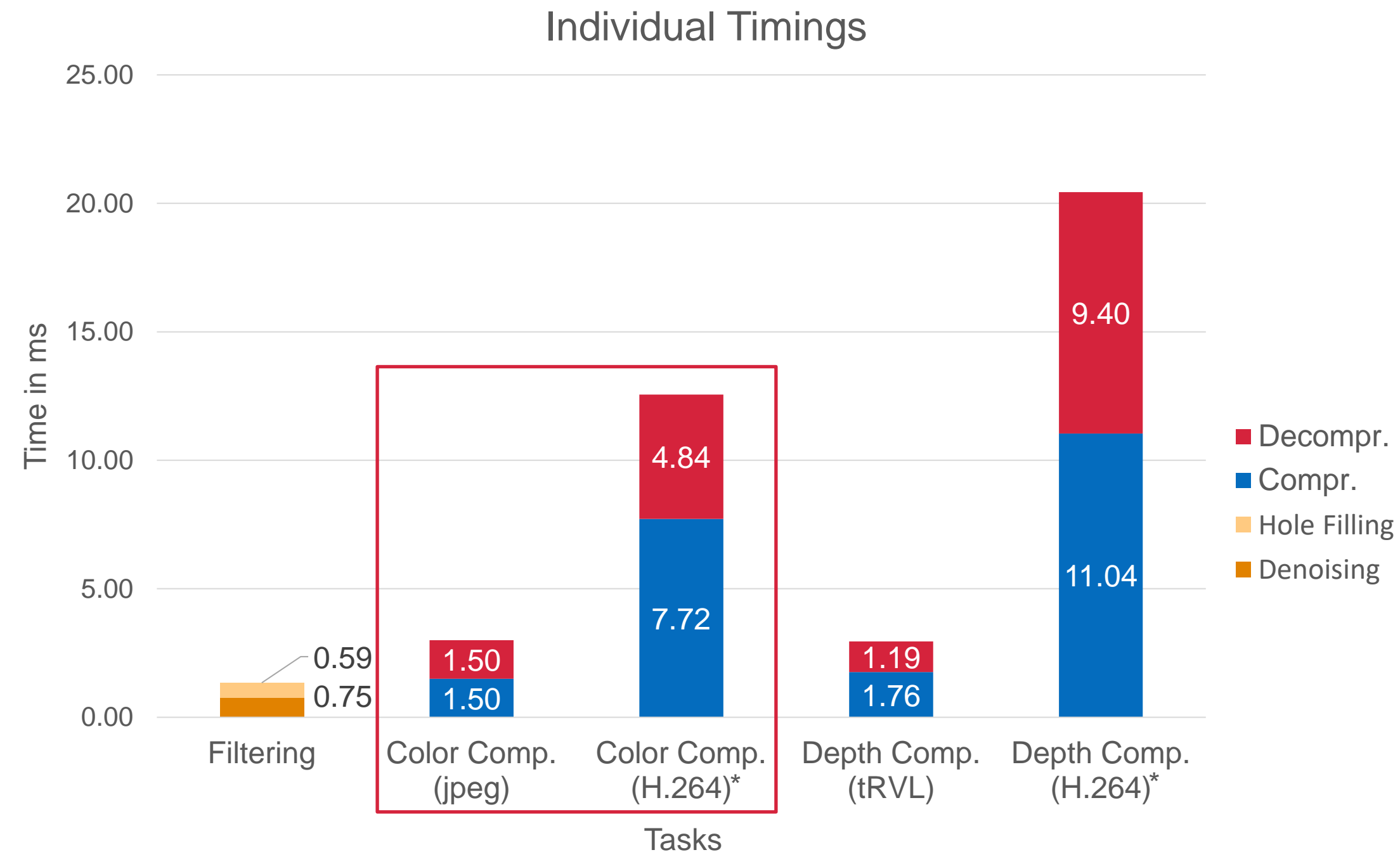


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

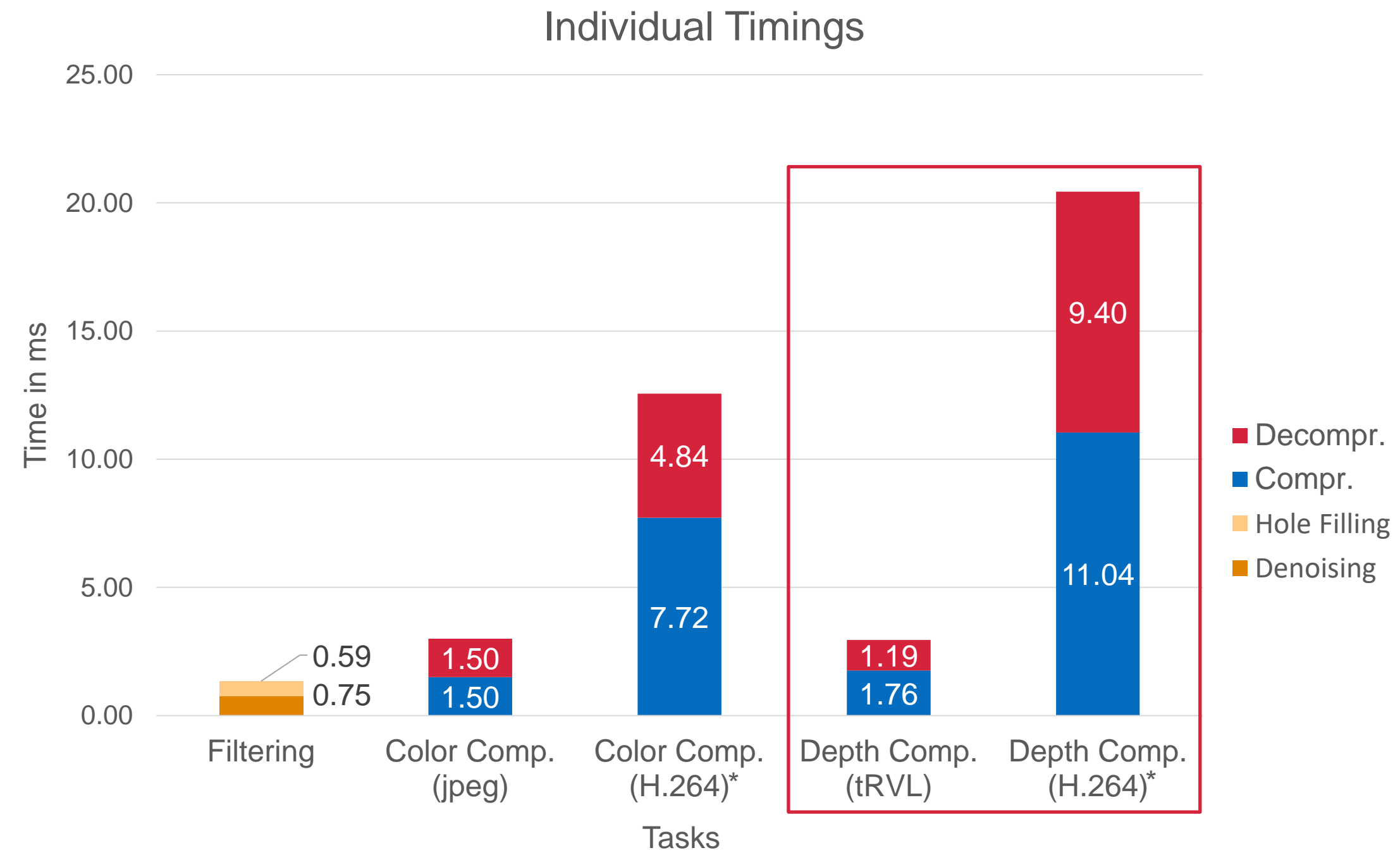


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

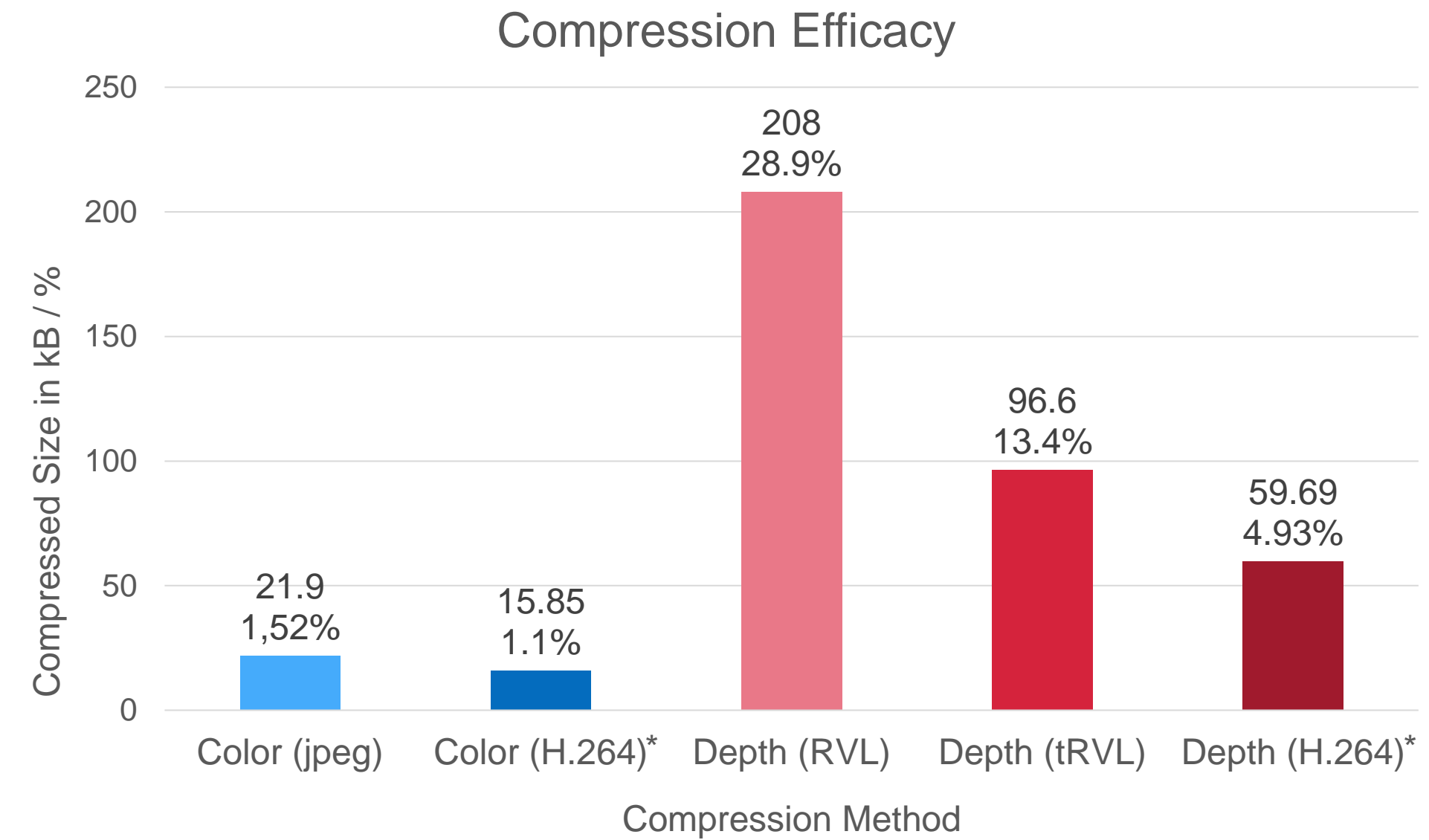
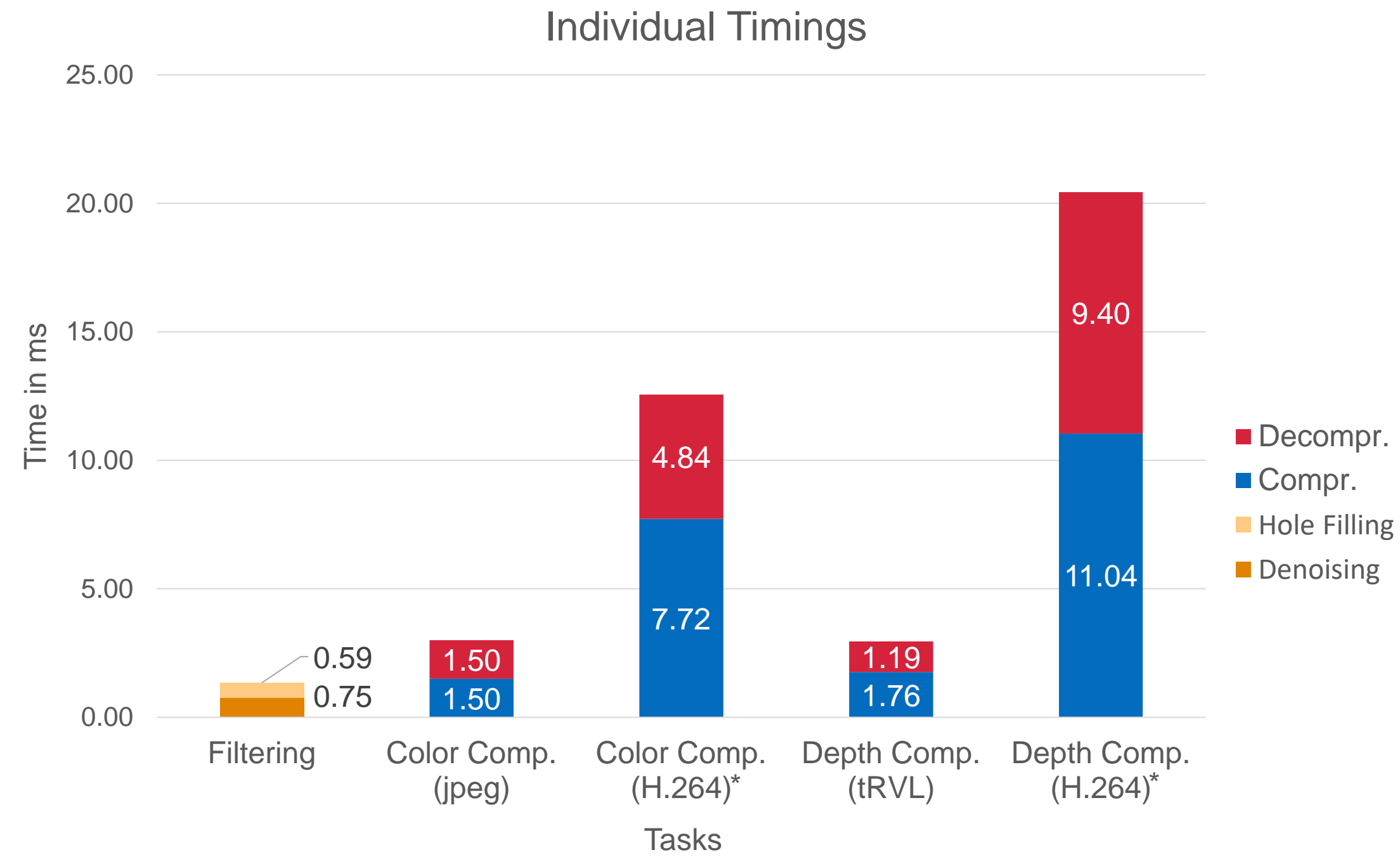


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

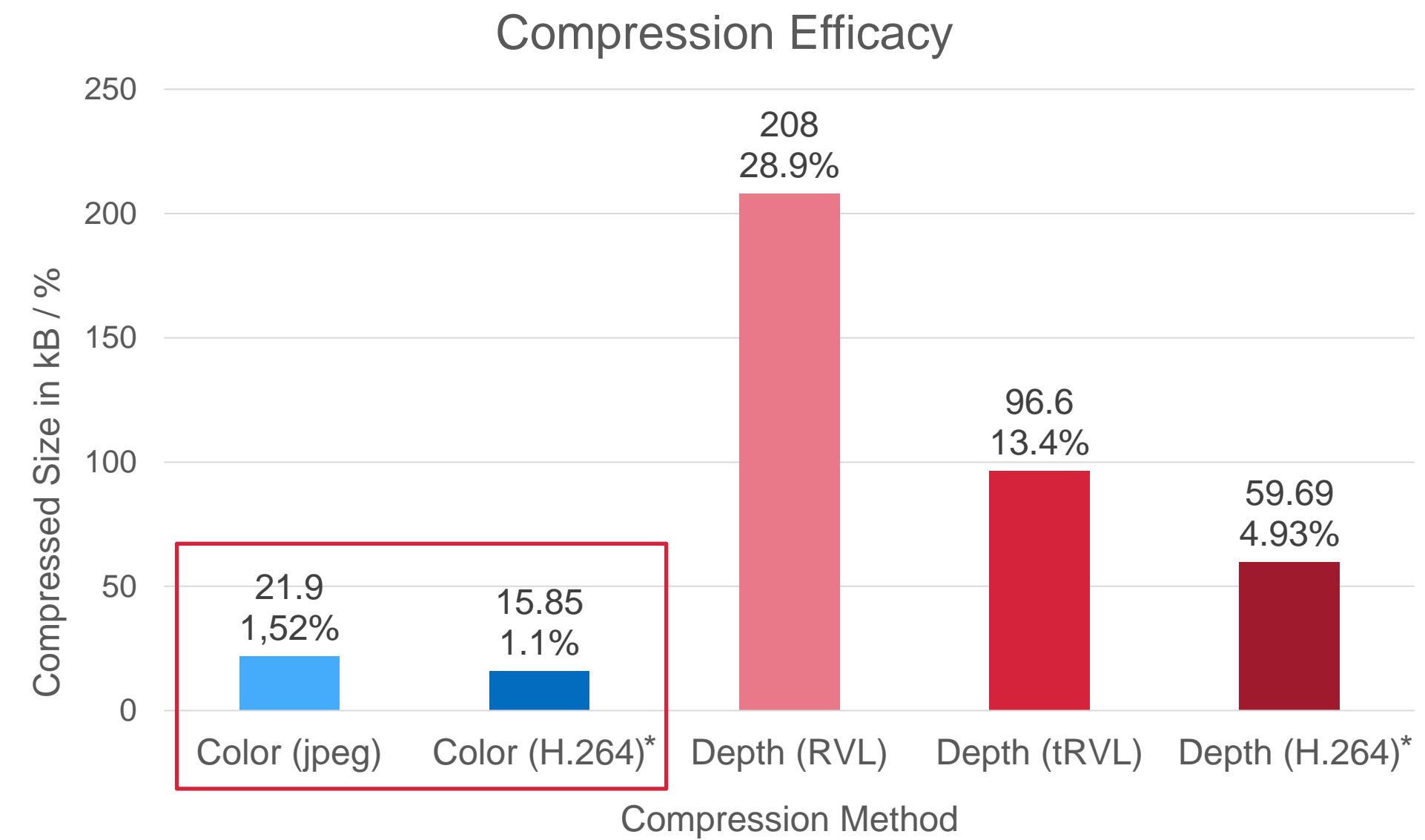
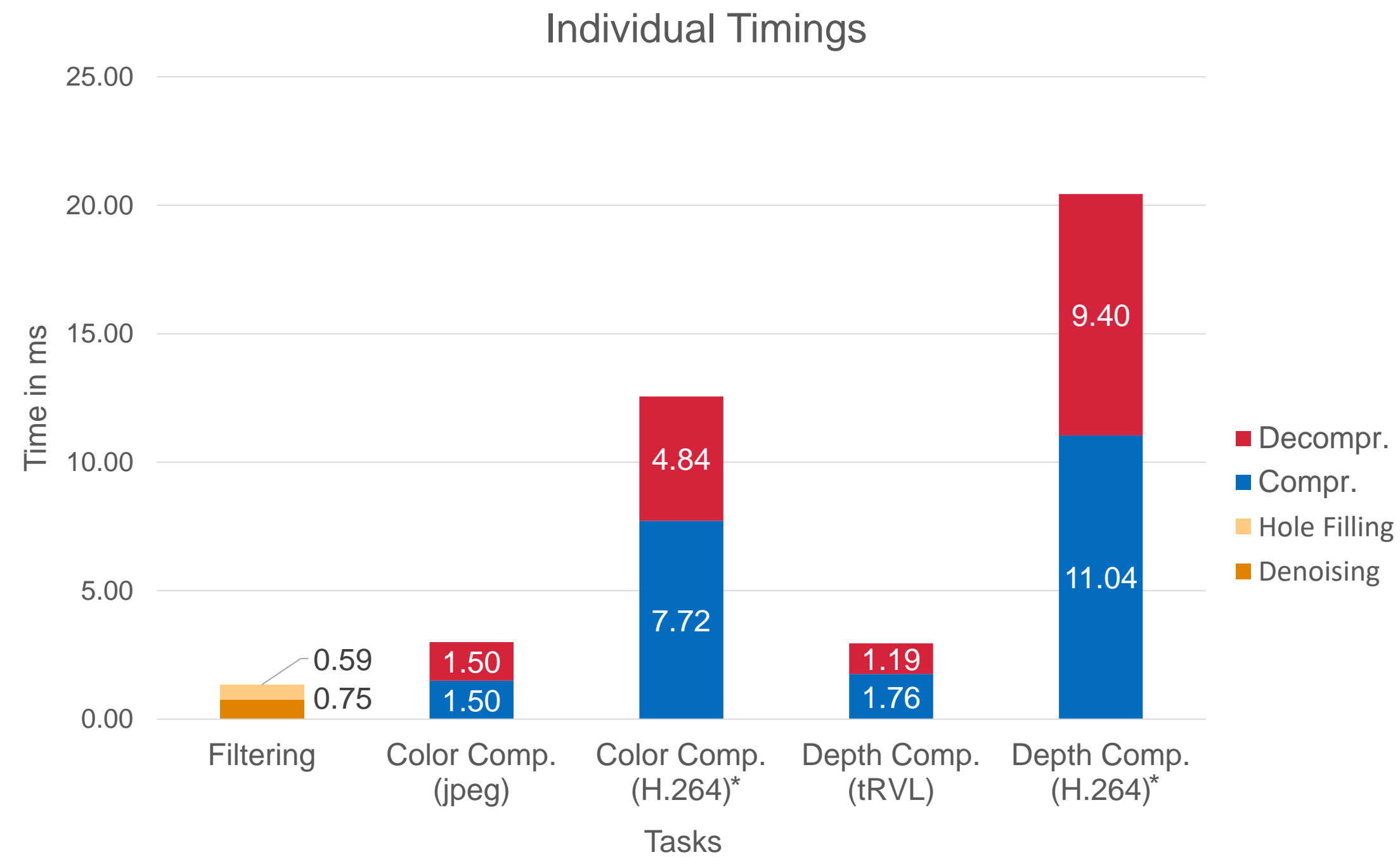


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

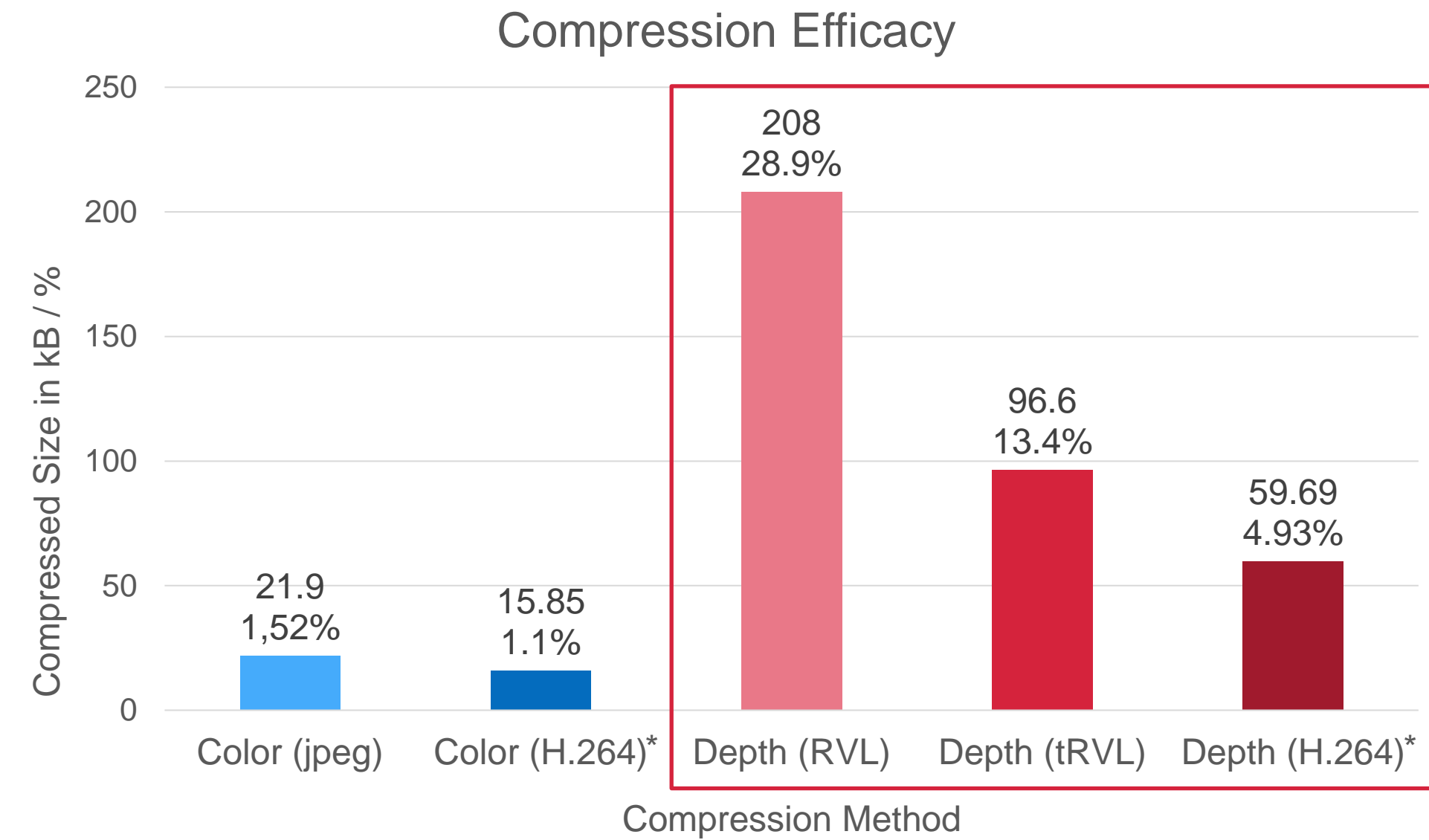
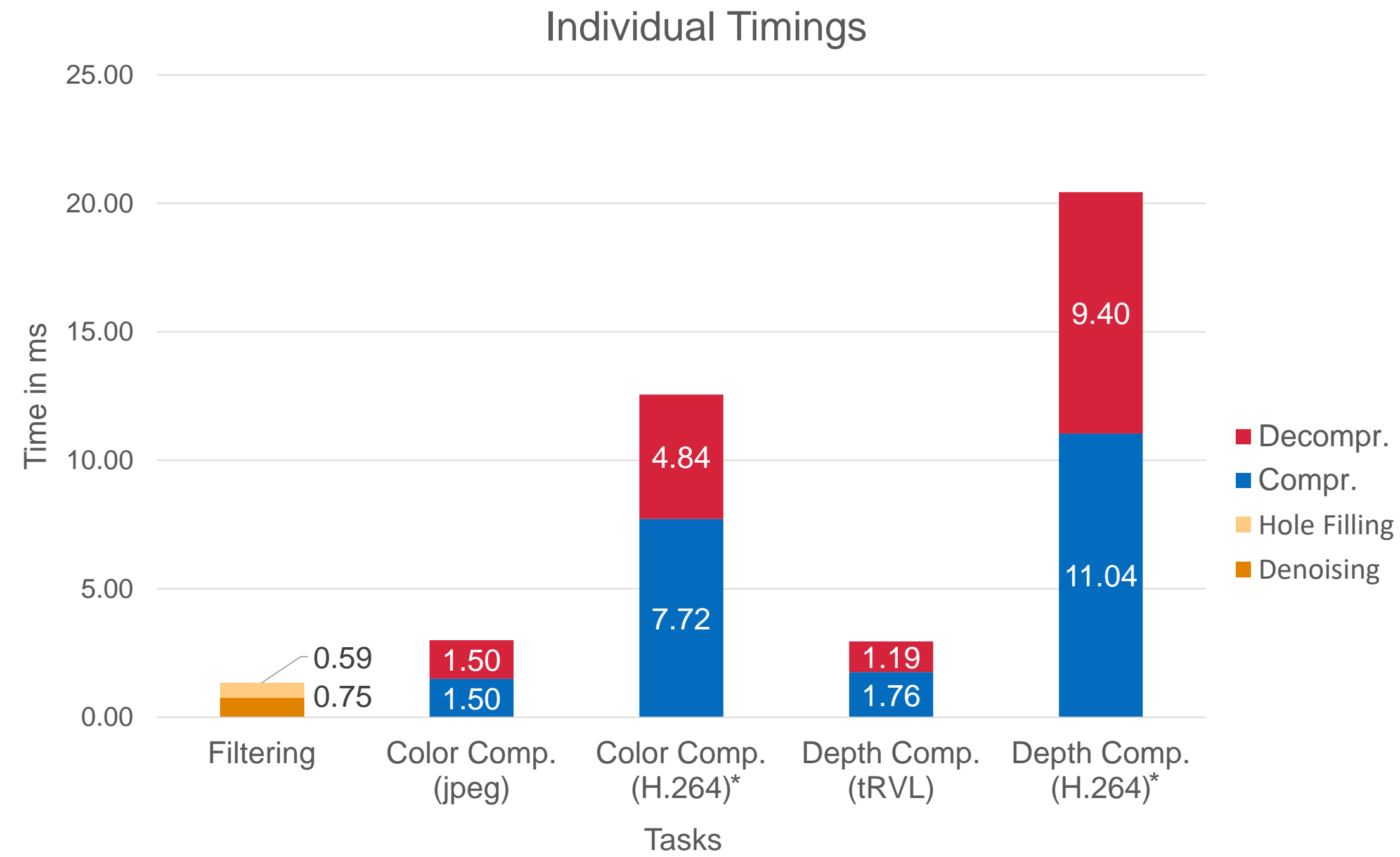


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

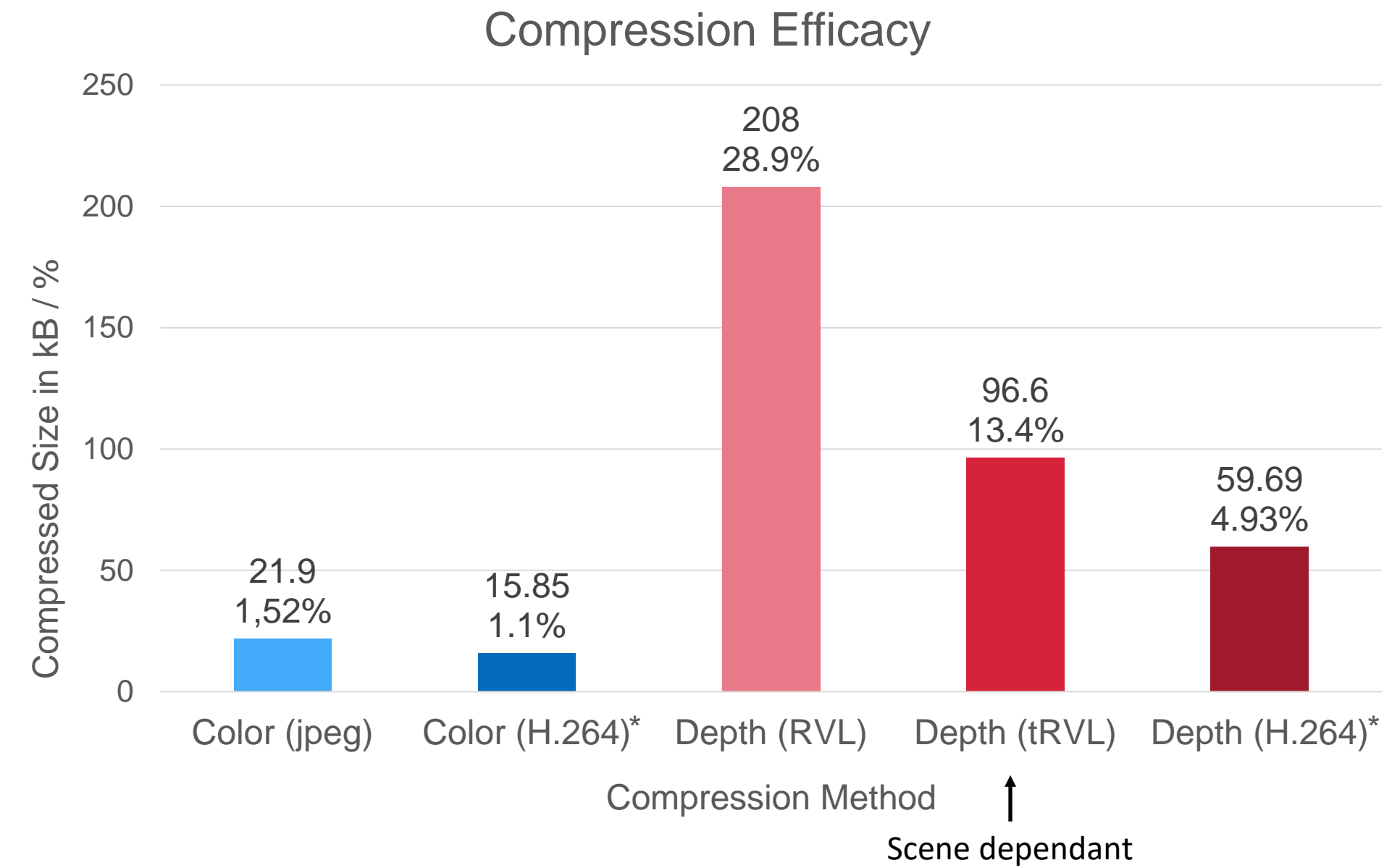
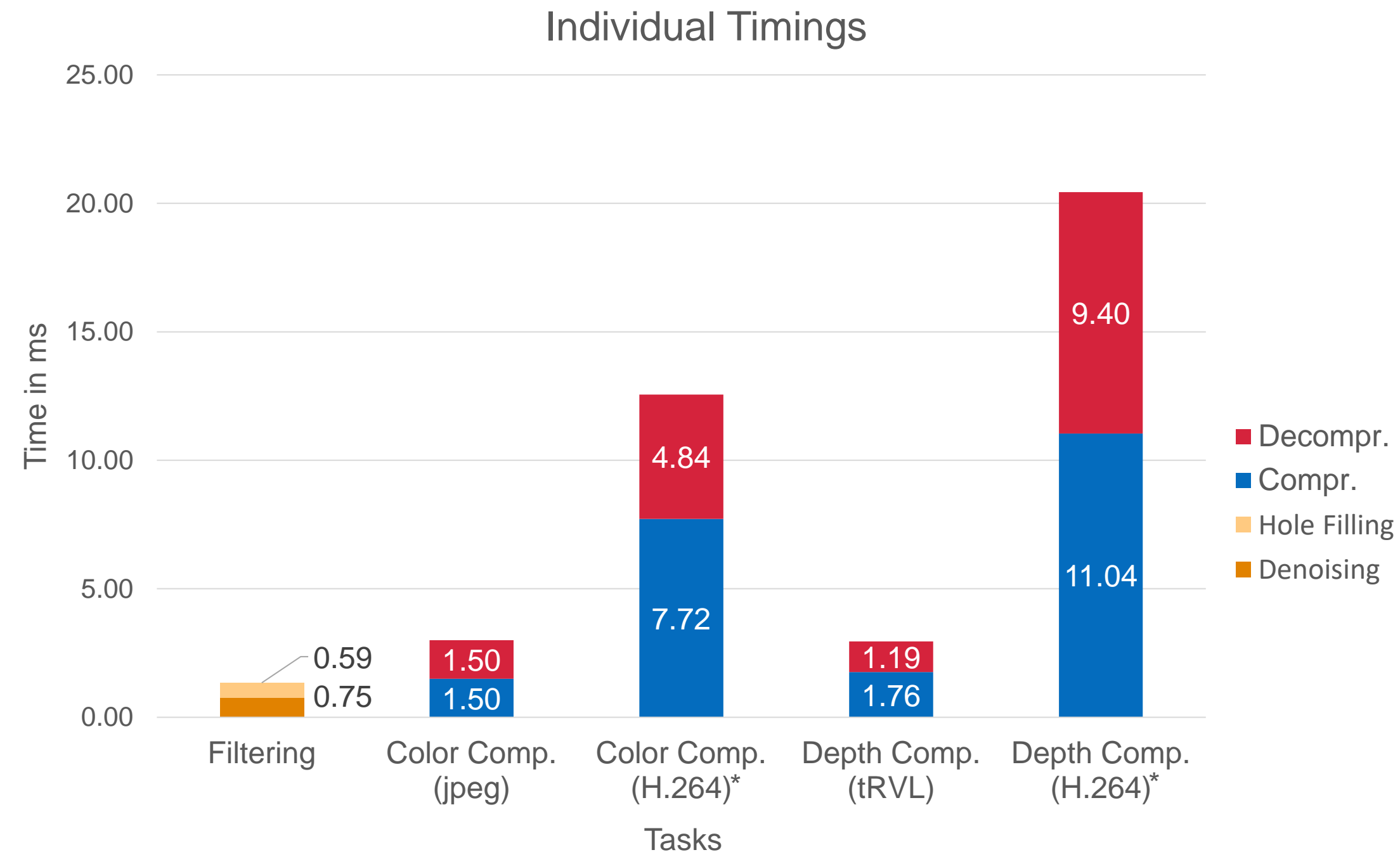


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

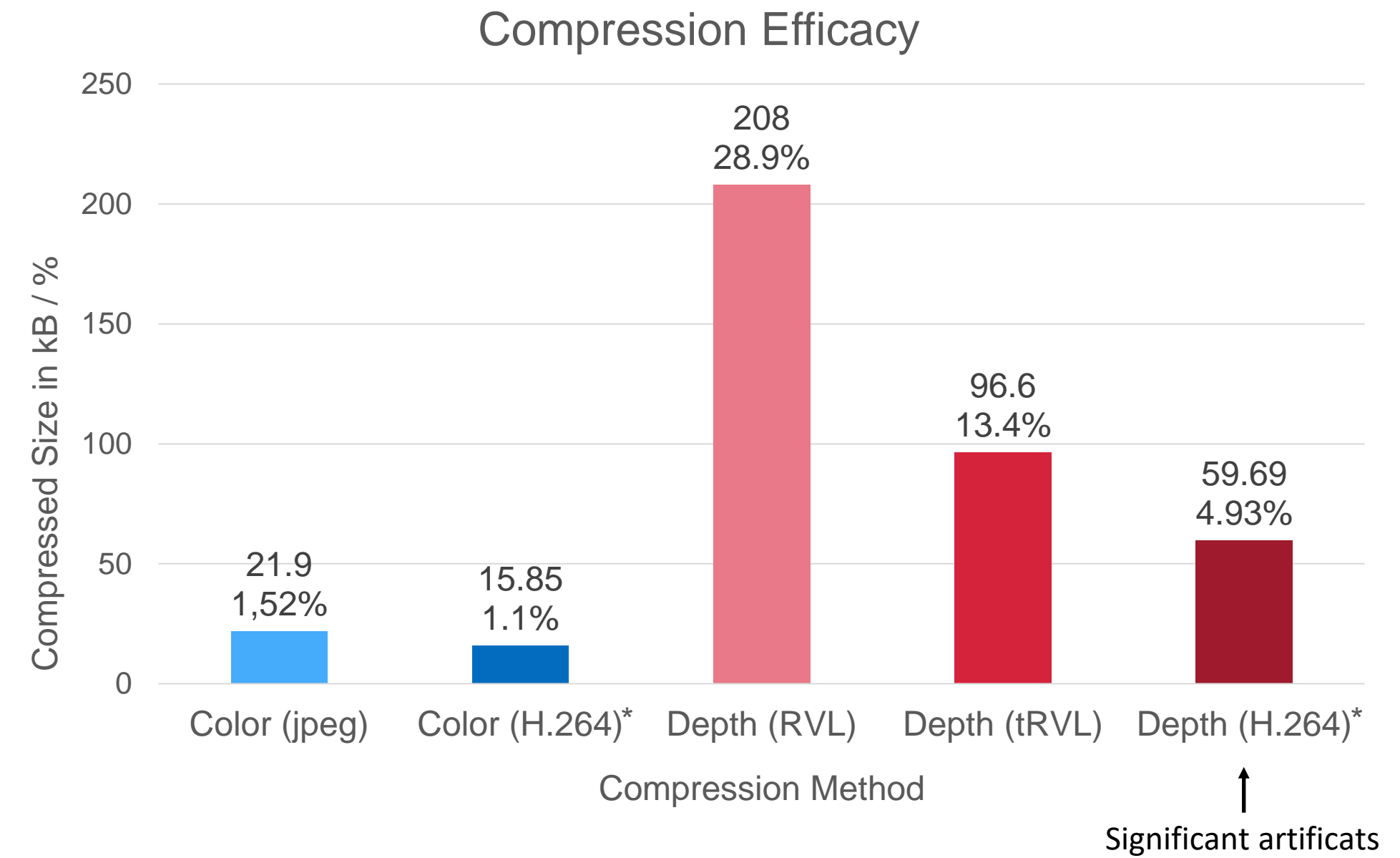
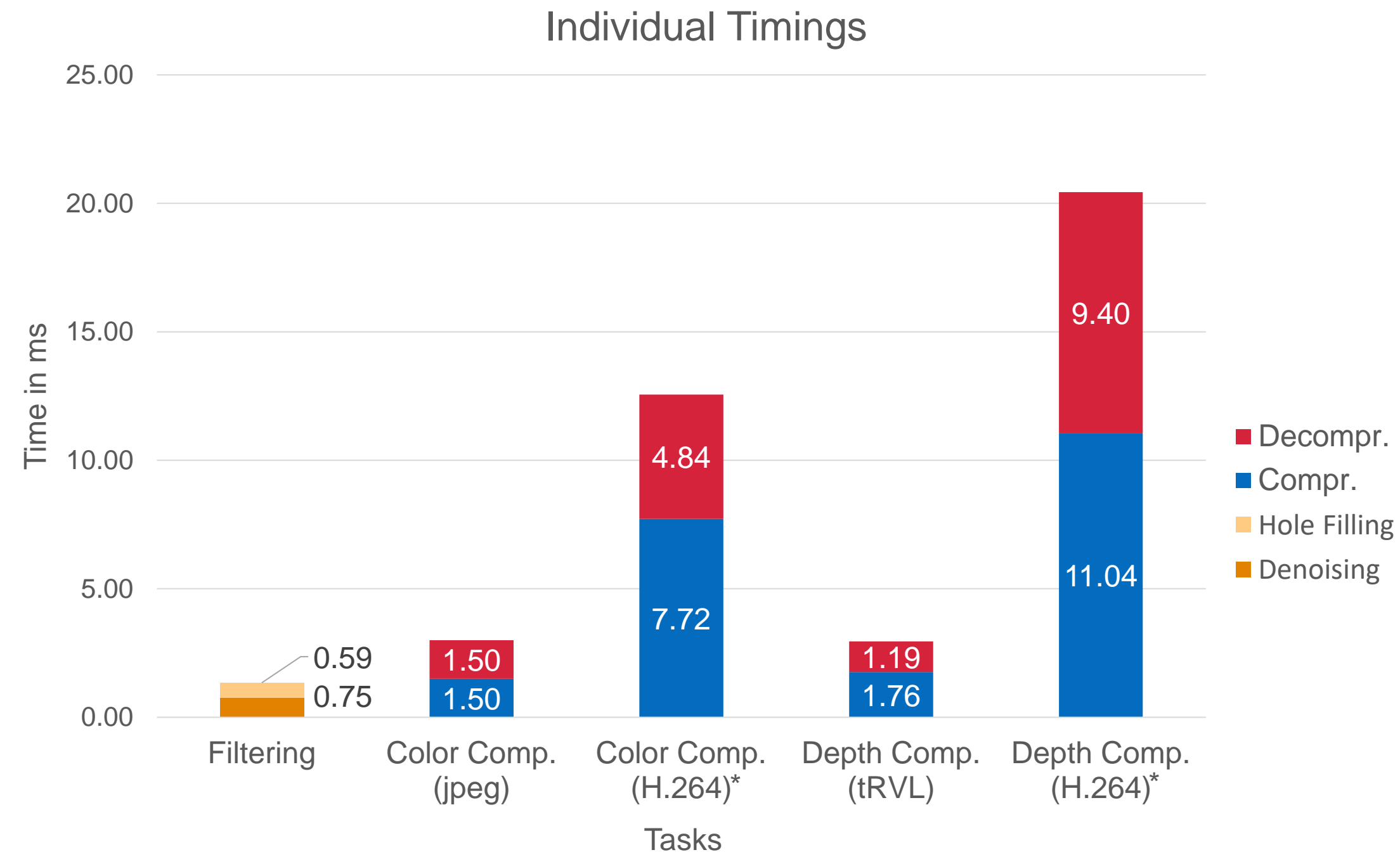


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression

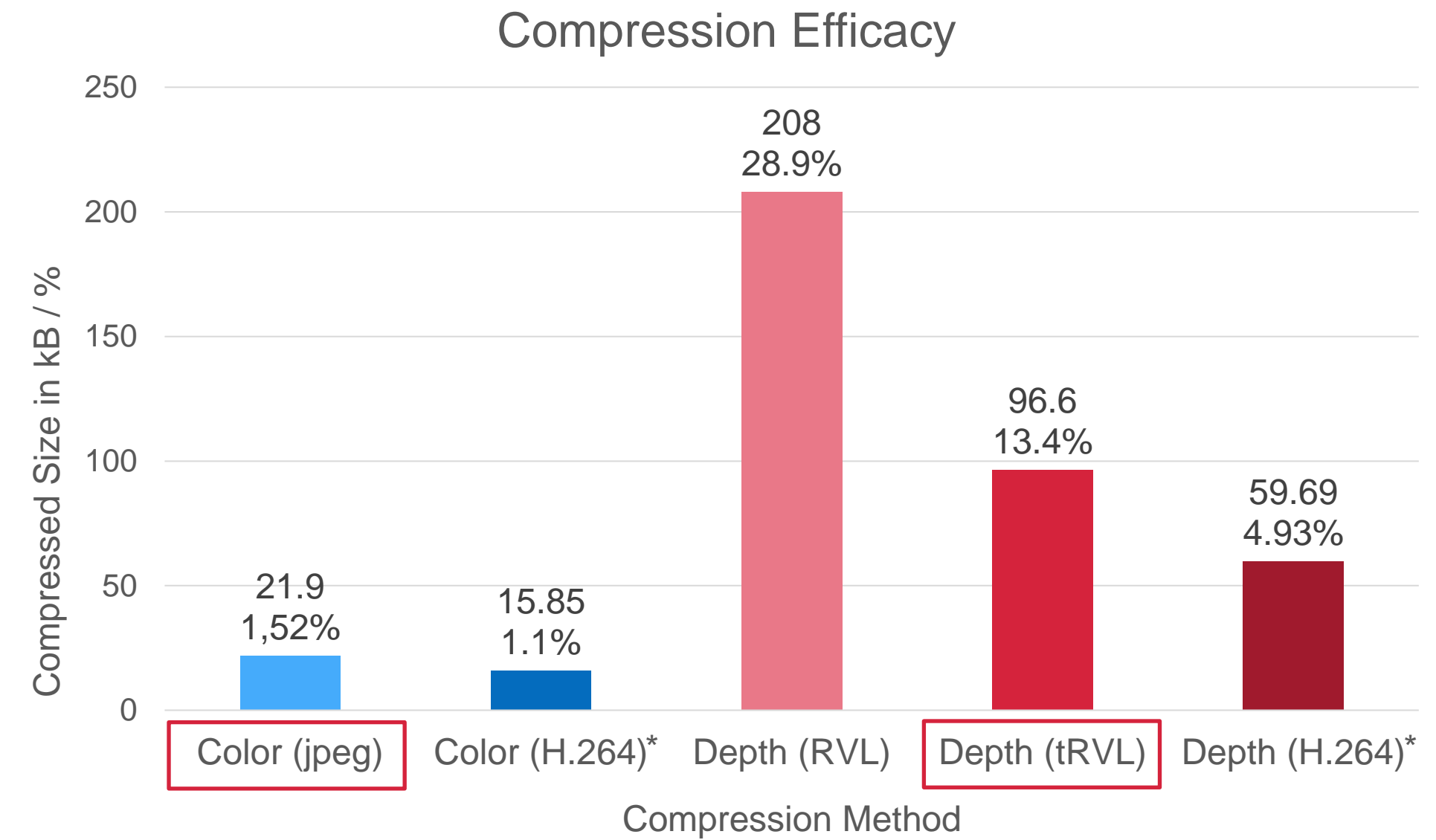
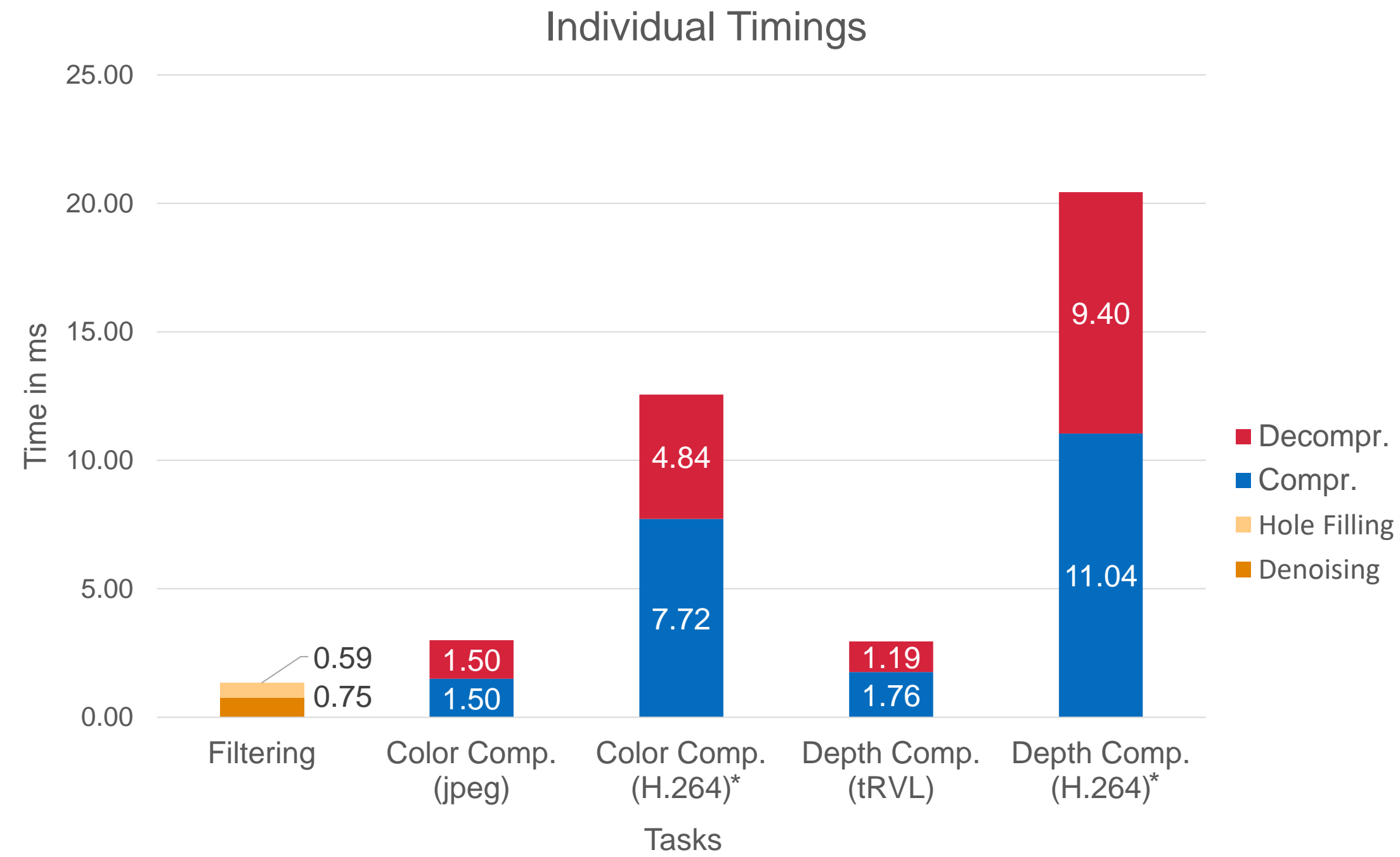


*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

Results: Timings & Compression



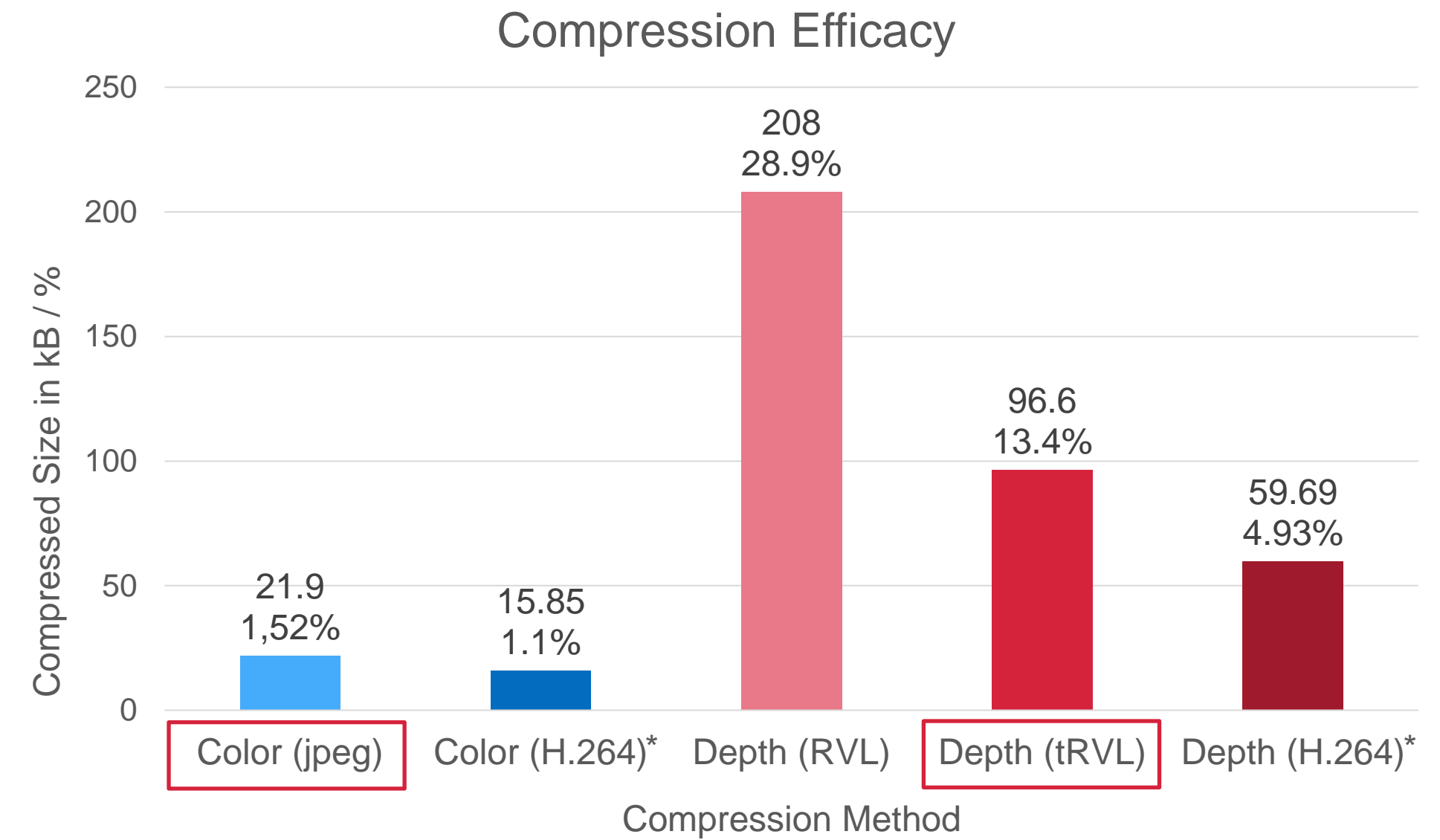
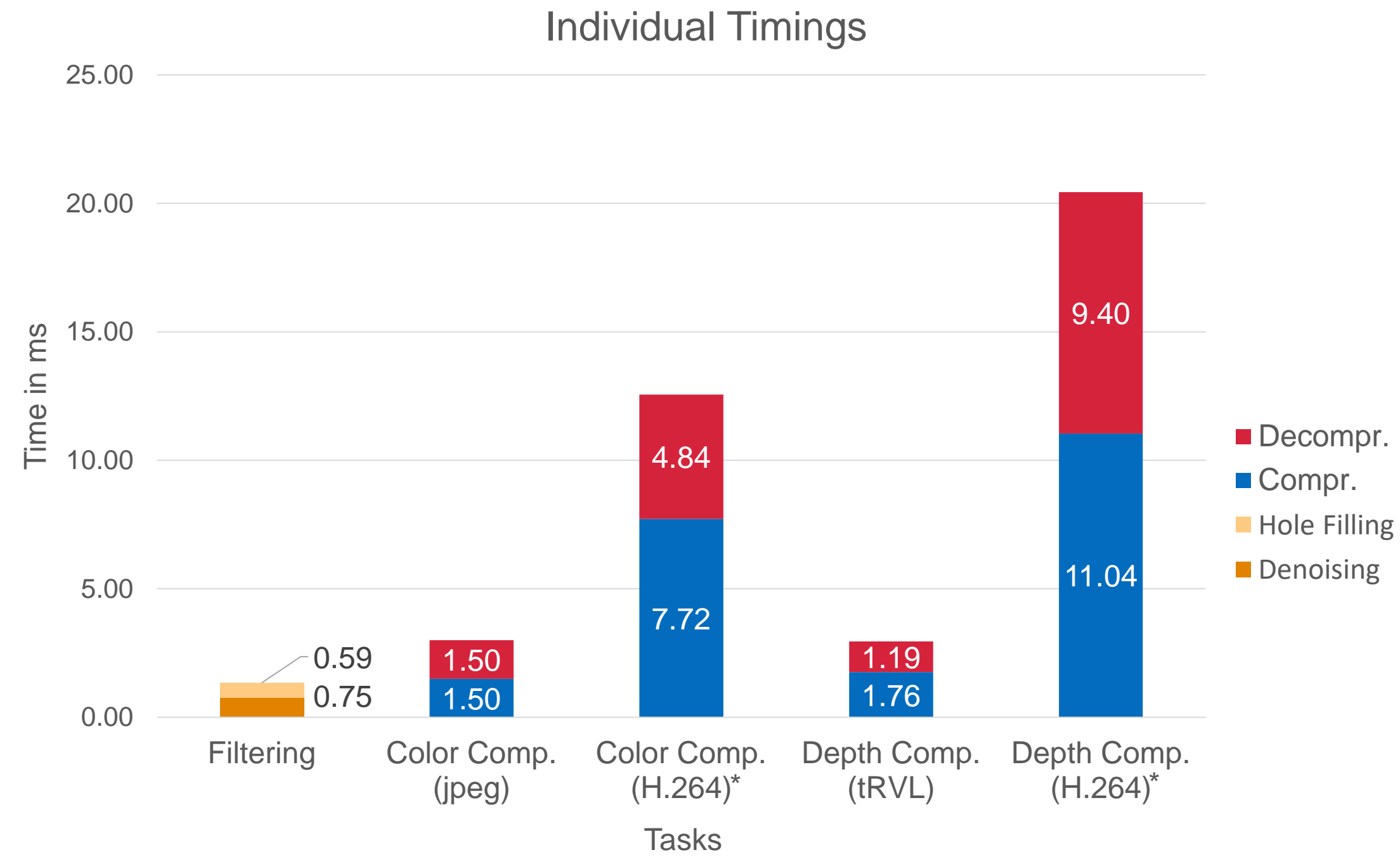
Bandwidth per cam (30fps):
3.55kB/s or 27.8 Mbps

System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

Results: Timings & Compression

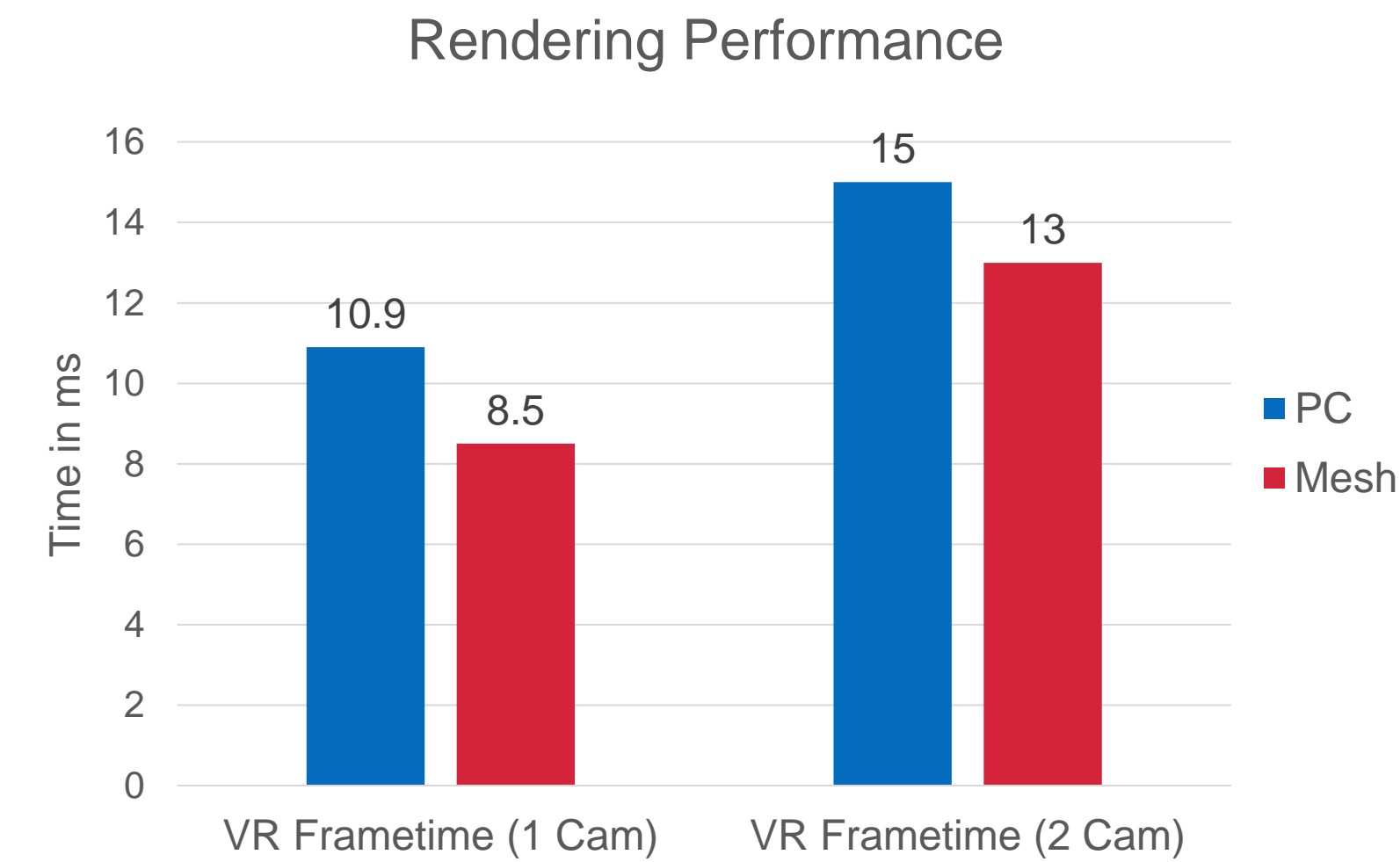


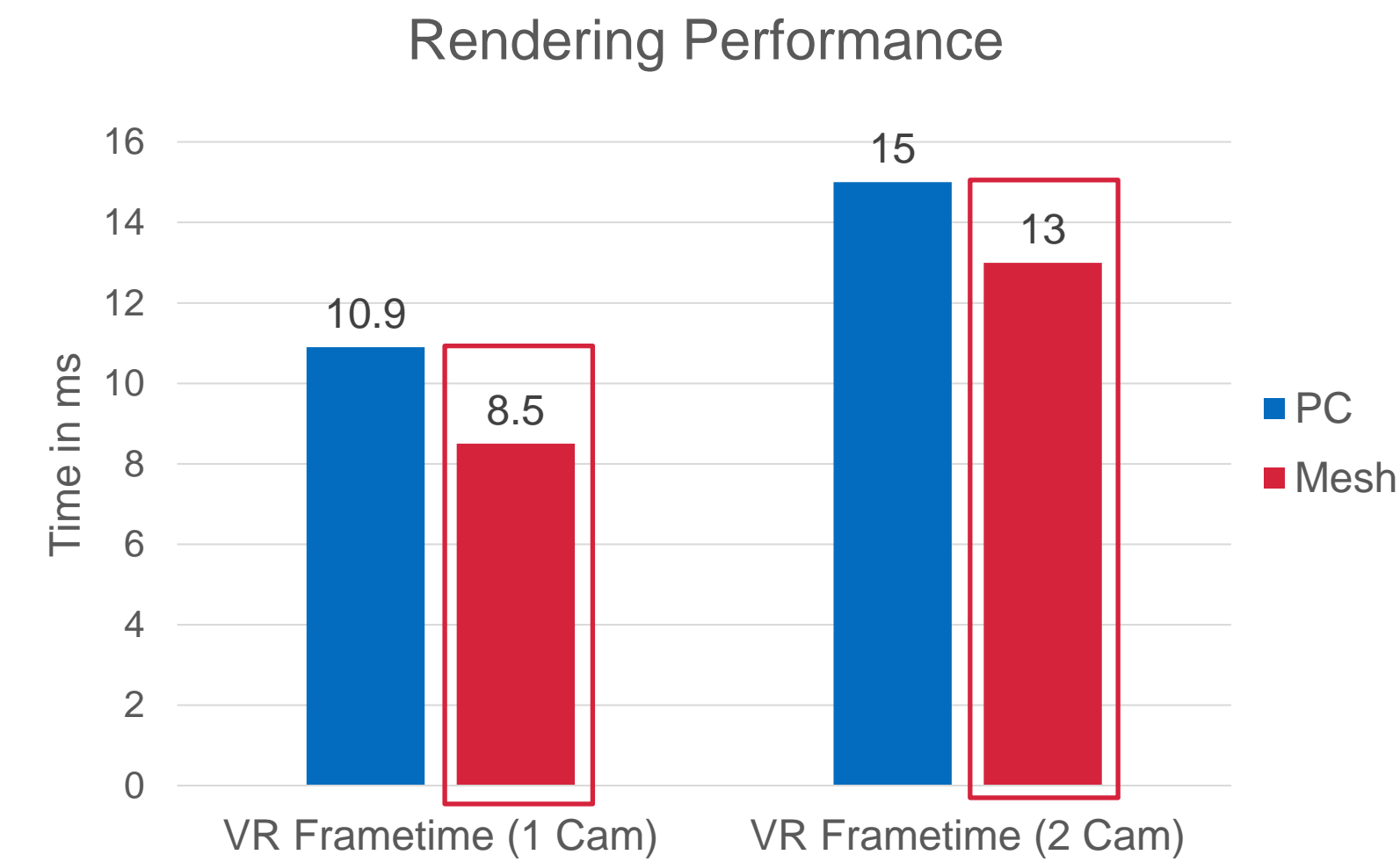
Bandwidth per cam (30fps):
3.55kB/s or 27.8 Mbps
18.8% lower with cropping

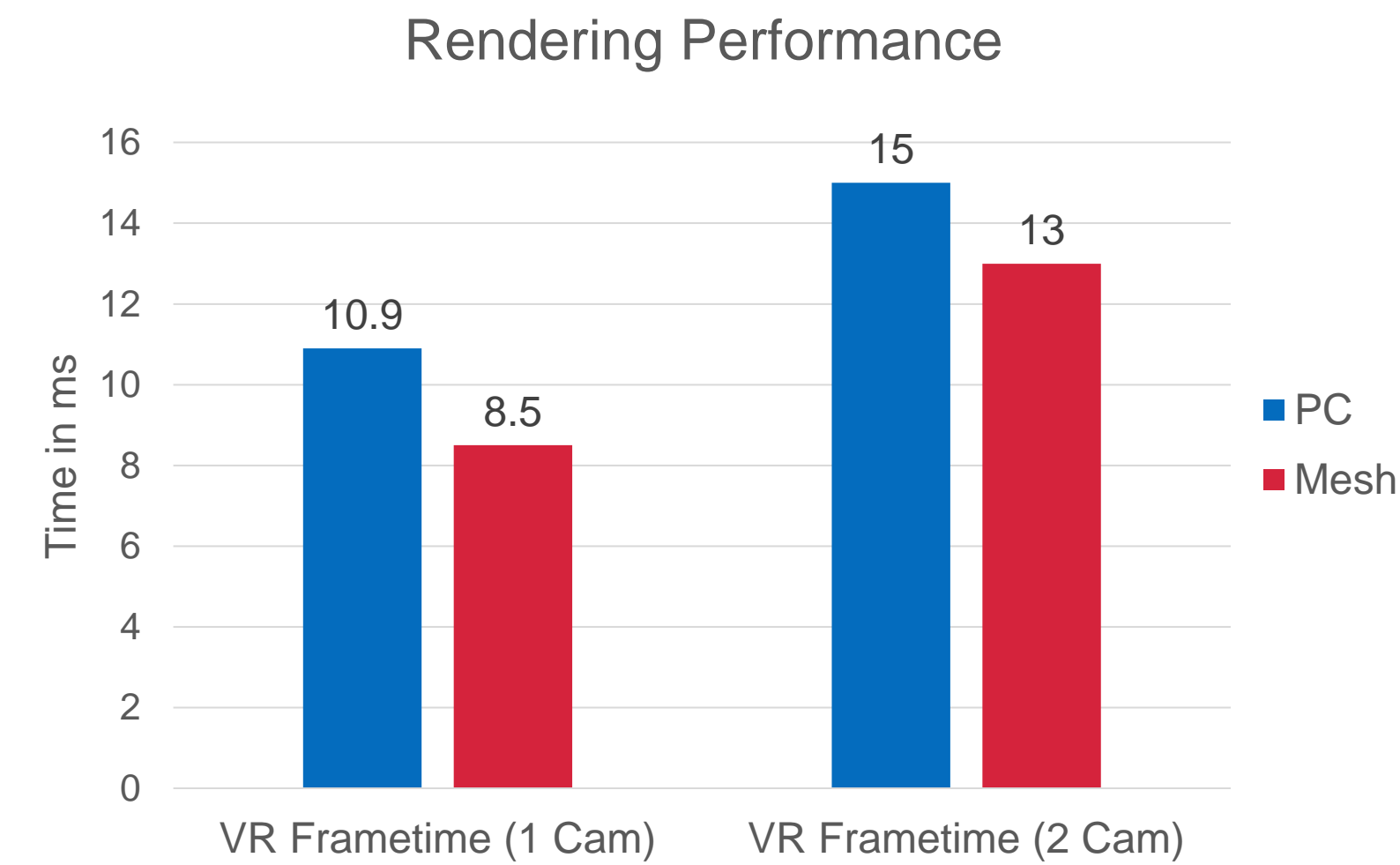
System Specification

Item	Details
HMD	HTC Vive Pro Eye
CPU	Intel Core i7 7800x
GPU	Nvidia GeForce 2070
RAM	16 GB
OS	Windows 10
Engine	Unreal Engine 4.26
Camera	Azure Kinect C: FullHD; D: 640x576

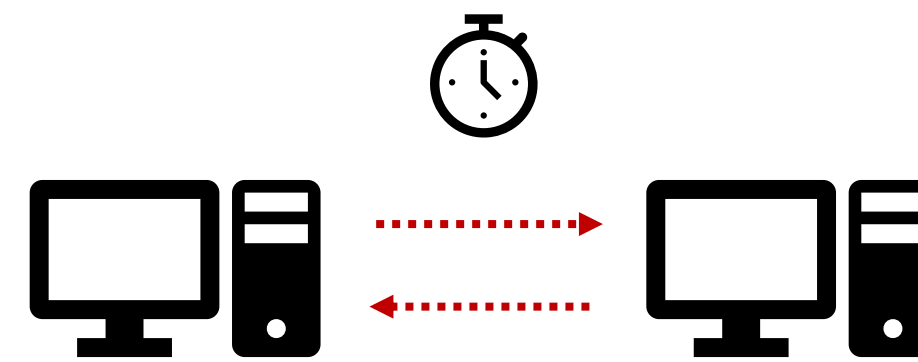
*H.264 (color) @ CRF20, (depth) CRF10, Ultrafast, Zerolatency

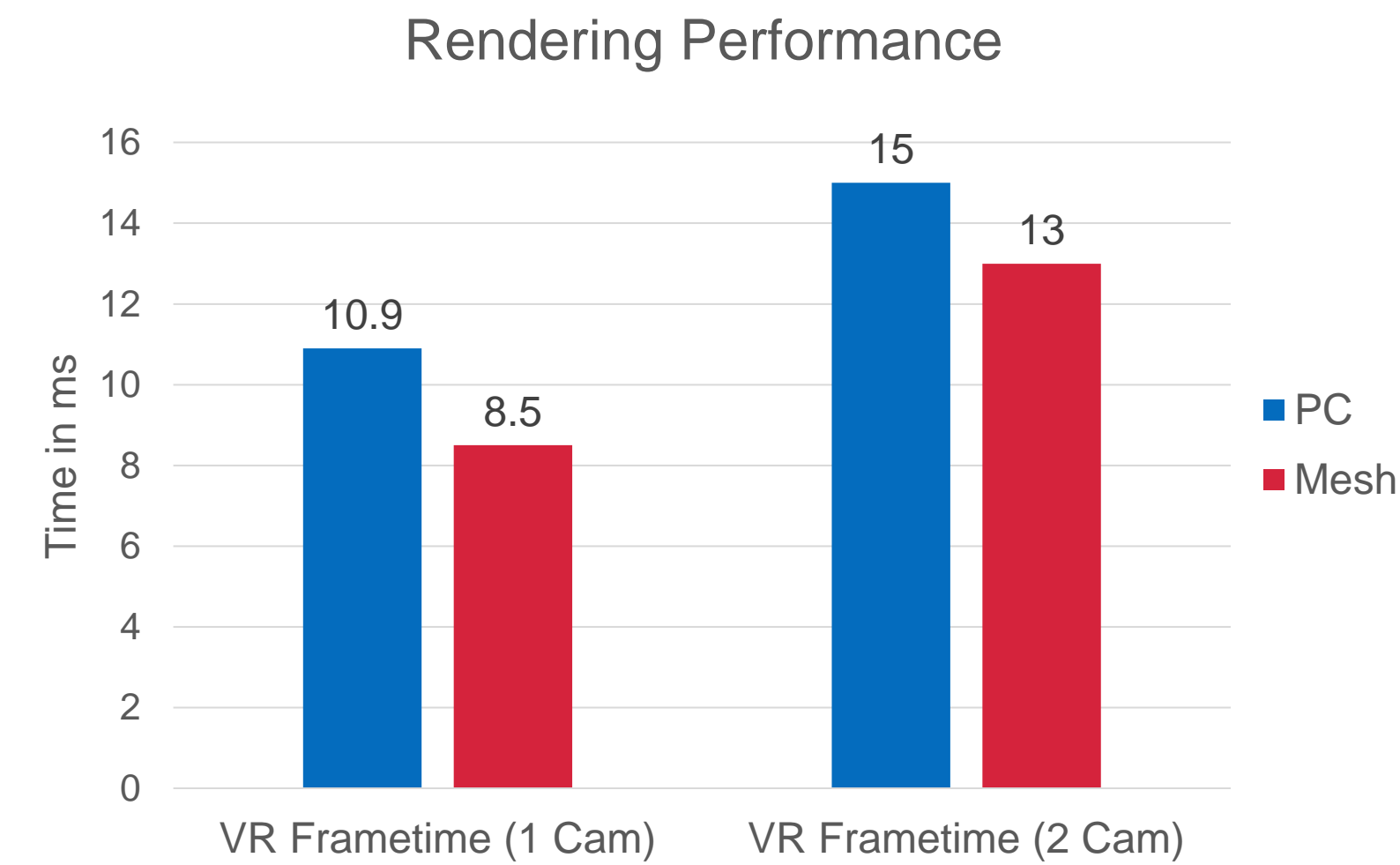




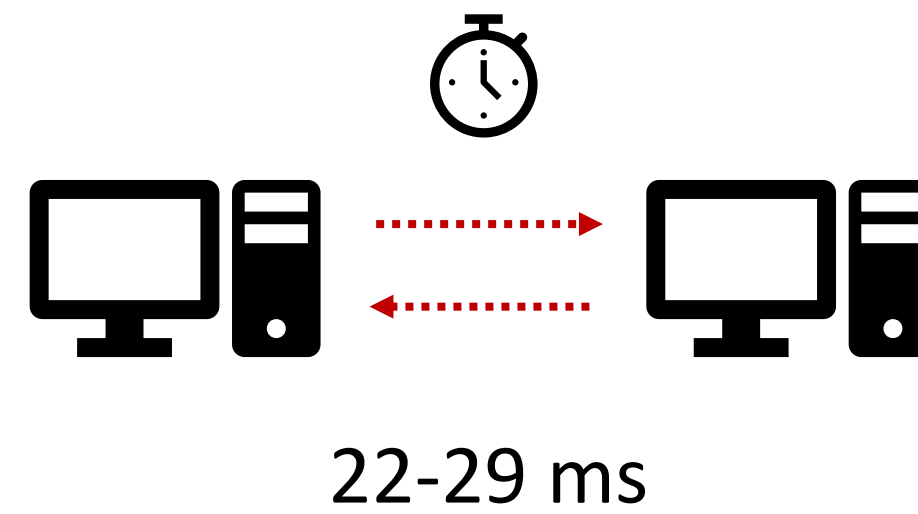


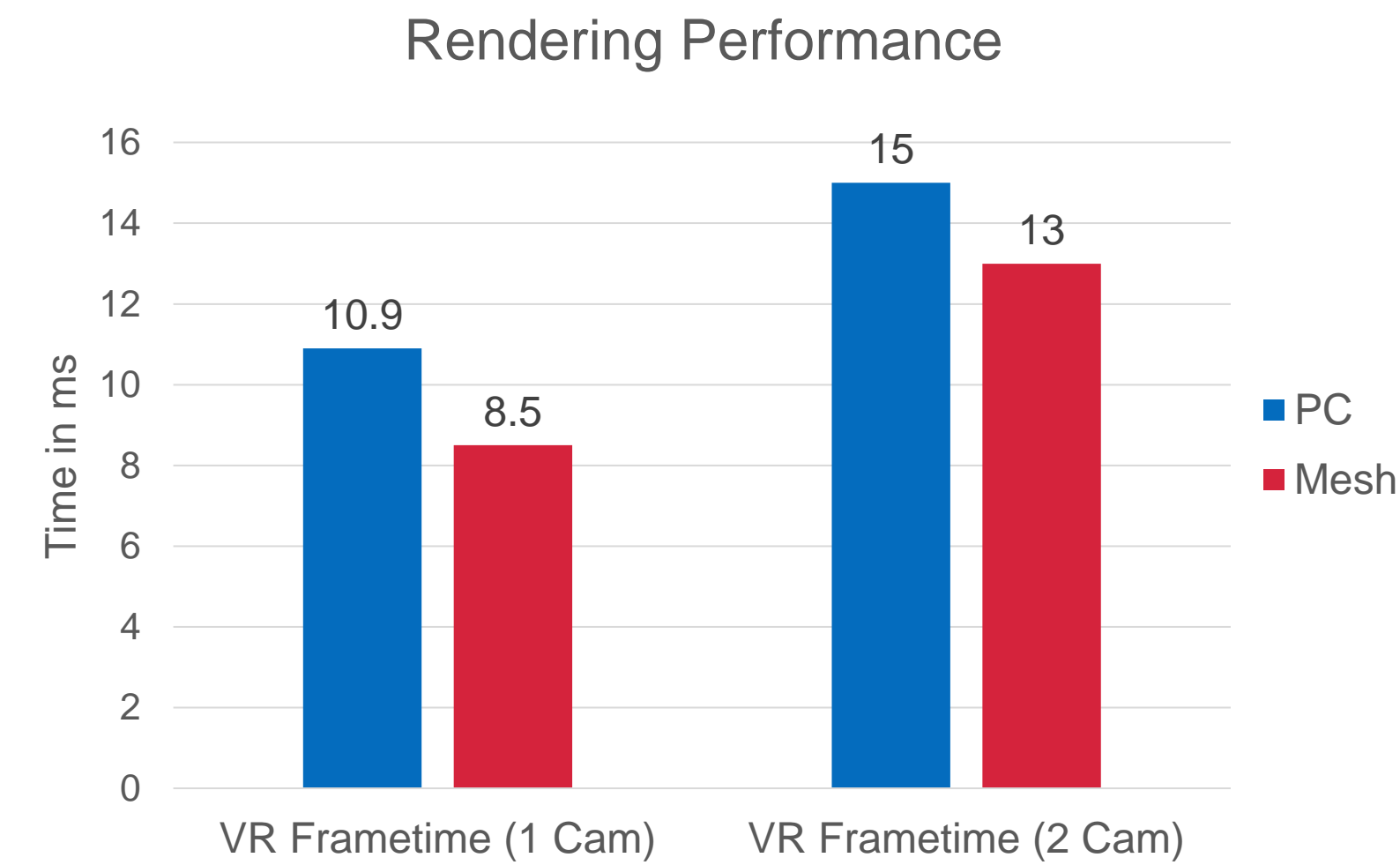
VR Round-Trip Time Latency



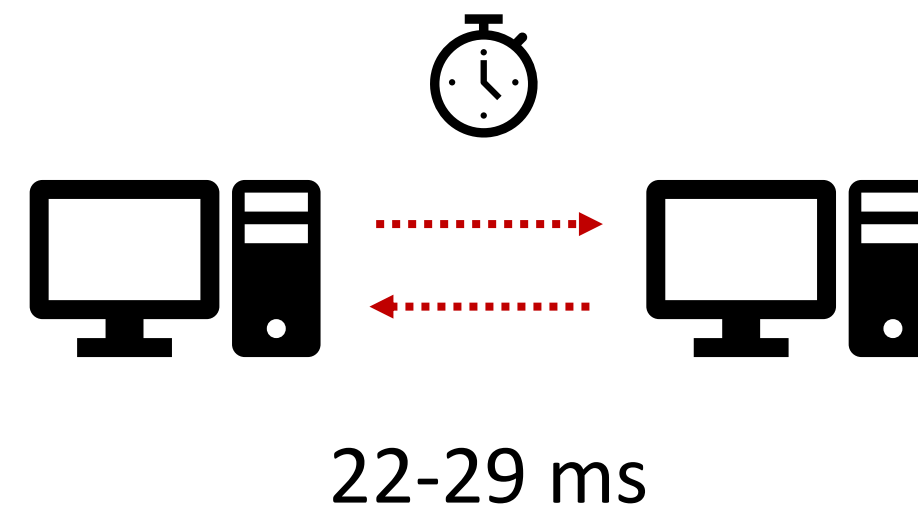


VR Round-Trip Time Latency

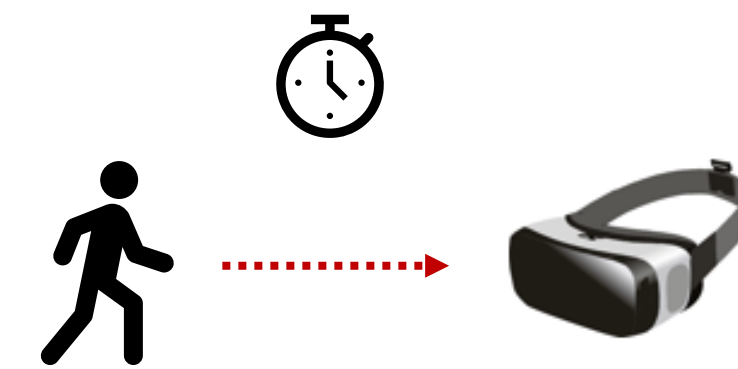


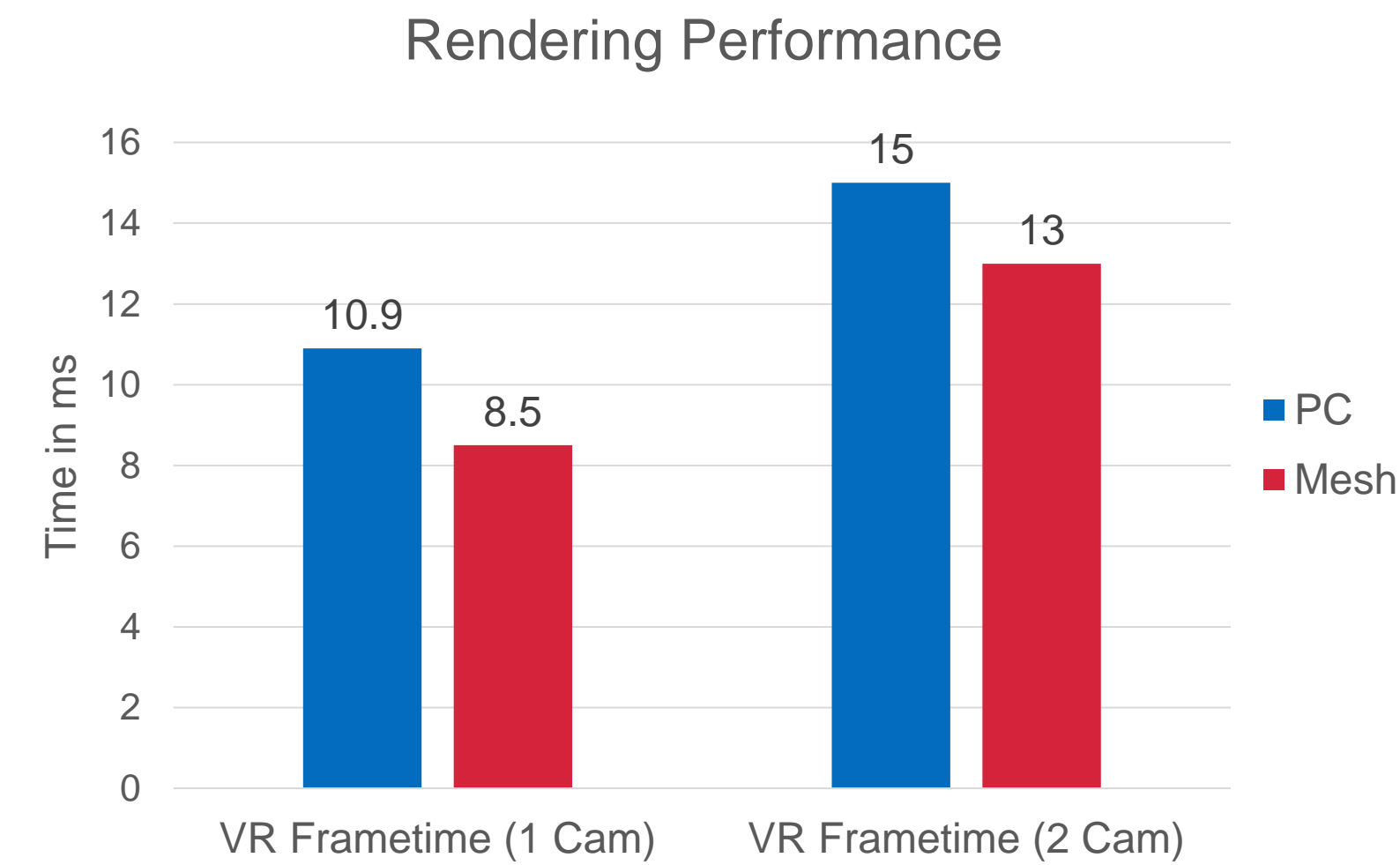


VR Round-Trip Time Latency

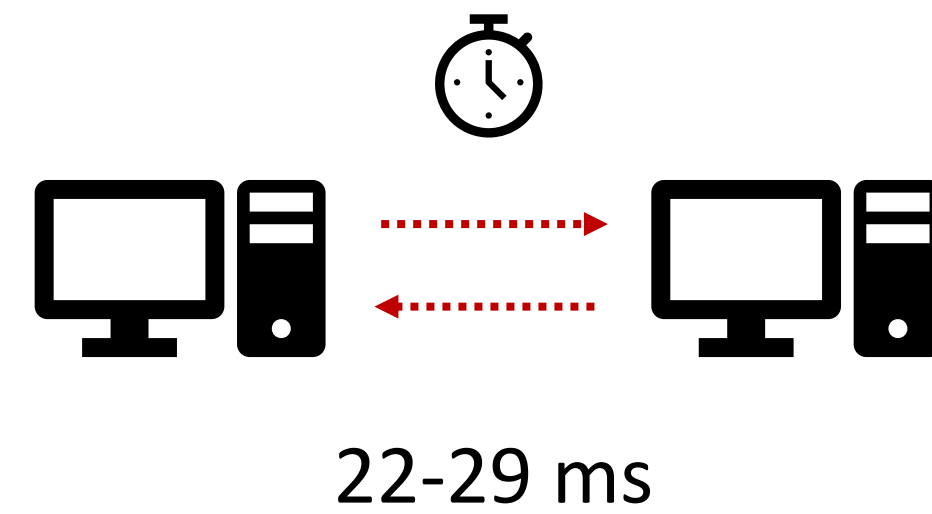


PC Motion-to-Photon

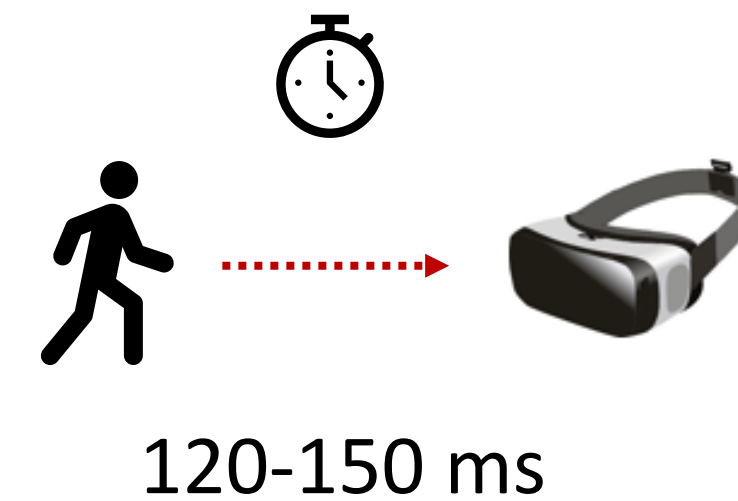


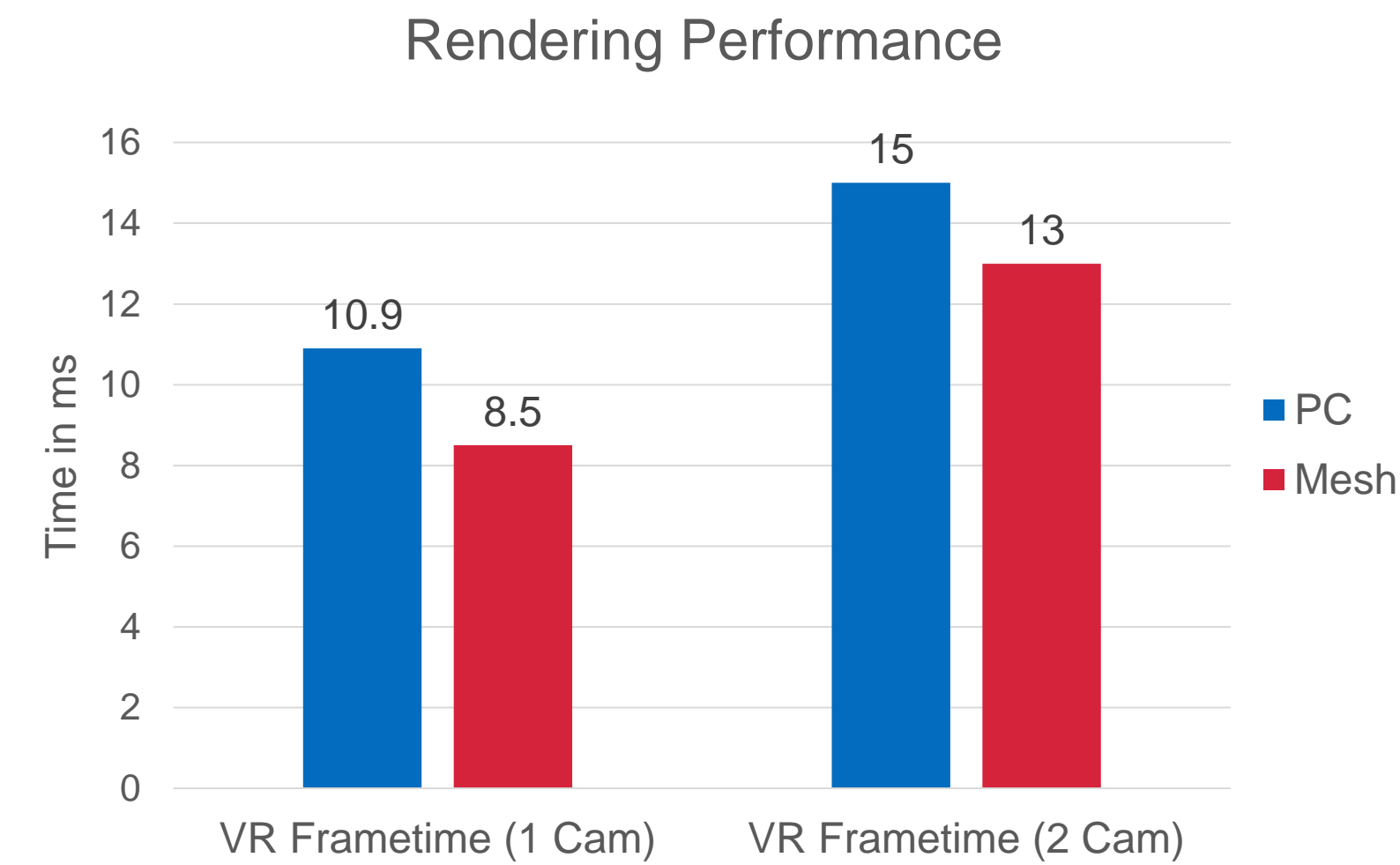


VR Round-Trip Time Latency

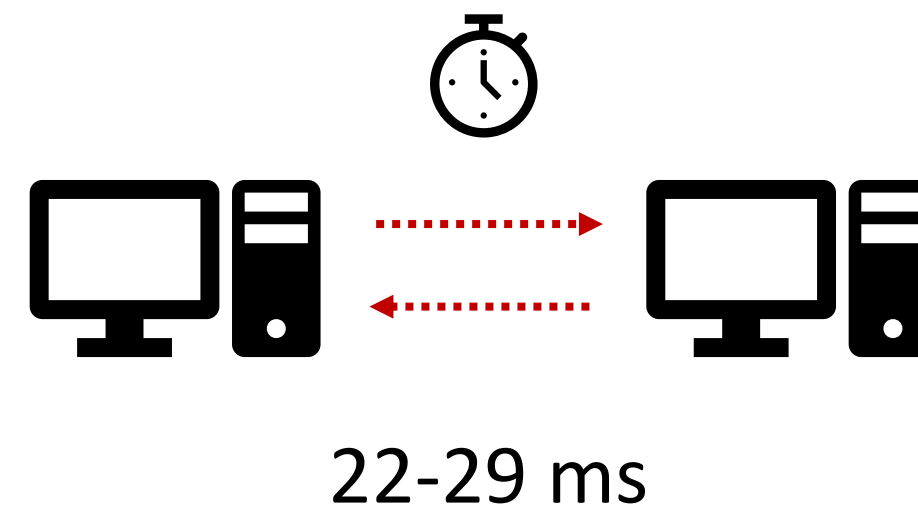


PC Motion-to-Photon

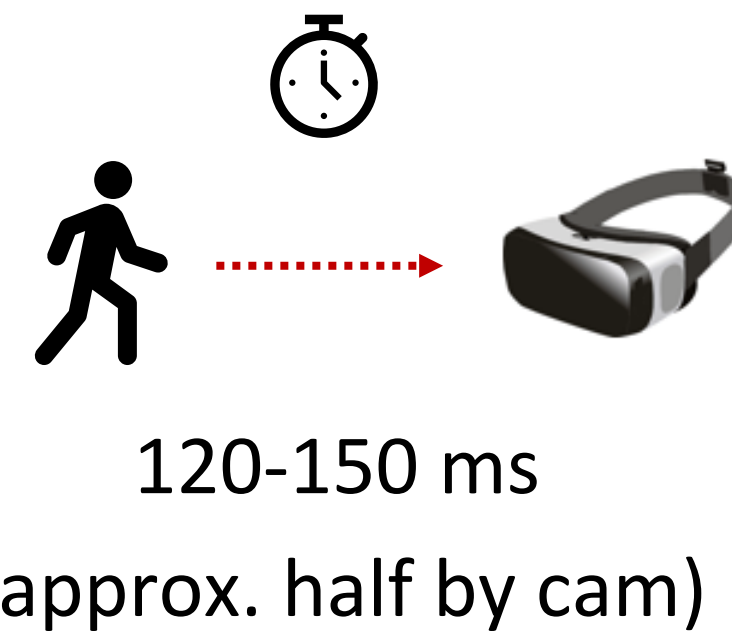




VR Round-Trip Time Latency



PC Motion-to-Photon



User Studies

- 7/12 doctors/medical students with OR experience

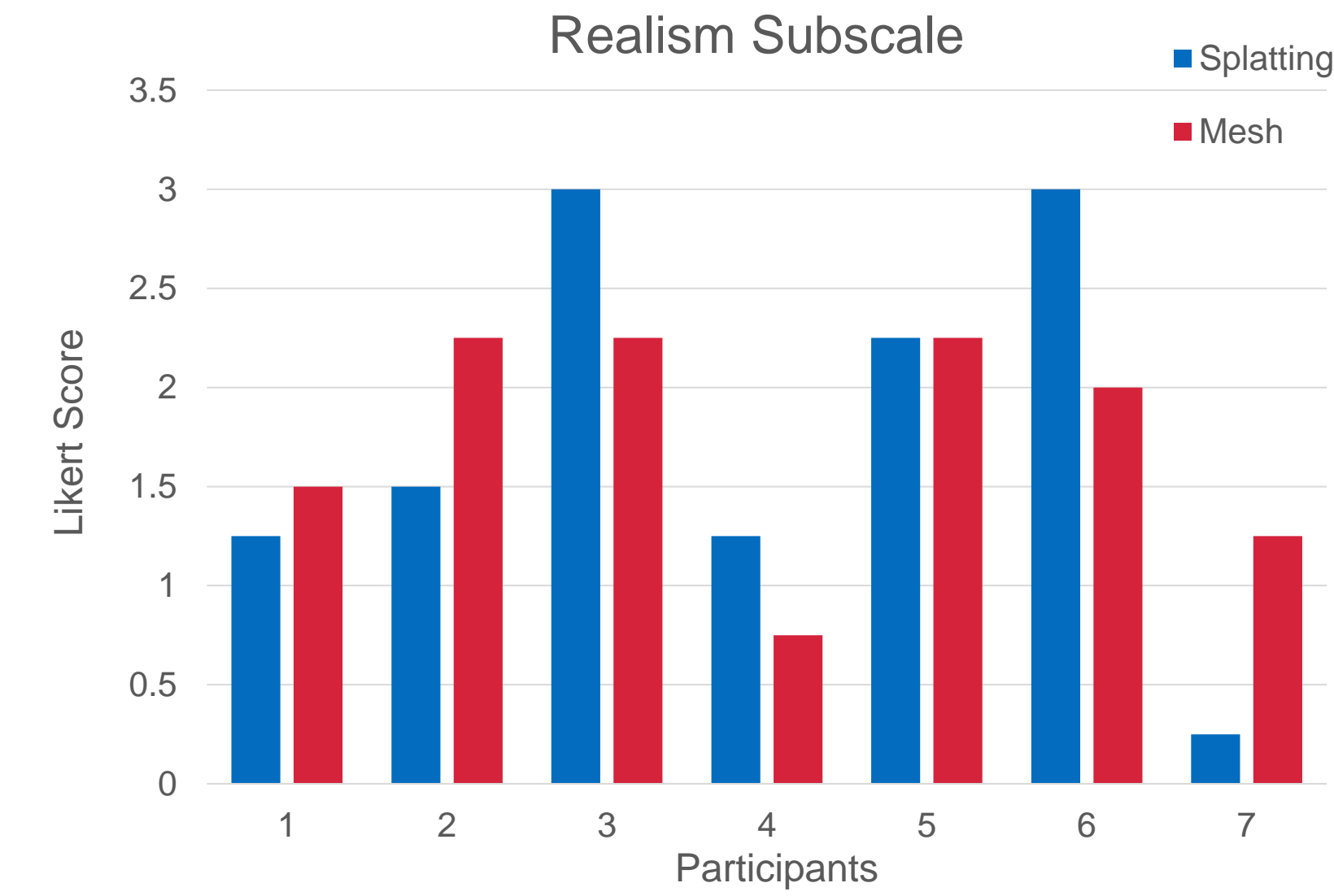
- 7/12 doctors/medical students with OR experience
- Task: remote lego building instructions while in VR and seeing point cloud vis.



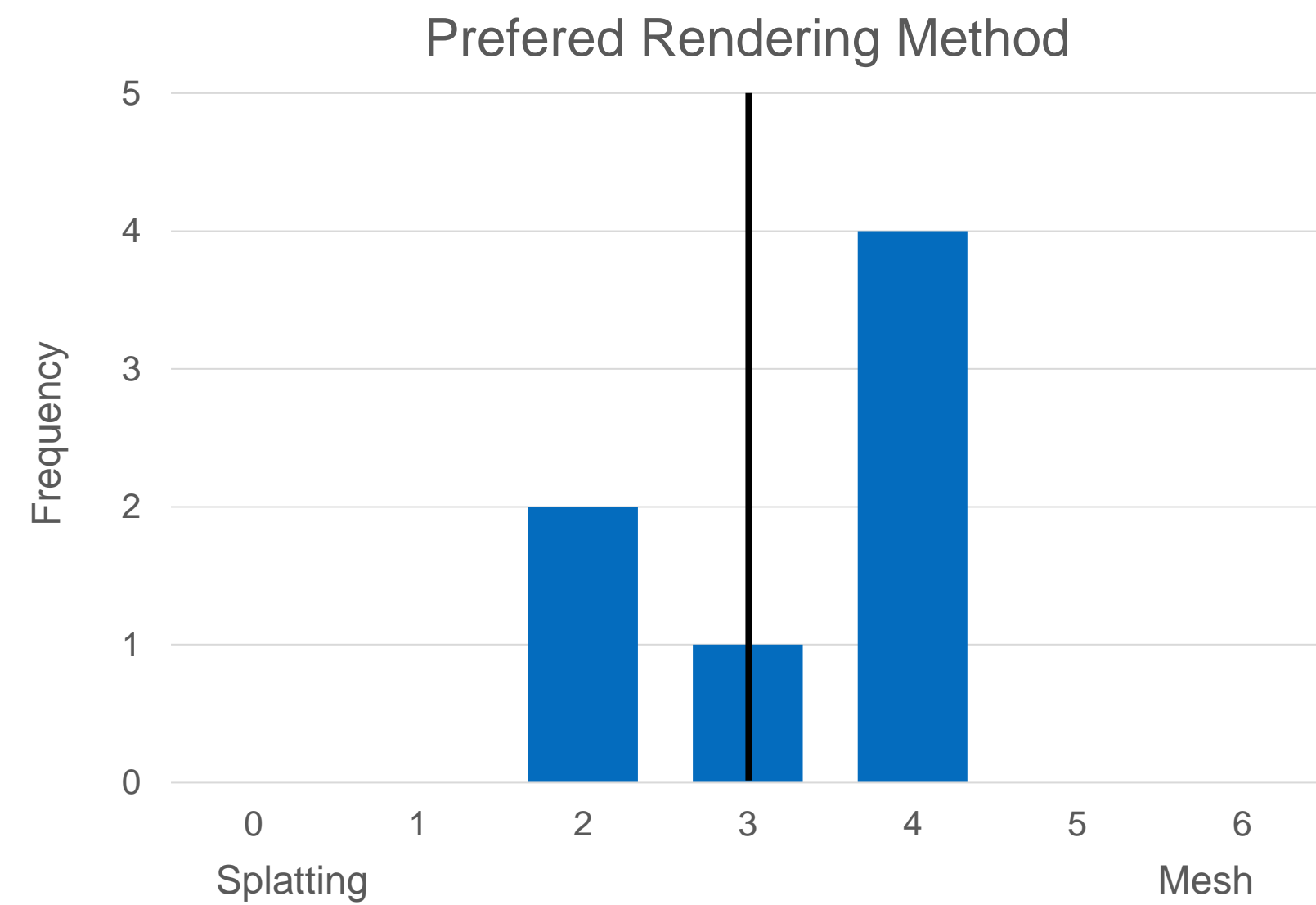
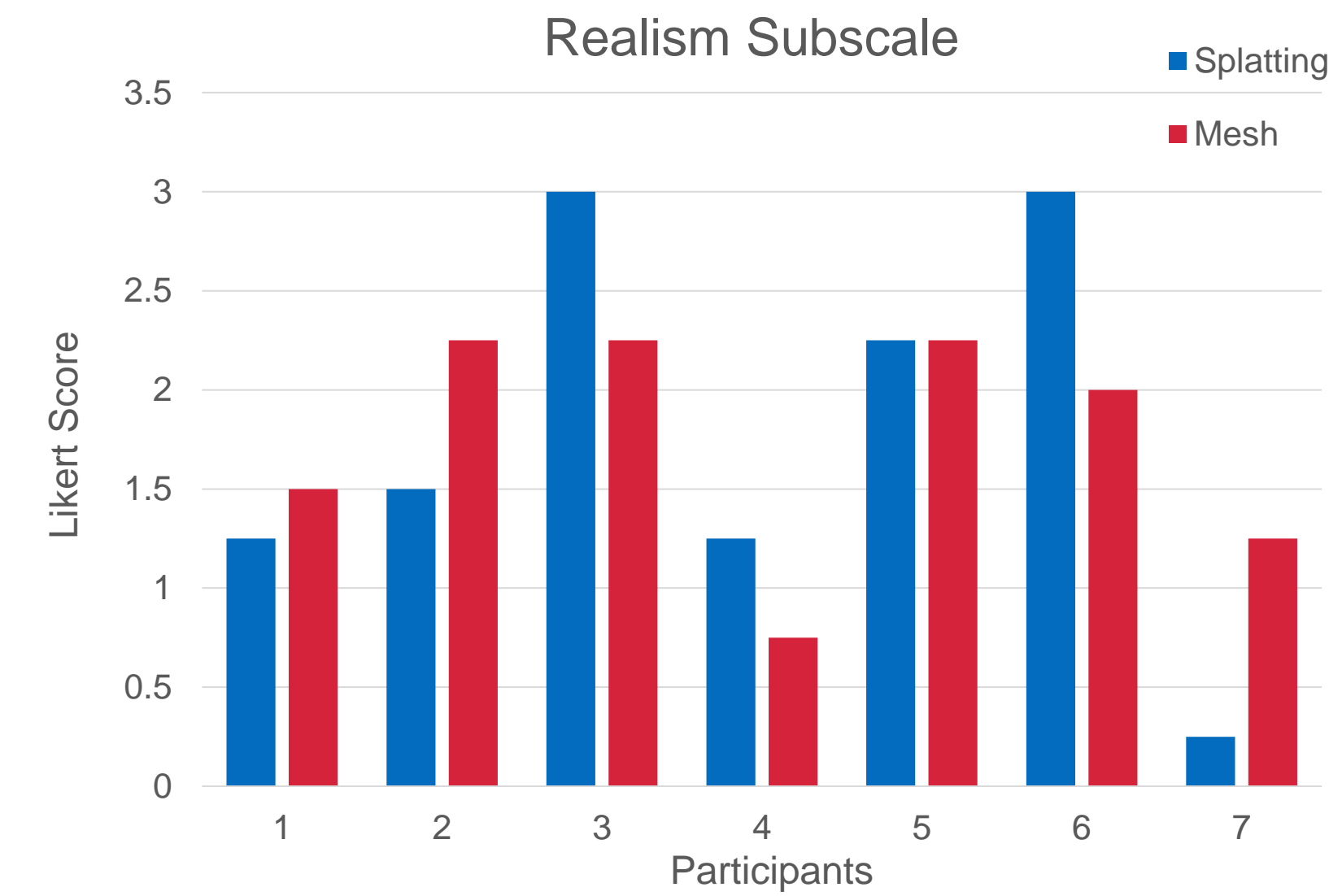
- 7/12 doctors/medical students with OR experience
- Task: remote lego building instructions while in VR and seeing point cloud vis.
- IPQ questionnaire (presence, involvement, realism) + custom questions



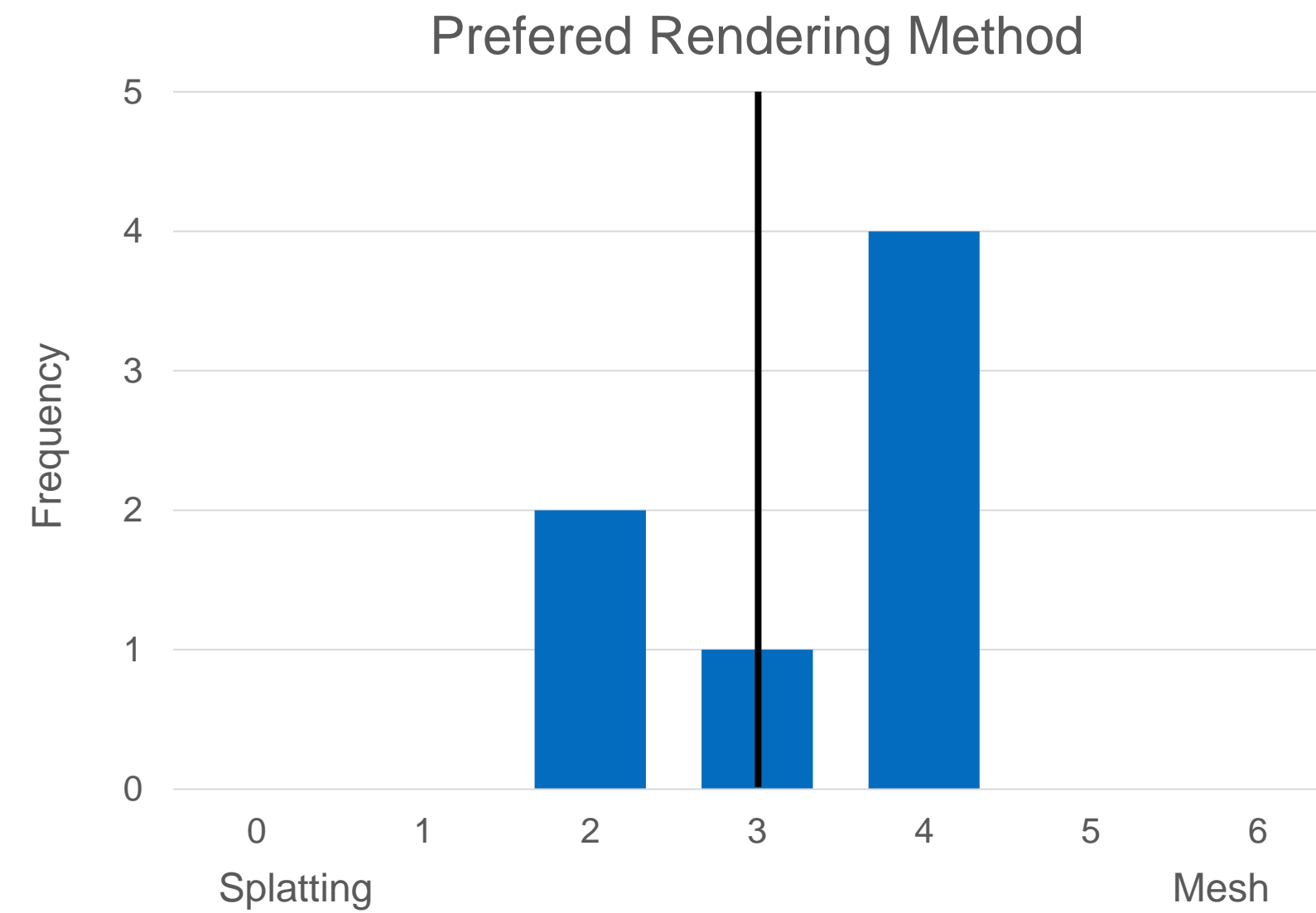
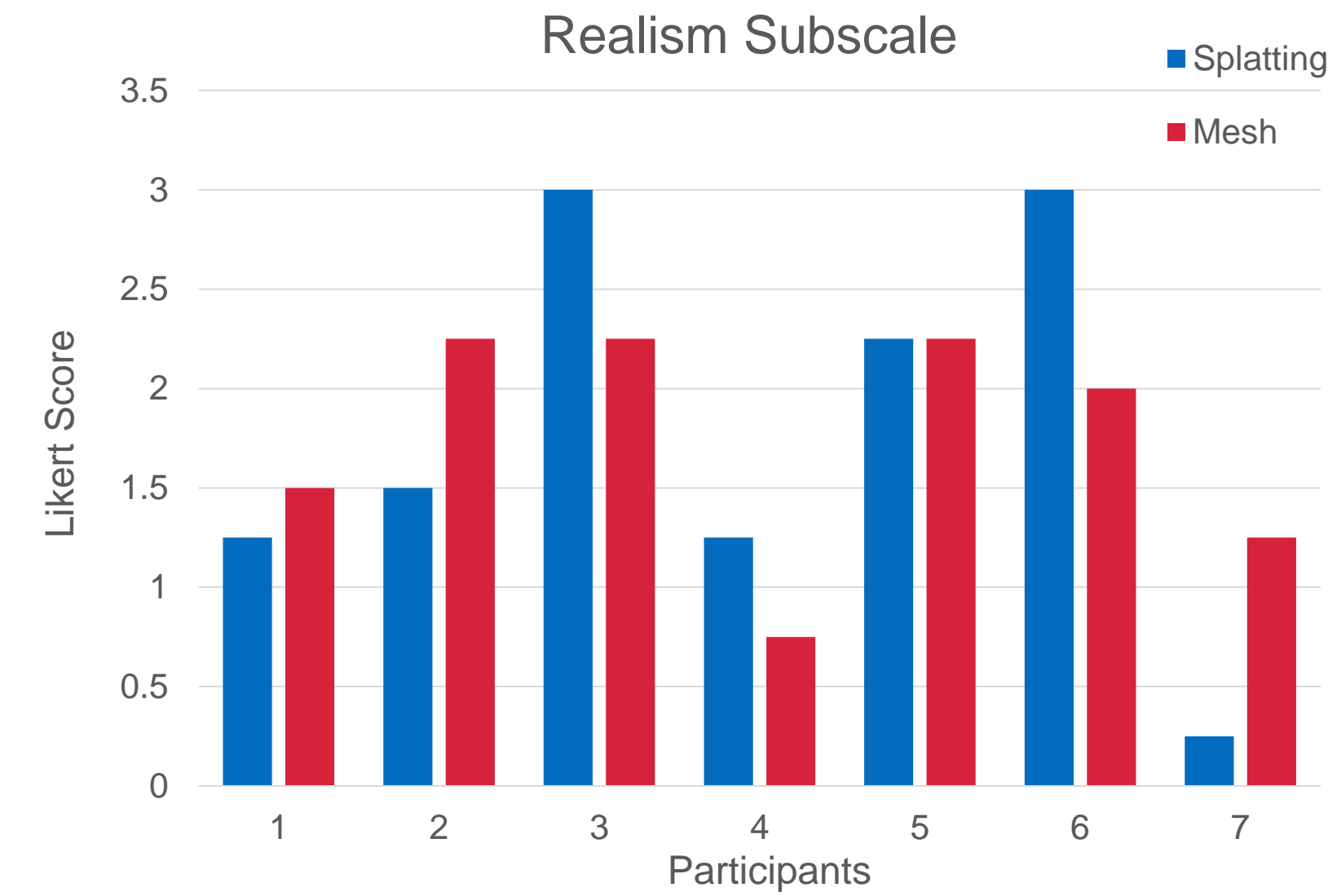
User Study 1: Renderers & Use Cases



User Study 1: Renderers & Use Cases

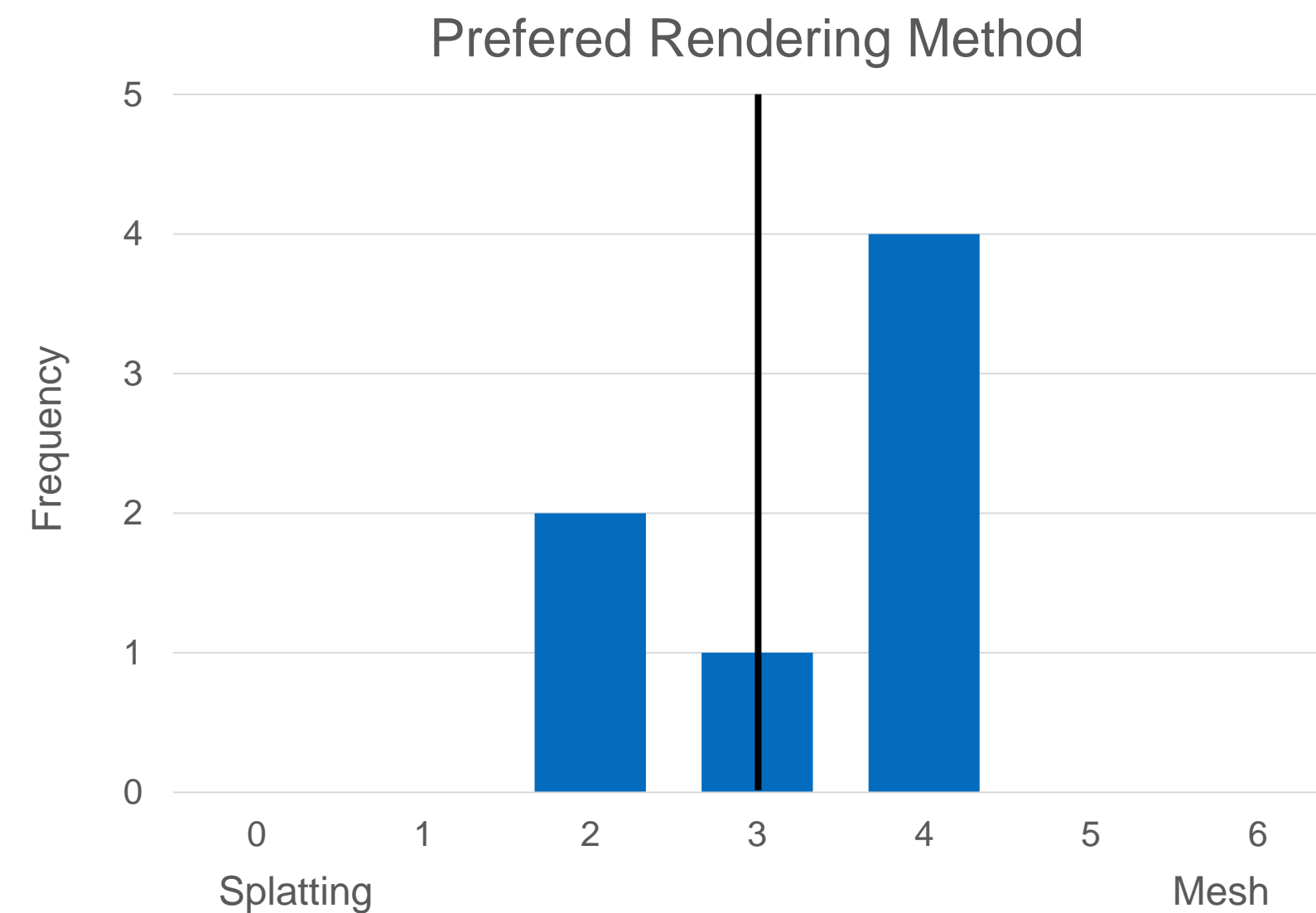
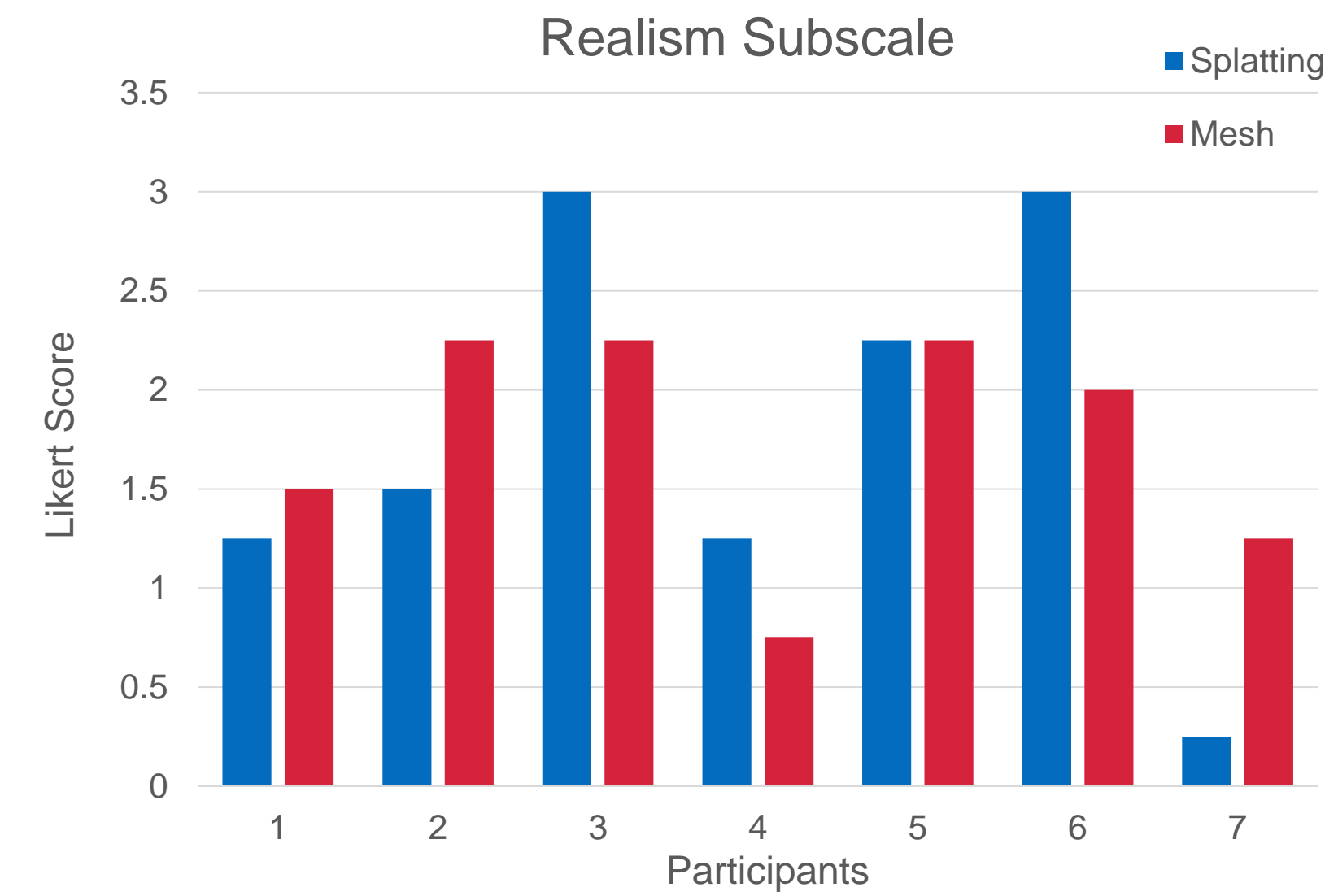


User Study 1: Renderers & Use Cases



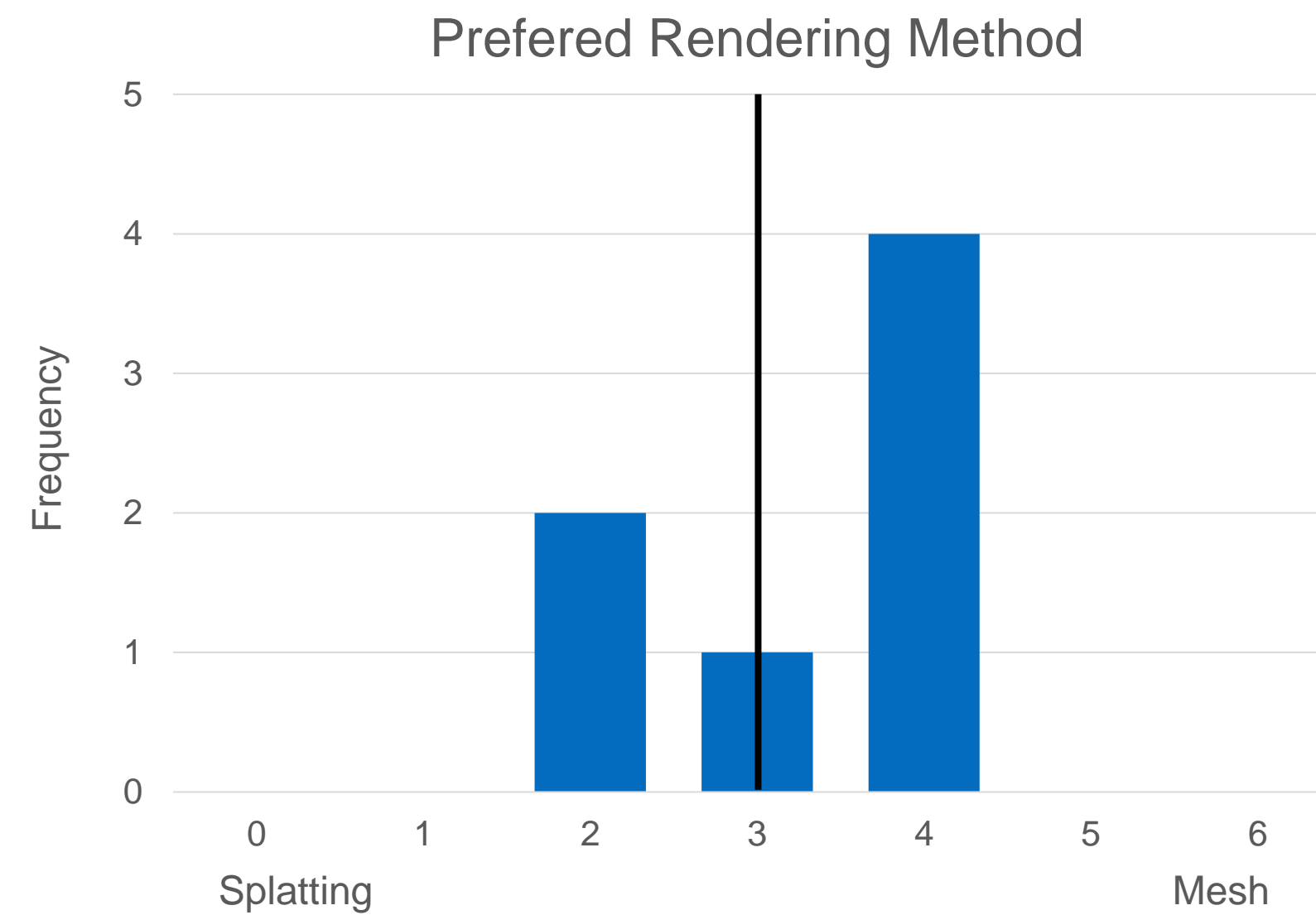
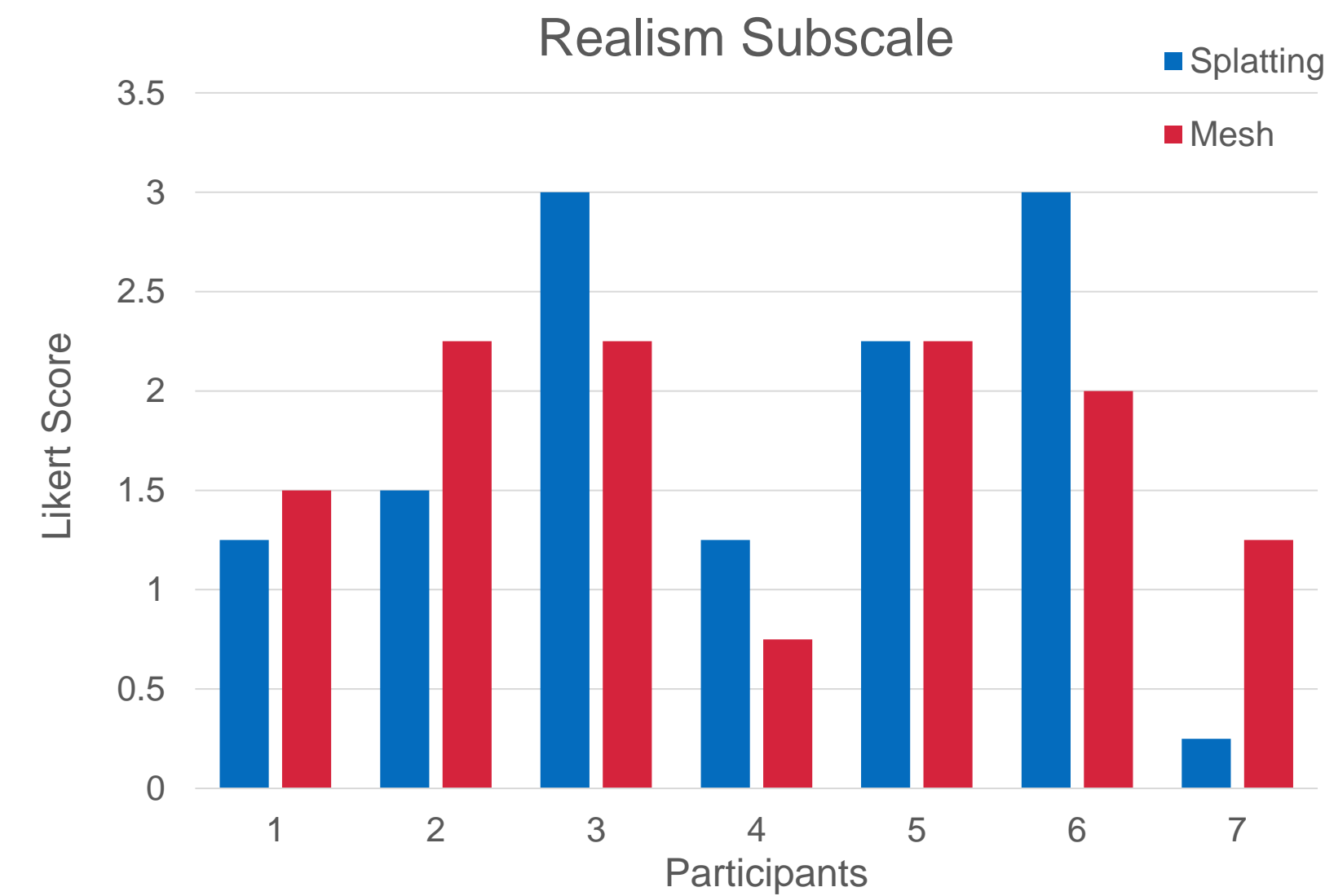
- Renderers rated similar, slight preference for mesh

User Study 1: Renderers & Use Cases



- Renderers rated similar, slight preference for mesh
- No significance due to low sample size (7)

User Study 1: Renderers & Use Cases



- Renderers rated similar, slight preference for mesh
- No significance due to low sample size (7)

• Use cases:

Emergency operations

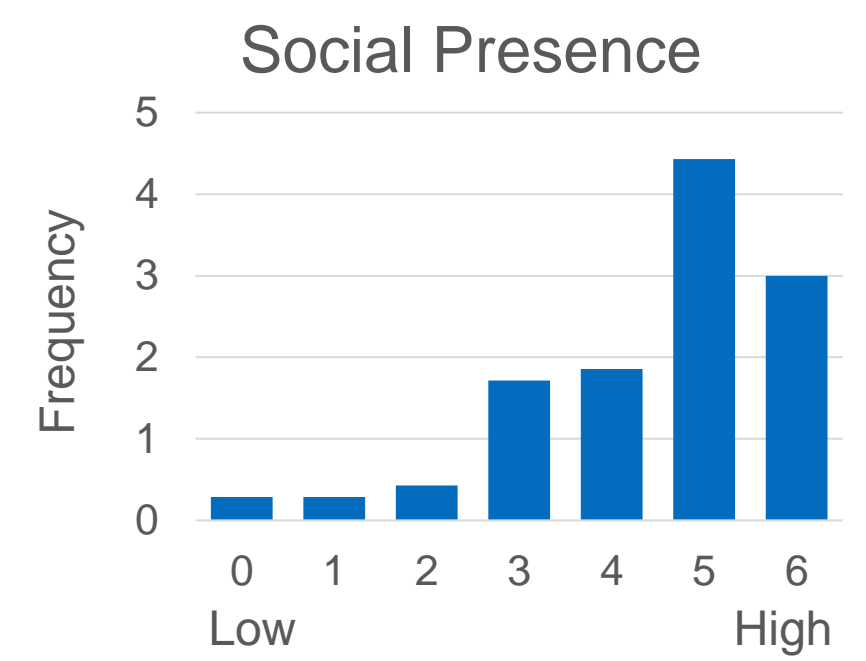
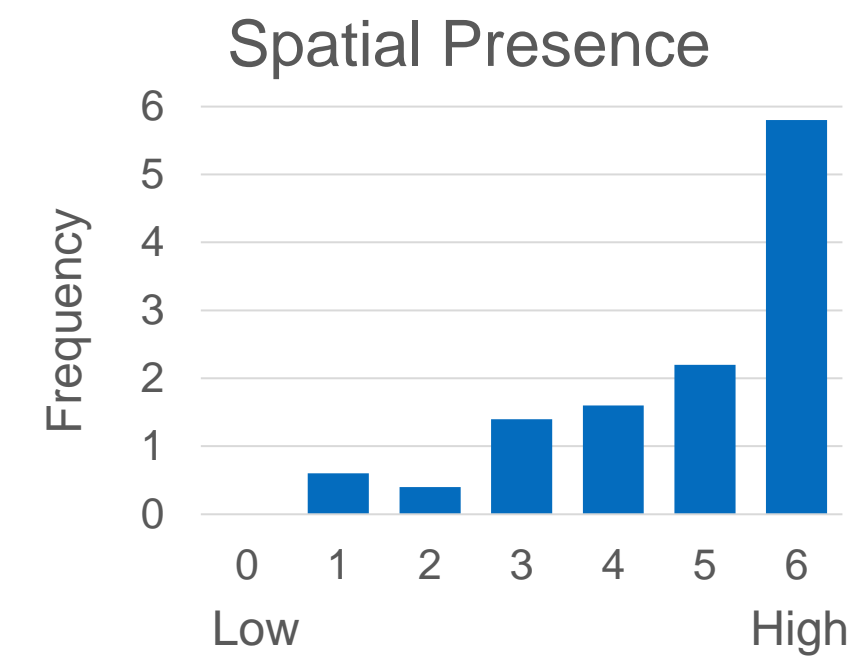
Inexperienced doctors on duty

Educational operations

Teaching in general

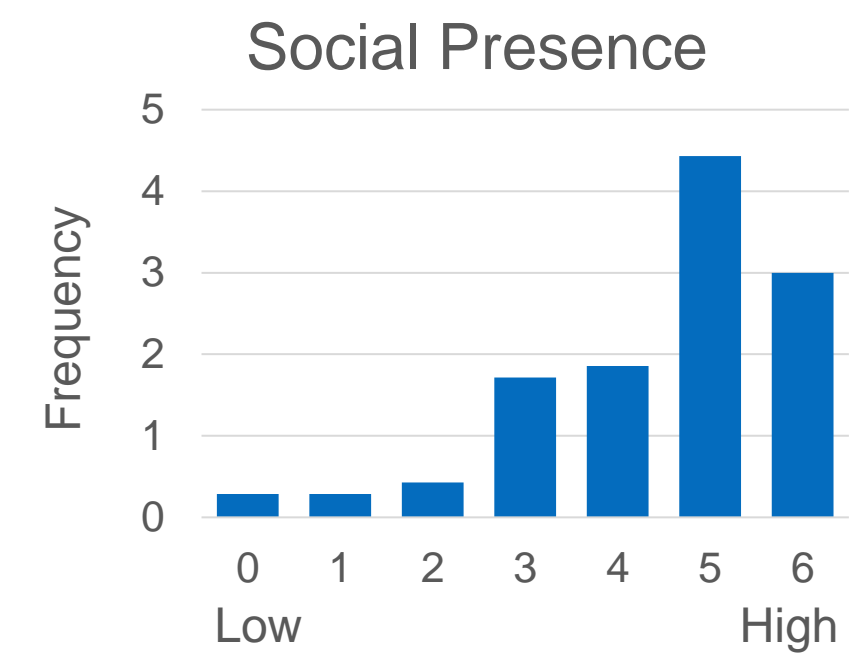
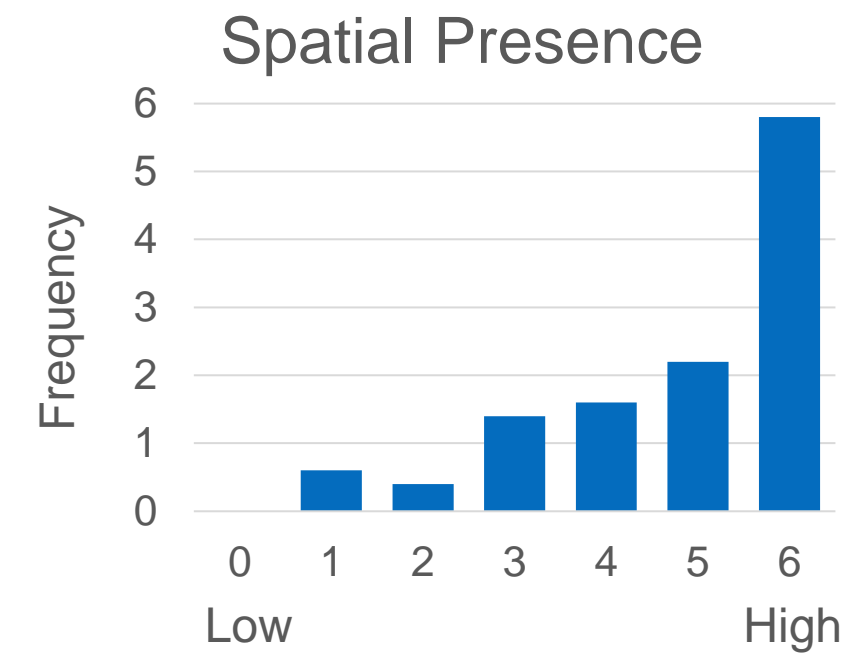
User Study 2 : General Feedback & Presence

- High spatial and social presence

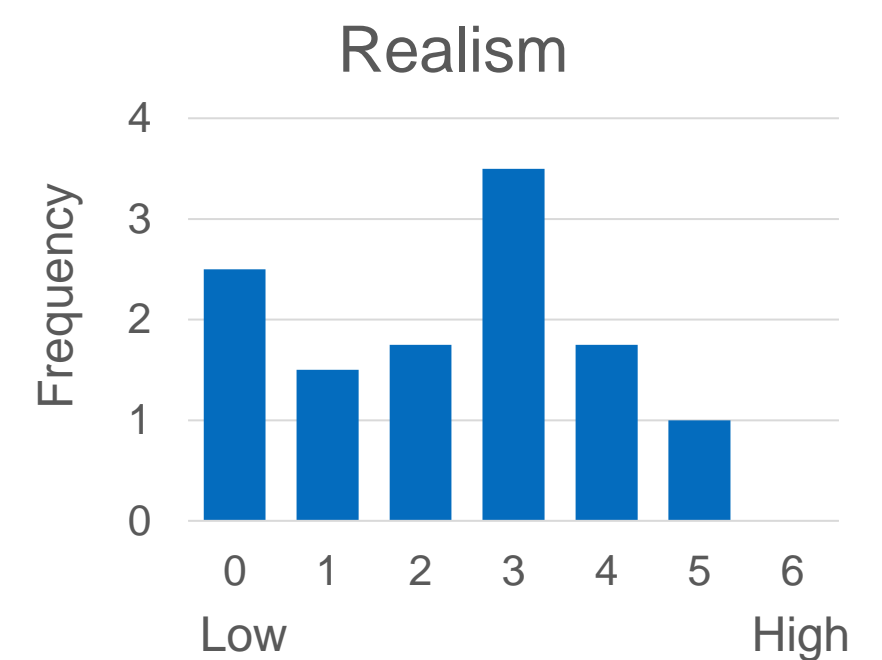
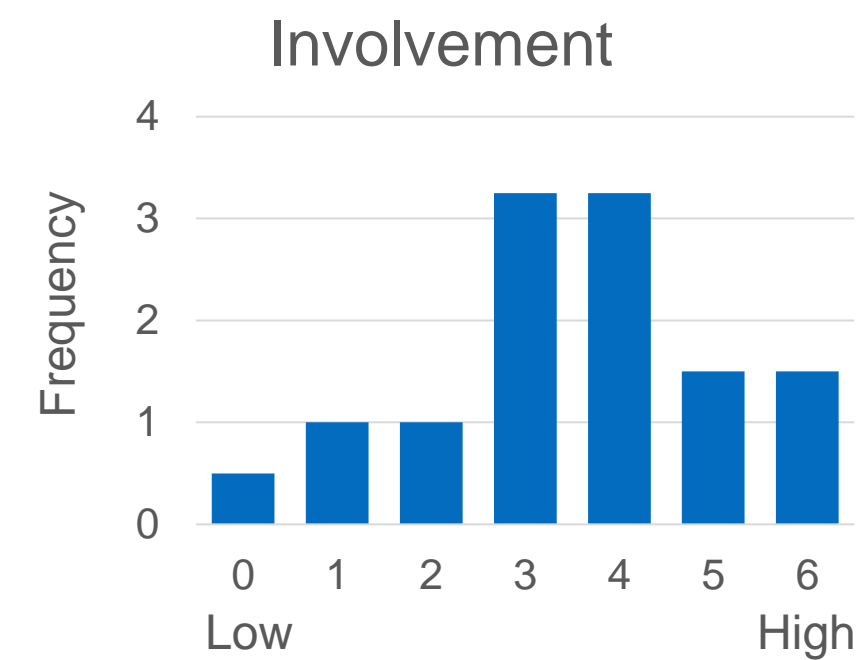


User Study 2 : General Feedback & Presence

- High spatial and social presence

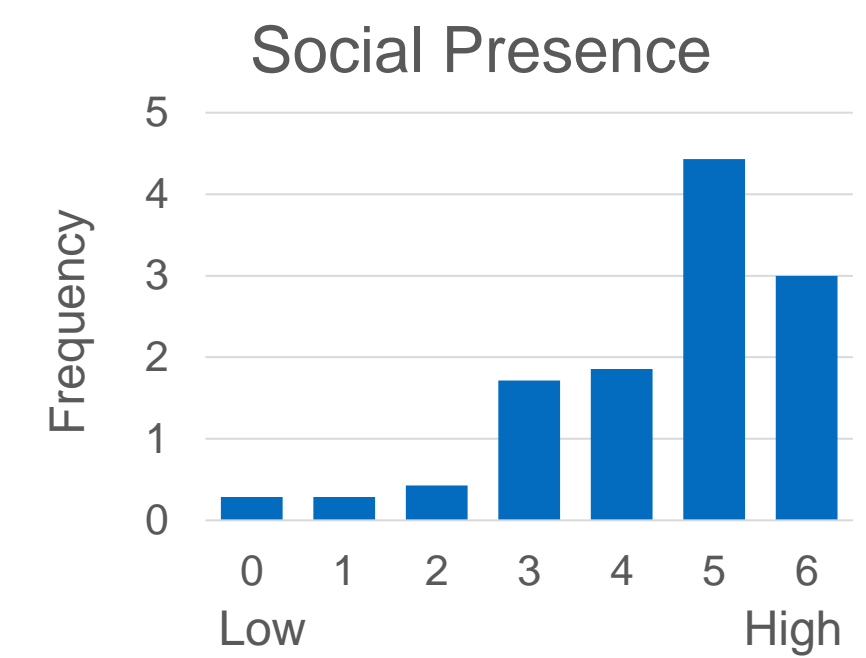
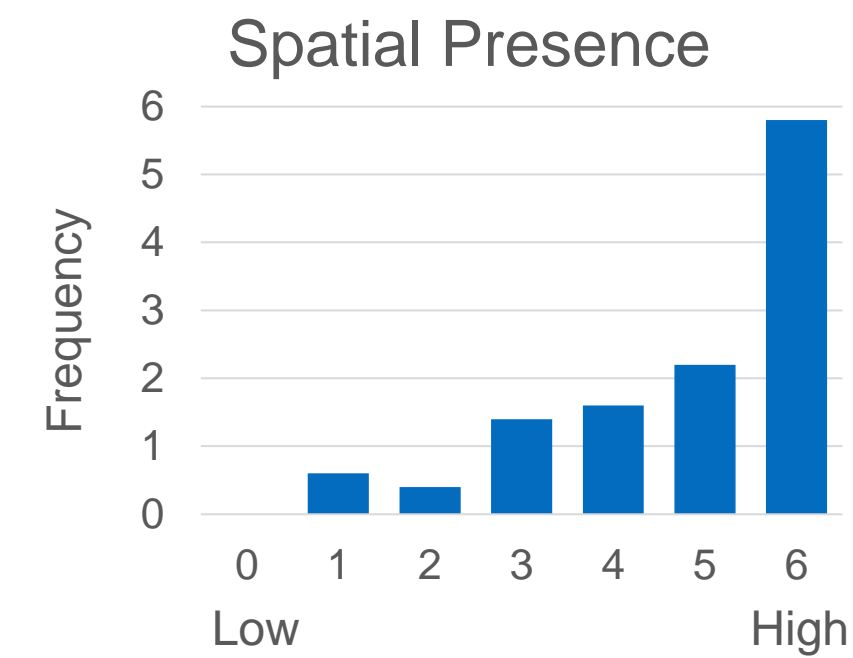


- Moderate involvement and realism

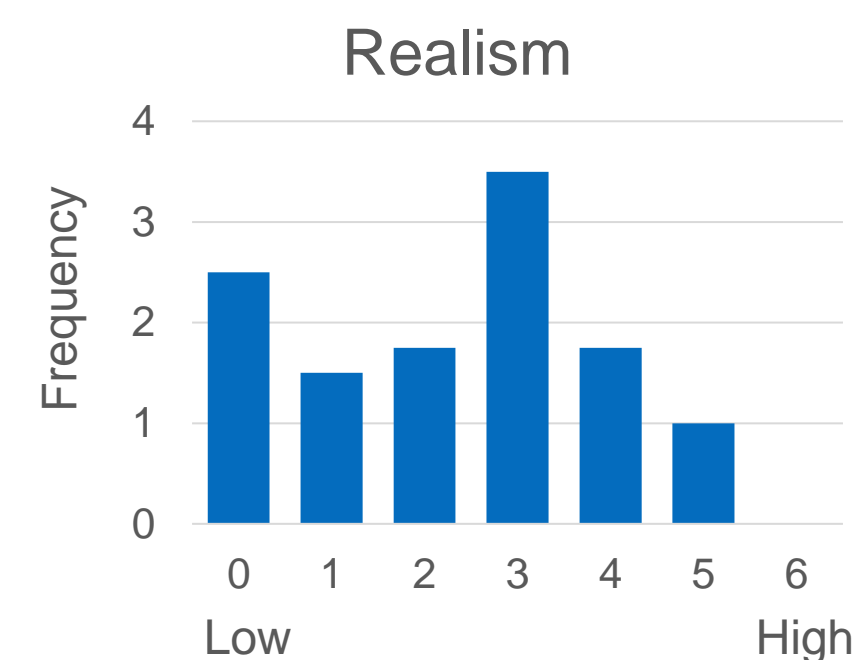
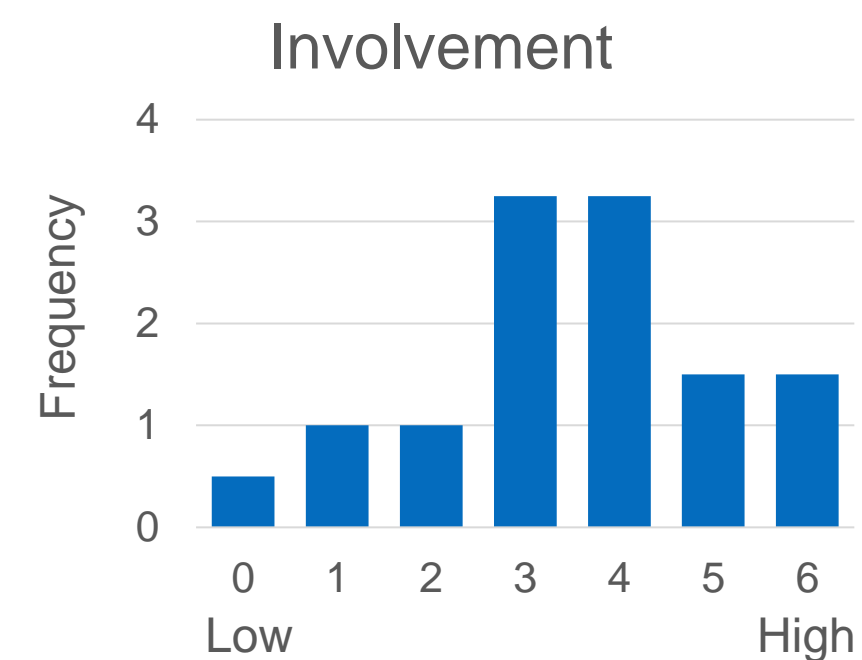


User Study 2 : General Feedback & Presence

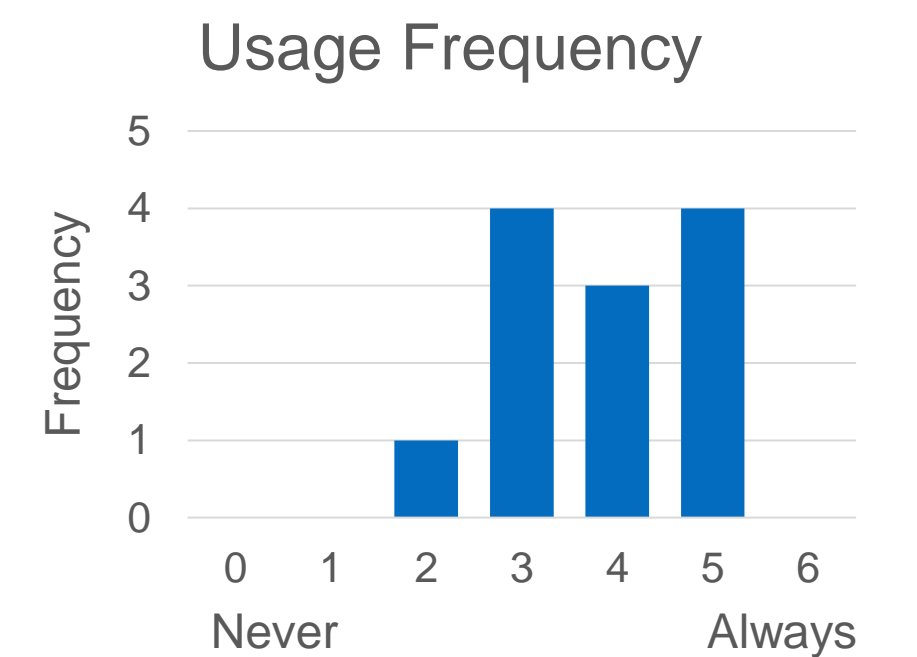
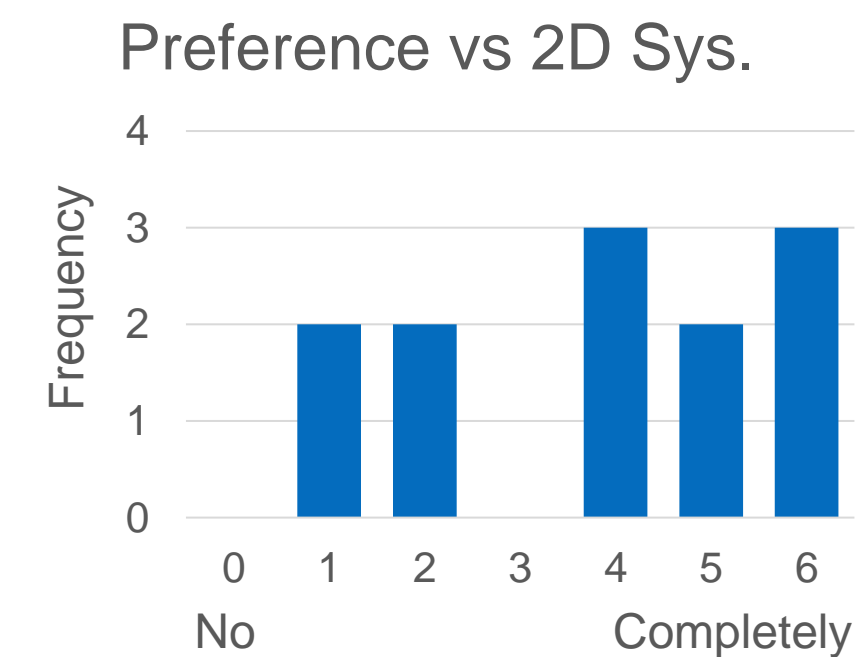
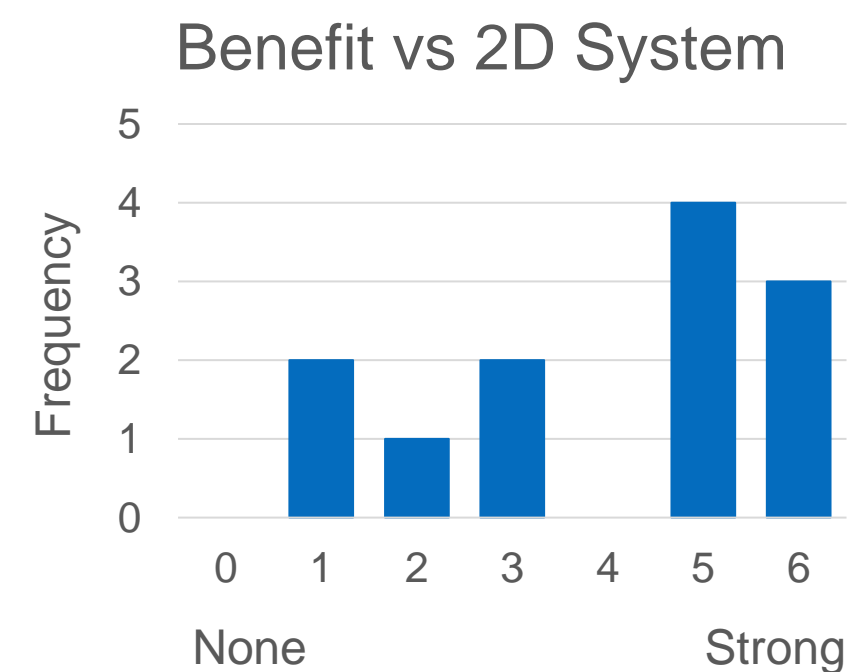
- High spatial and social presence



- Moderate involvement and realism

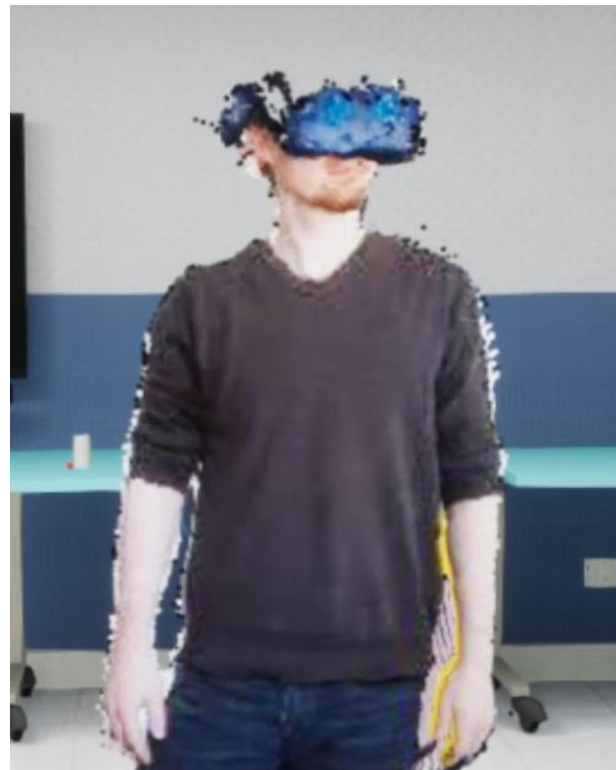


- Most doctors see advantages over video and would like to use it



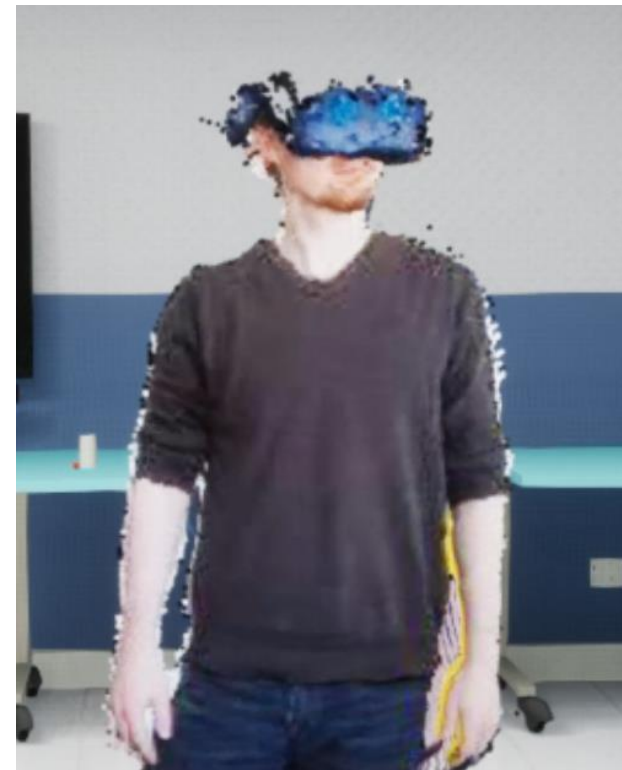
Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction



Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction
 - Eye + mouth tracking, 3D Morphable Face Model



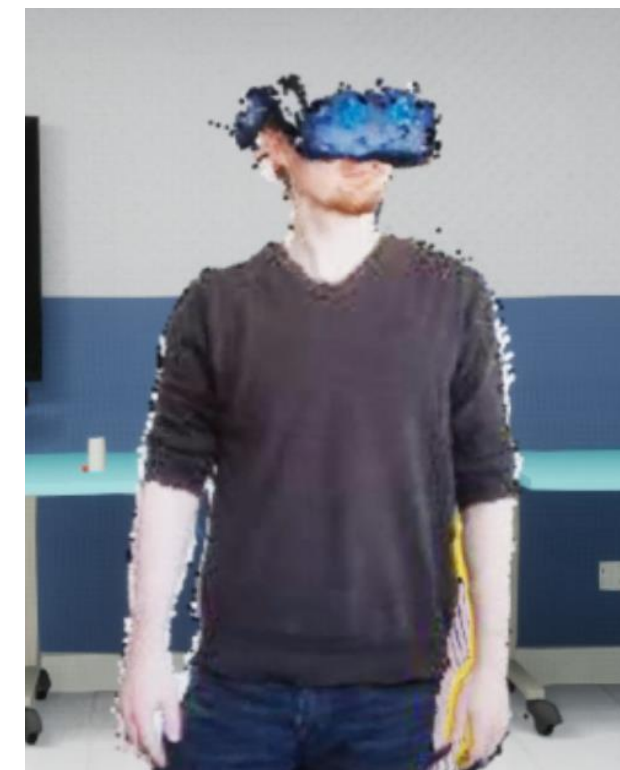
Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction
 - Eye + mouth tracking, 3D Morphable Face Model
 - 6 pre-animated facial expressions + self-trained neural network for selection



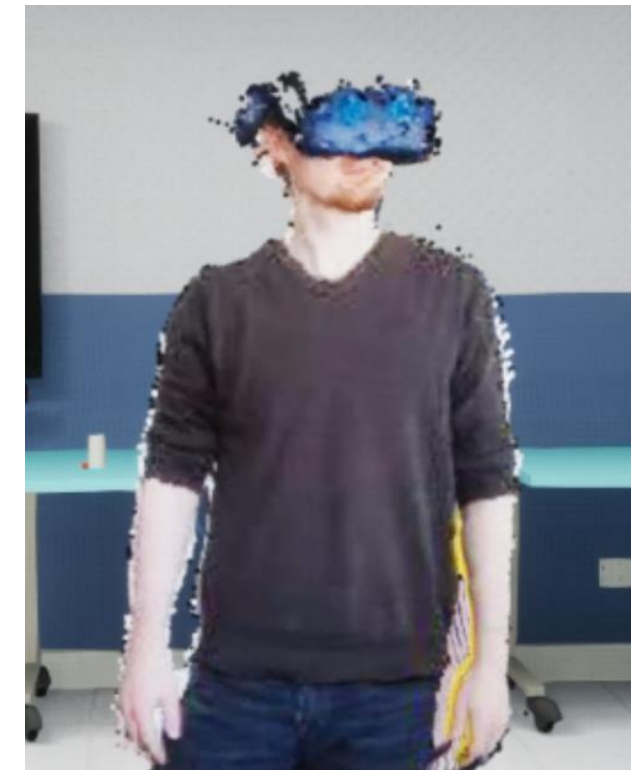
Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction
 - Eye + mouth tracking, 3D Morphable Face Model
 - 6 pre-animated facial expressions + self-trained neural network for selection
1. Initially per user: take 6 images, adapt face model



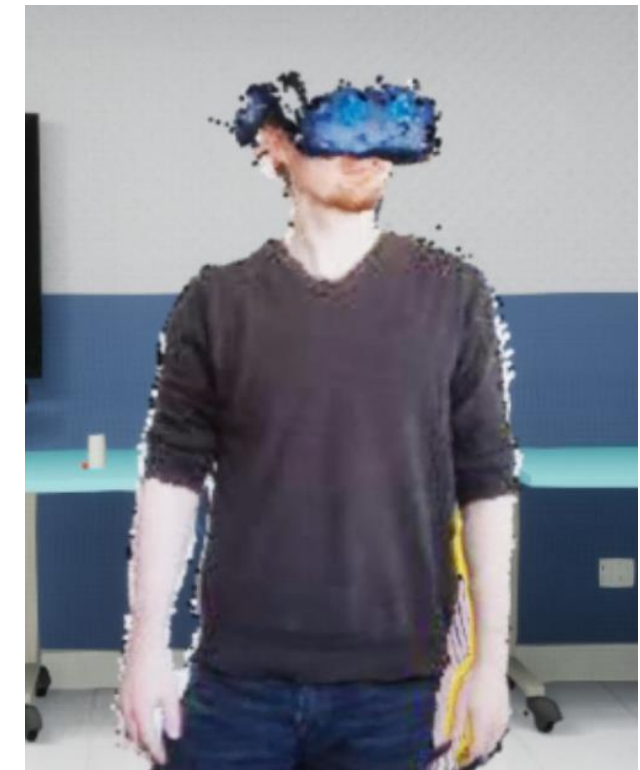
Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction
 - Eye + mouth tracking, 3D Morphable Face Model
 - 6 pre-animated facial expressions + self-trained neural network for selection
1. Initially per user: take 6 images, adapt face model
 2. At runtime: select and interpolate facial expressions and textures



Limitation: Occluded Avatar Faces

- Work in progress: face reconstruction
 - Eye + mouth tracking, 3D Morphable Face Model
 - 6 pre-animated facial expressions + self-trained neural network for selection
1. Initially per user: take 6 images, adapt face model
 2. At runtime: select and interpolate facial expressions and textures
 3. Render face model



- New immersive VR telepresence system for remote collaboration

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline
 - 120-150 ms motion-to-photon latency

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline
 - 120-150 ms motion-to-photon latency
 - 23.4 Mbit/s per camera (lossless)

Conclusion

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline
 - 120-150 ms motion-to-photon latency
 - 23.4 Mbit/s per camera (lossless)
- Promising avatar face reconstruction prototype

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline
 - 120-150 ms motion-to-photon latency
 - 23.4 Mbit/s per camera (lossless)
- Promising avatar face reconstruction prototype
- High spatial and social presence, preferred by most doctors

- New immersive VR telepresence system for remote collaboration
- Efficient RGB-D streaming and point cloud rendering pipeline
 - 120-150 ms motion-to-photon latency
 - 23.4 Mbit/s per camera (lossless)
- Promising avatar face reconstruction prototype
- High spatial and social presence, preferred by most doctors
- Slight preference for mesh rendering

- Improve pre-processing using deep learning

- Improve pre-processing using deep learning
- Improve compression using adapted/hybrid video compression

Future Work

- Improve pre-processing using deep learning
- Improve compression using adapted/hybrid video compression
- Improve fidelity using AI up-sampling and dedicated color/stereo cameras

Future Work

- Improve pre-processing using deep learning
- Improve compression using adapted/hybrid video compression
- Improve fidelity using AI up-sampling and dedicated color/stereo cameras
- Explore point cloud/mesh fusion

Future Work

- Improve pre-processing using deep learning
- Improve compression using adapted/hybrid video compression
- Improve fidelity using AI up-sampling and dedicated color/stereo cameras
- Explore point cloud/mesh fusion
- Integrate AR HMDs

Thank you for your attention!

Questions?

