

Human-Centric Bandwidth-Efficient Volumetric Video Streaming

Presenter: *SHI Yuang*

Supervisor: *Prof. OOI Wei Tsang*

National University of Singapore

Table of Content

- Volumetric Video
- Volumetric Video Streaming
 - Point Cloud Compression
 - Tiling (Segmentation)
 - Quality Assessment
- Summary of Our Contributions and Future Work

Human-Centric Bandwidth-Efficient Volumetric Video Streaming

Volumetric Video

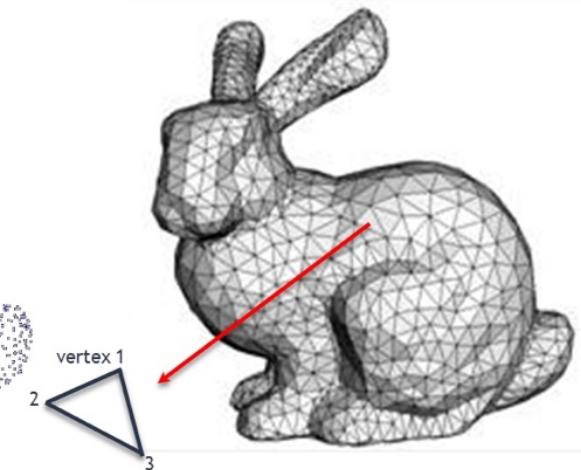
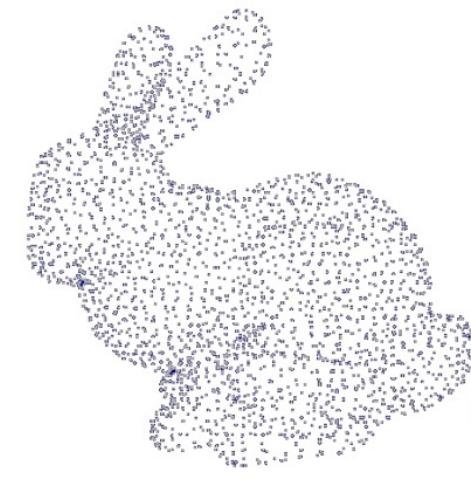
- Captures actors or subjects in 3D.
- Viewers can watch from any conceivable angle.



Video Source: https://www.youtube.com/watch?v=9qd276AJgo&embeds_referring_euri=https%3A%2F%2Fhubblecontent.osi.office.net%2F&source_ve_path=MjM4NTE&feature=emb_title

Point Cloud VS. Mesh

- Point Cloud
 - A set of points
 - Each point consists of position (X,Y,Z), texture (R, G, B)
- Mesh
 - Consists of polygons that are made of X, Y, and Z coordinates and lines

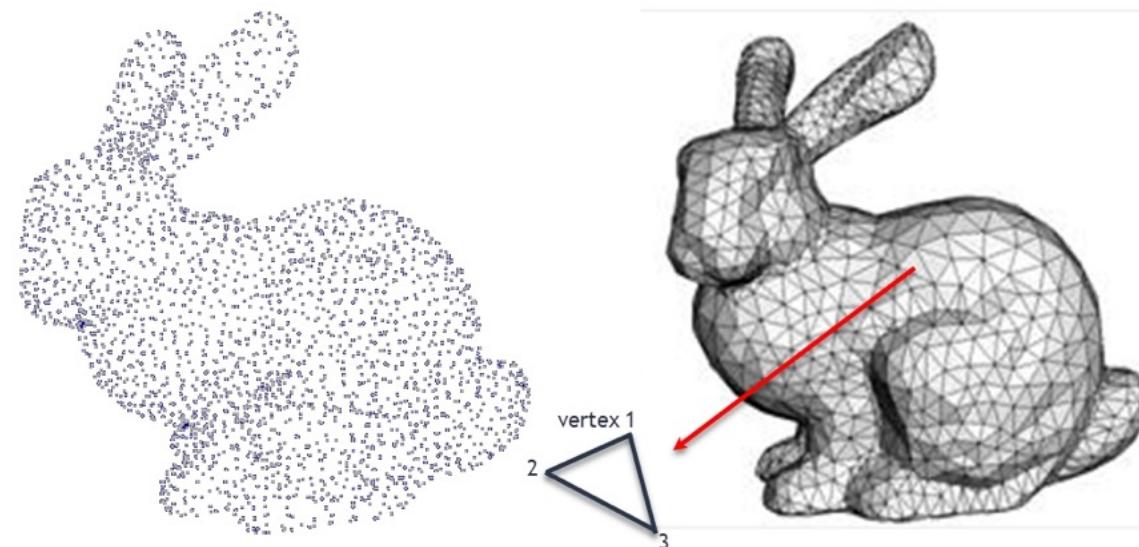


Point Cloud

Mesh

Point Cloud VS. Mesh

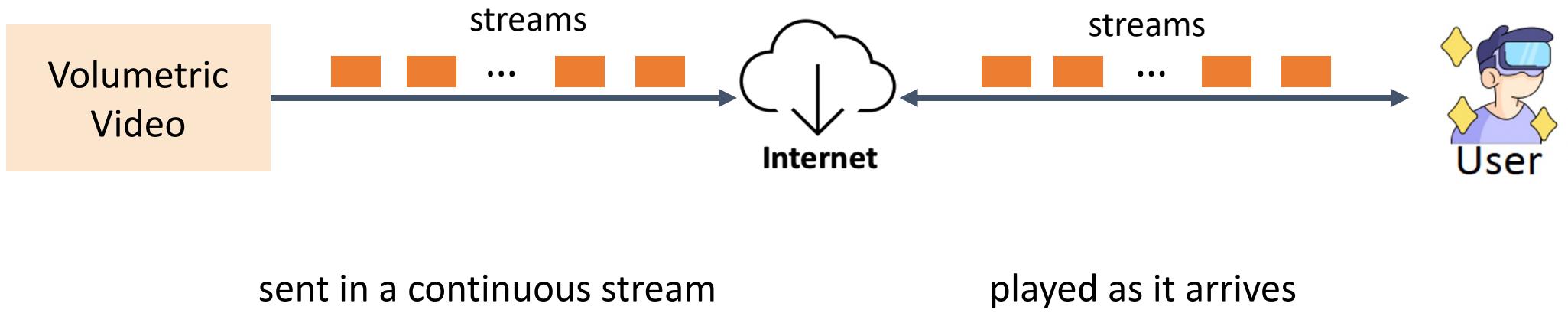
- **Point Cloud**: each point can be treated **independently**. Natural and straight forward.
- Mesh: relies on the **topology** and **connectivity** between vertices.



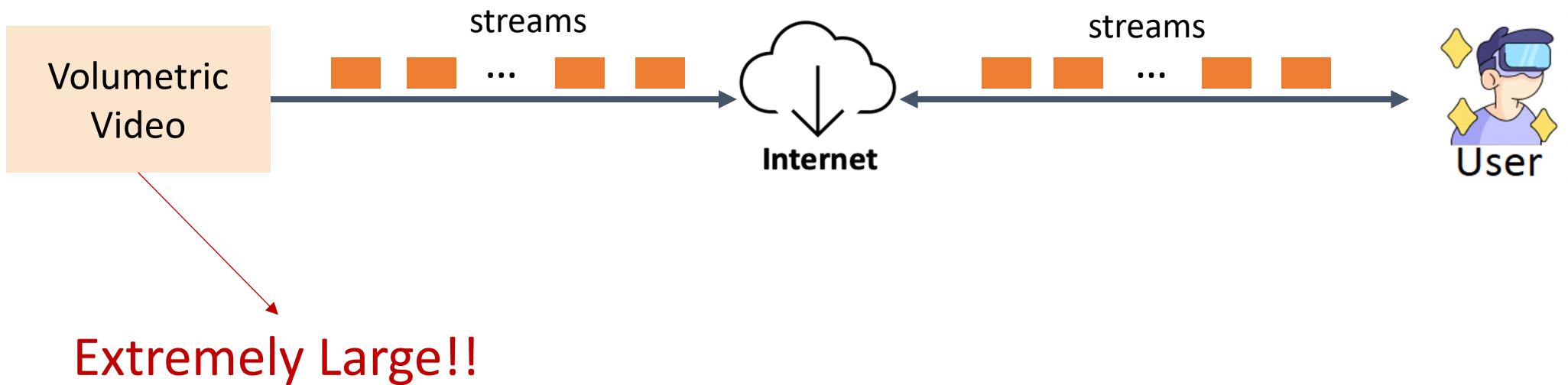
Point Cloud is relatively **simple** and **flexible** to handle.

Human-Centric Bandwidth-Efficient Volumetric Video Streaming

Volumetric Video Streaming

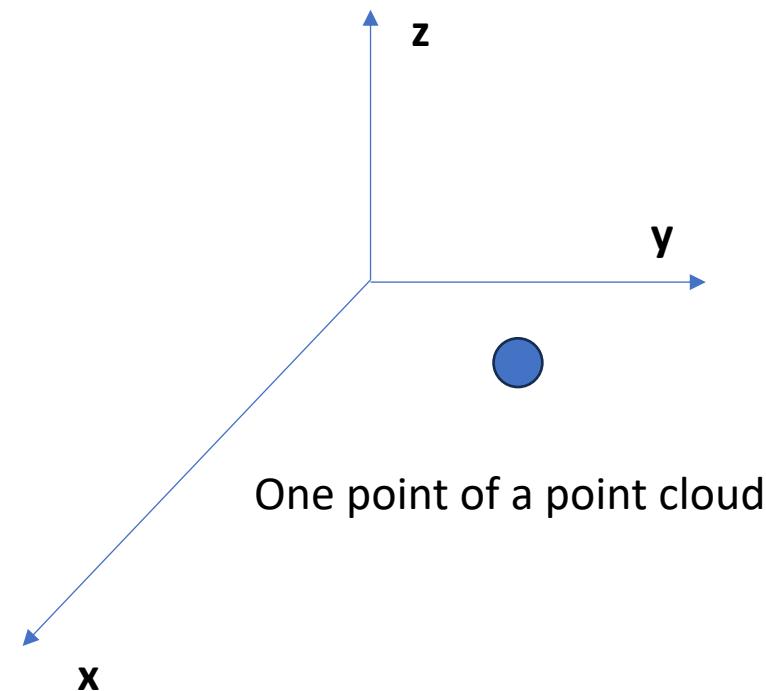


Volumetric Video Streaming



Point Cloud

- A point consists of
 - Position: (X, Y, Z)
 - 3 bytes per coordinate
 - Texture: (R,G,B)
 - 1 bytes per channel
 - **15 bytes per point**



Point Cloud

- This frame is a point cloud
 - 800K points
 - 15 bytes per point
 - **12 MB**



Point Cloud

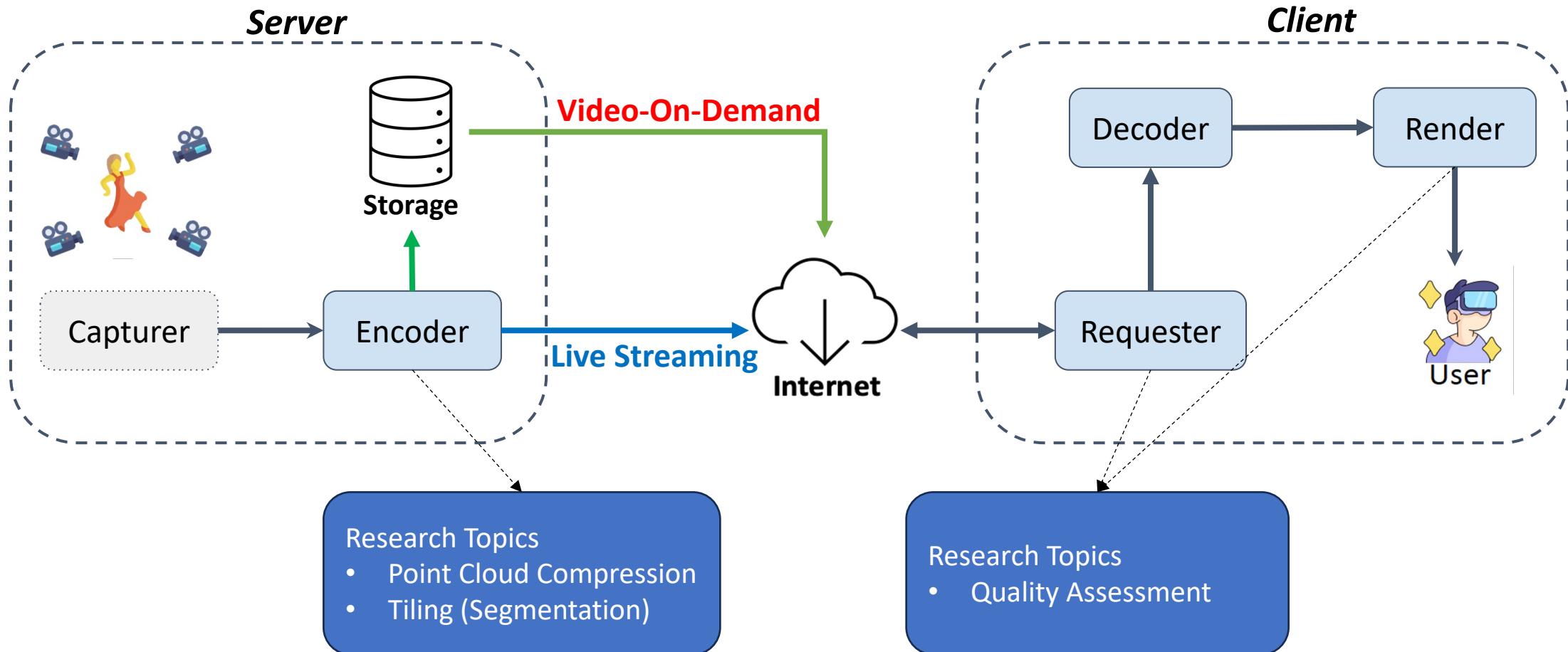
- For this 30-fps volumetric video:
 - $12 \text{ MB} * 30 \text{ fps} = 2.9 \text{ Gbps}$

Rank	Country/Territory	Median download speed (Mbit/s) (Ookla) ^[1]
1	Romania	245.50
2	South Korea	241.58
3	Hong Kong	240.8
4	Monaco	220.35
5	Singapore	0.2 Gbps 214.33
6	Denmark	210.71
7	Thailand	206.81
8	Hungary	193.80
9	Liechtenstein	193.79
10	France	192.25

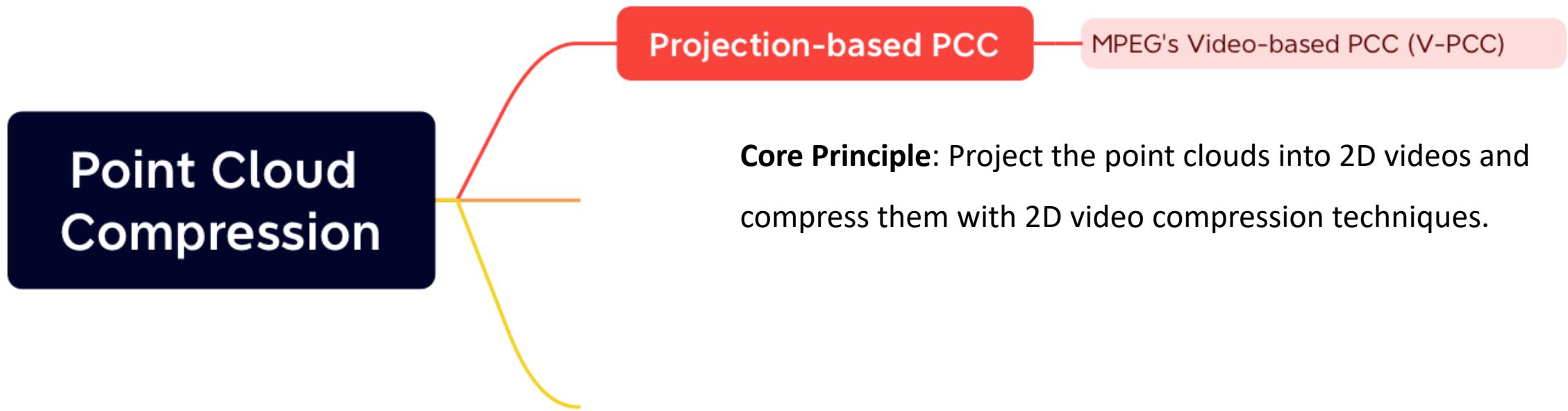
The Ultimate Question

How to stream such large video over the Internet while keeping the video quality high?

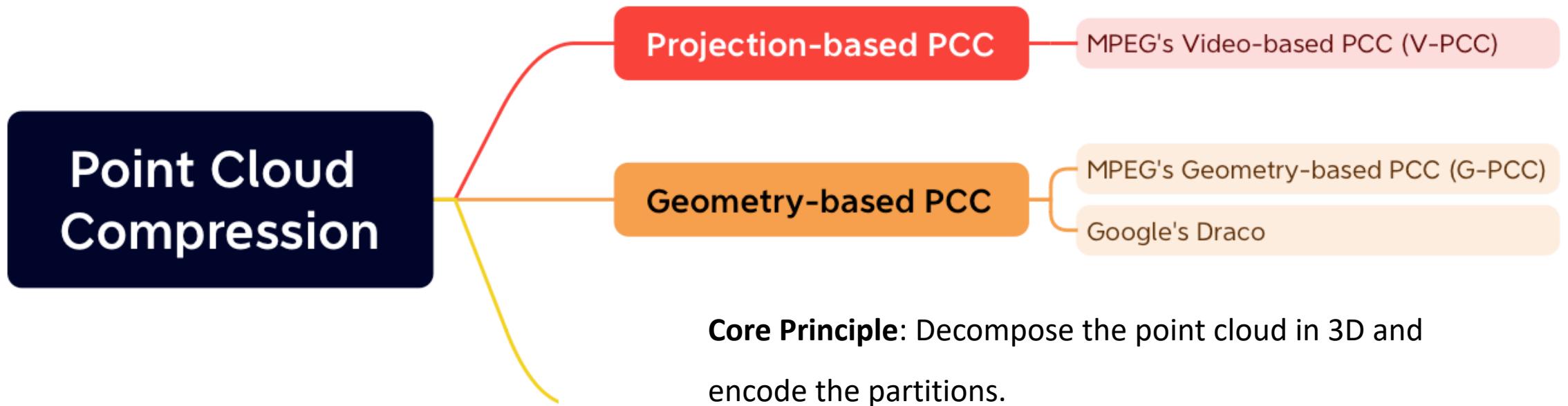
Volumetric Video Streaming System



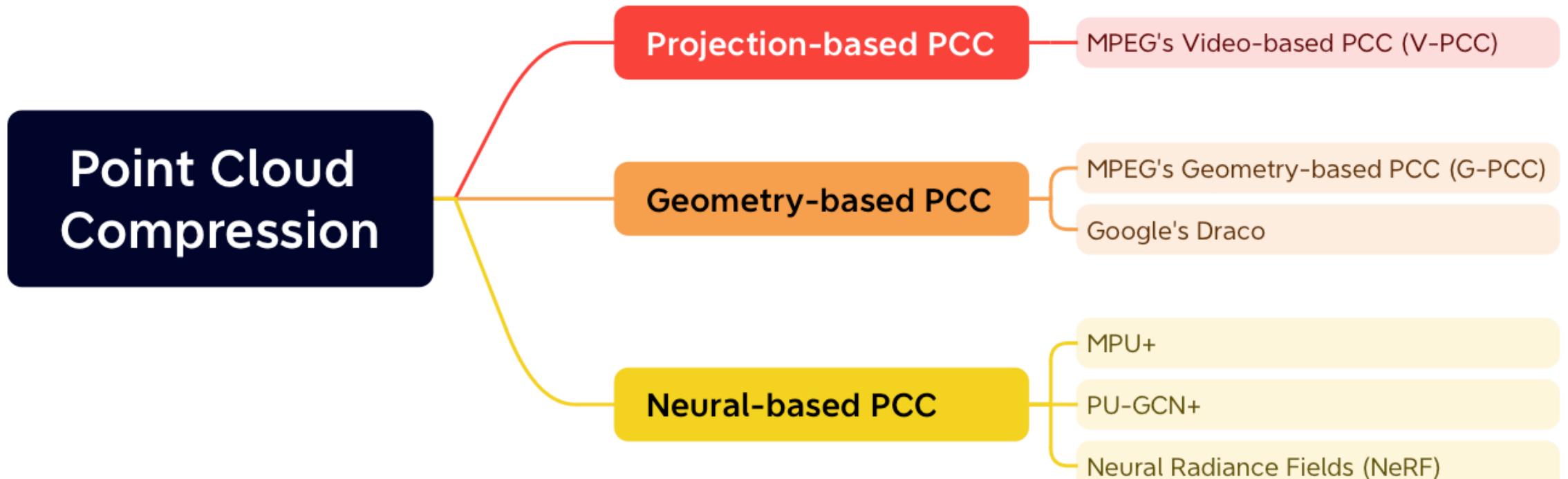
Point Cloud Compression (PCC)



Point Cloud Compression (PCC)



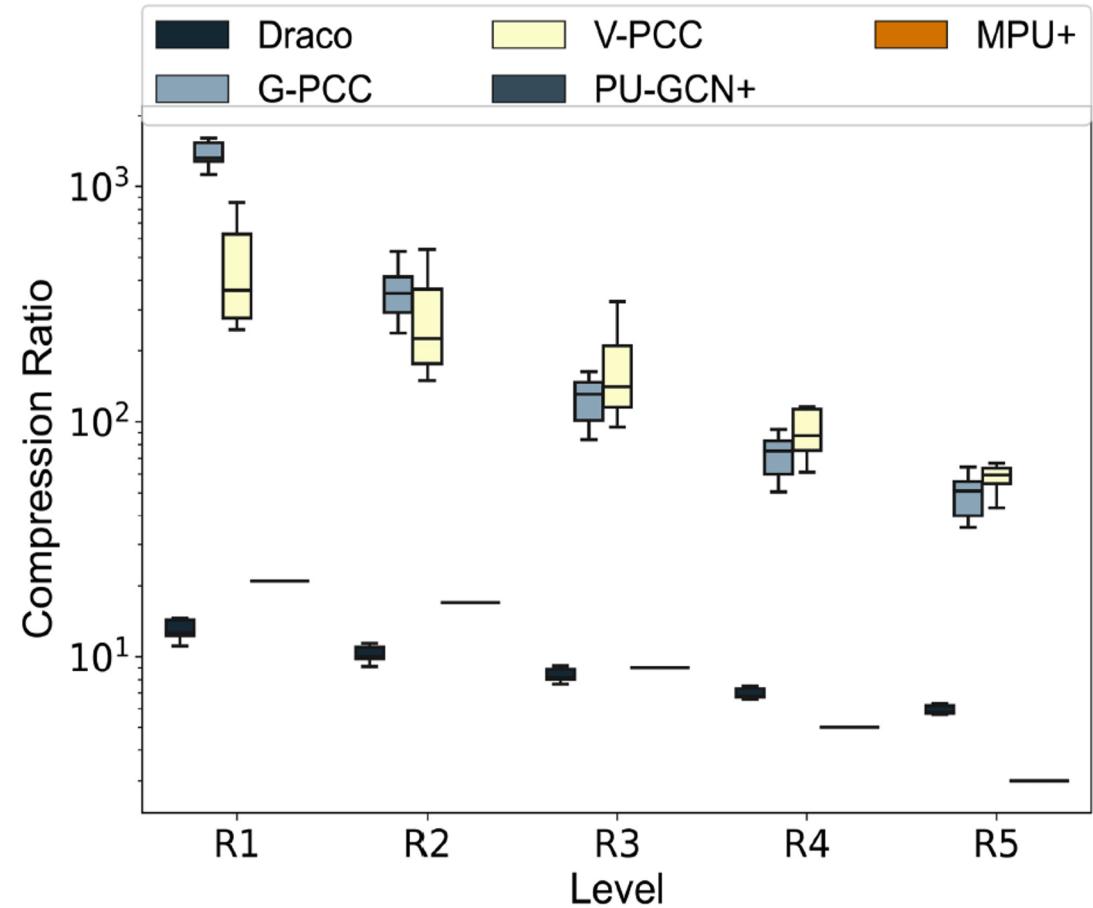
Point Cloud Compression (PCC)



Core Principle: Use neural networks to learn the geometry and attribute of point clouds.

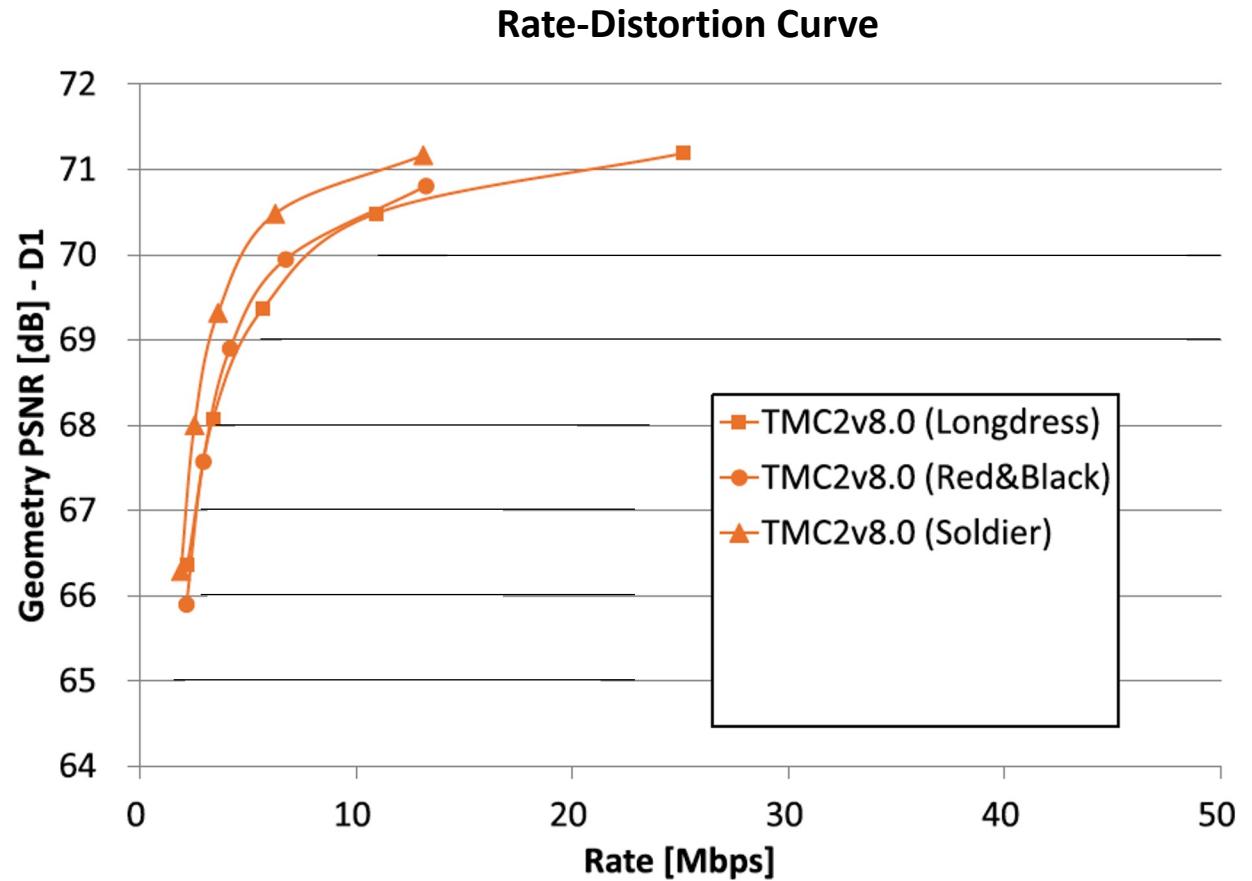
Point Cloud Compression (PCC)

- Video-based PCC (V-PCC),
- Geometry-based PCC (G-PCC),
- Google's Draco,
- MPU+
- PU-GCN+



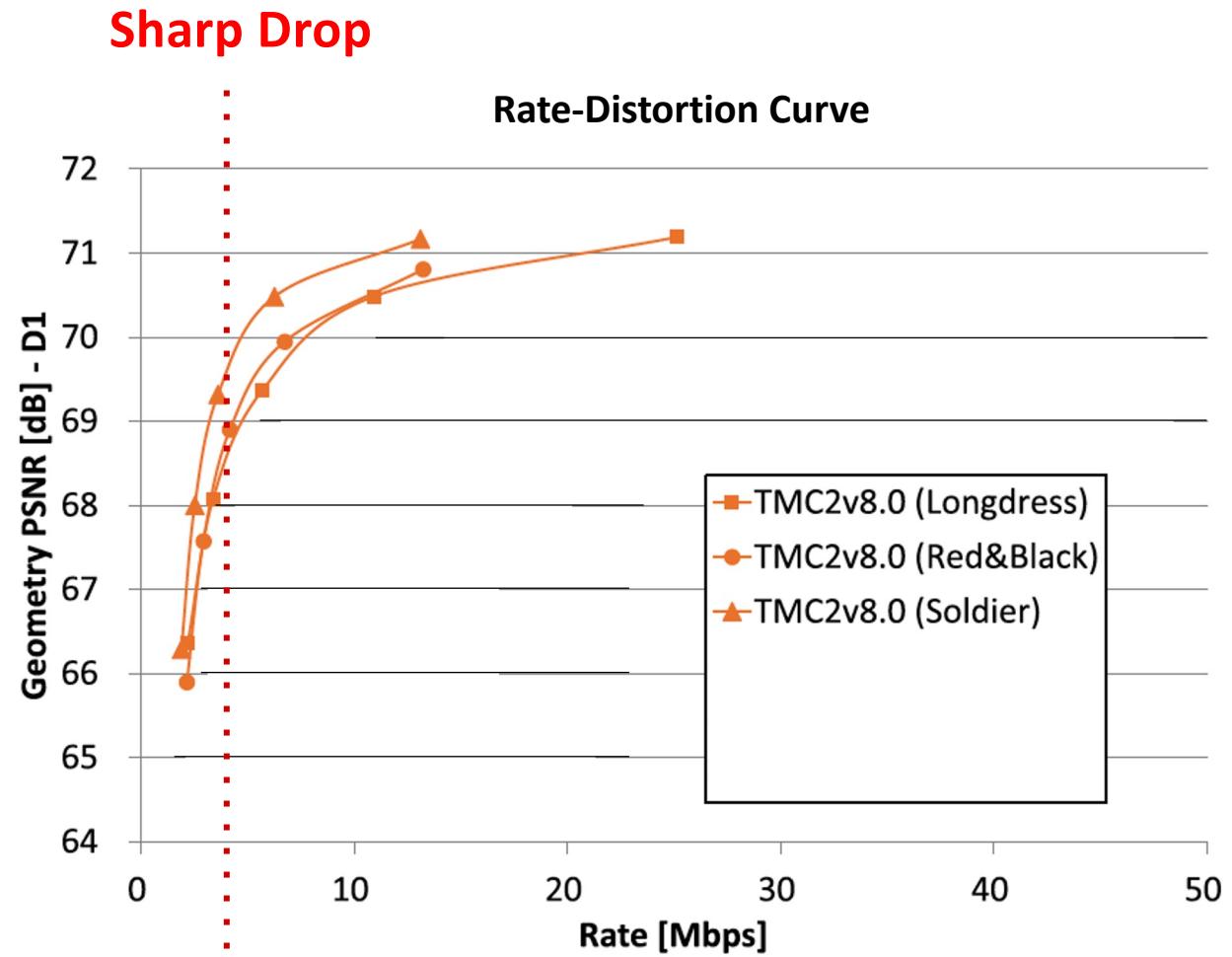
Problem of V-PCC

- **Highly dependent on 2D video codec.**
- The video codecs have a **bit-rate range** that they are designed to operate in



Problem of V-PCC

- Highly dependent on 2D video codec.
- The video codecs have a **bit-rate range** that they are designed to operate in



bit-rate = 2.8 Mbps



bit-rate = 1.8 Mbps



Enabling Low Bit-Rate MPEG V-PCC-encoded Volumetric Video Streaming with 3D Sub-sampling

Yuang Shi, Pranav Venkatram, Yifan Ding, Wei Tsang Ooi

The 14th Conference on ACM Multimedia Systems.

Our Contribution

V-PCC

same bitrate
better quality

**Our
method**



- We can **improve the quality** of V-PCC encoded point clouds **at low bit-rate.**

Intuitions

Intuitions

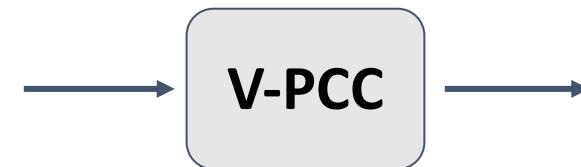
- Overall, we can obtain **better rate-distortion trade-offs** on the rendered images of the point clouds

Down-sample

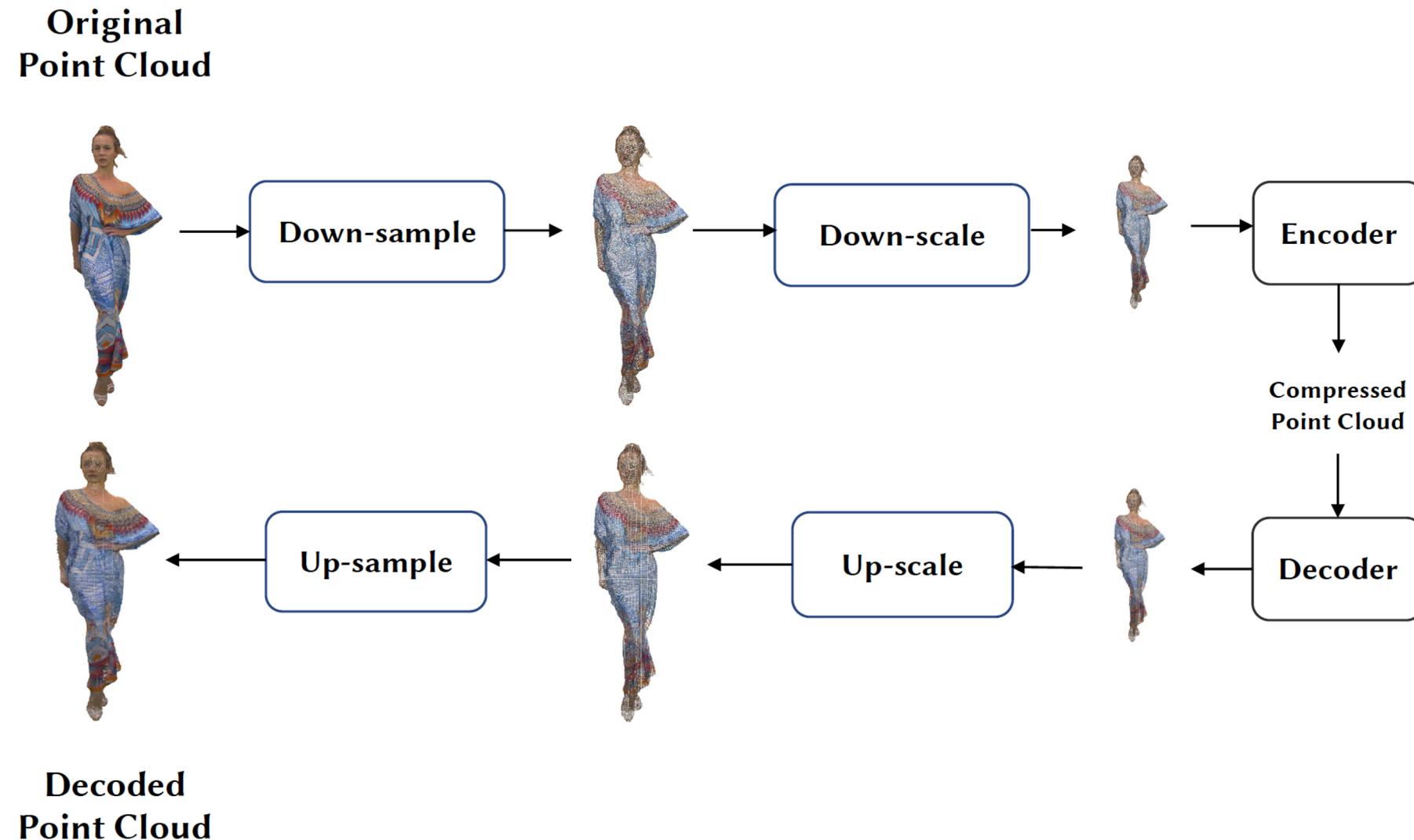


Up-sample

save bit-rate



Pipeline of Our Method



4.6 Mbps
V-PCC



$S = 0.7$
Our Method



2.2 Mbps
V-PCC



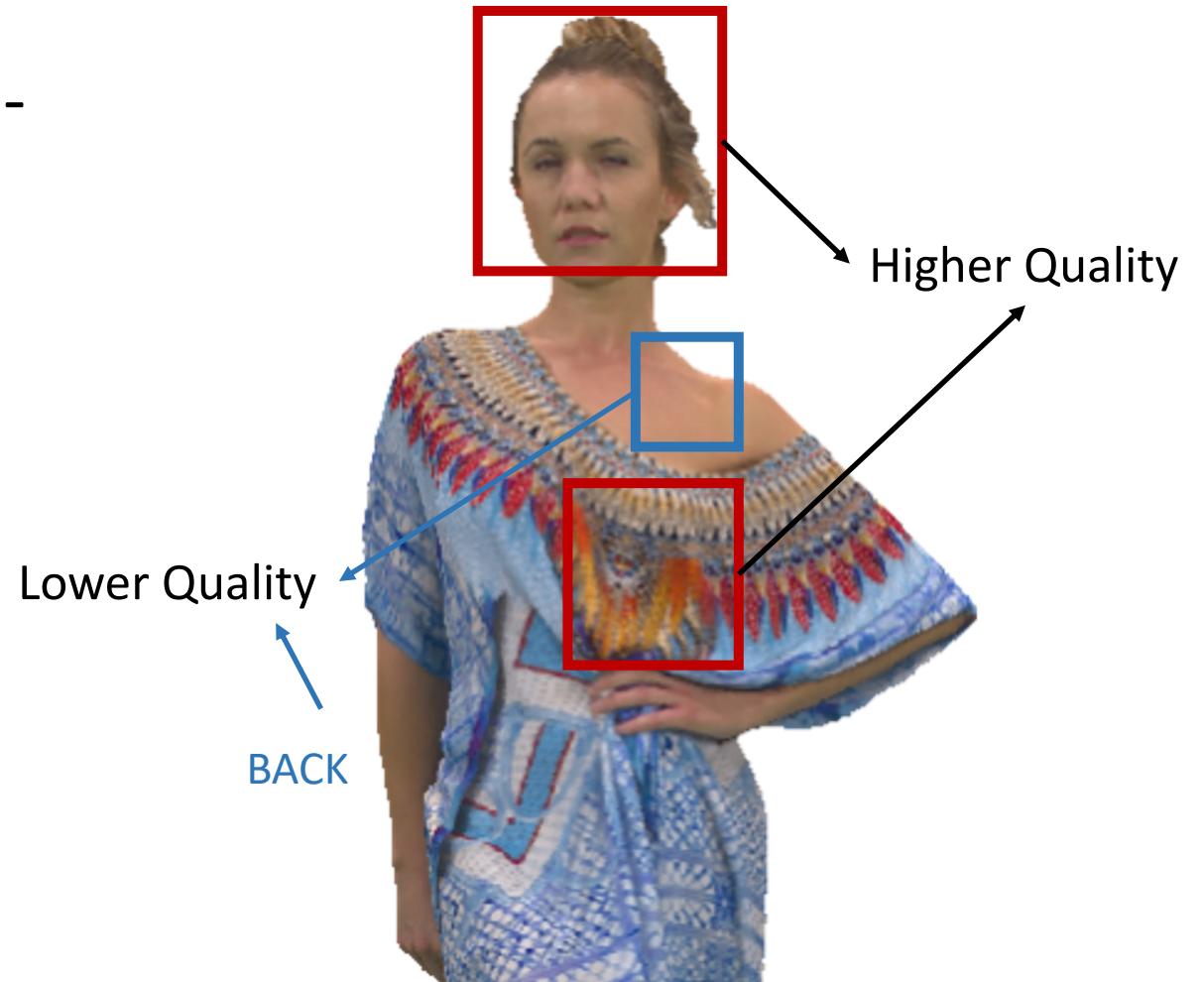
$S = 0.7$
Our Method



At same bitrate: 33.3% improvement in visual quality
At same quality: 48.5% saving in bandwidth

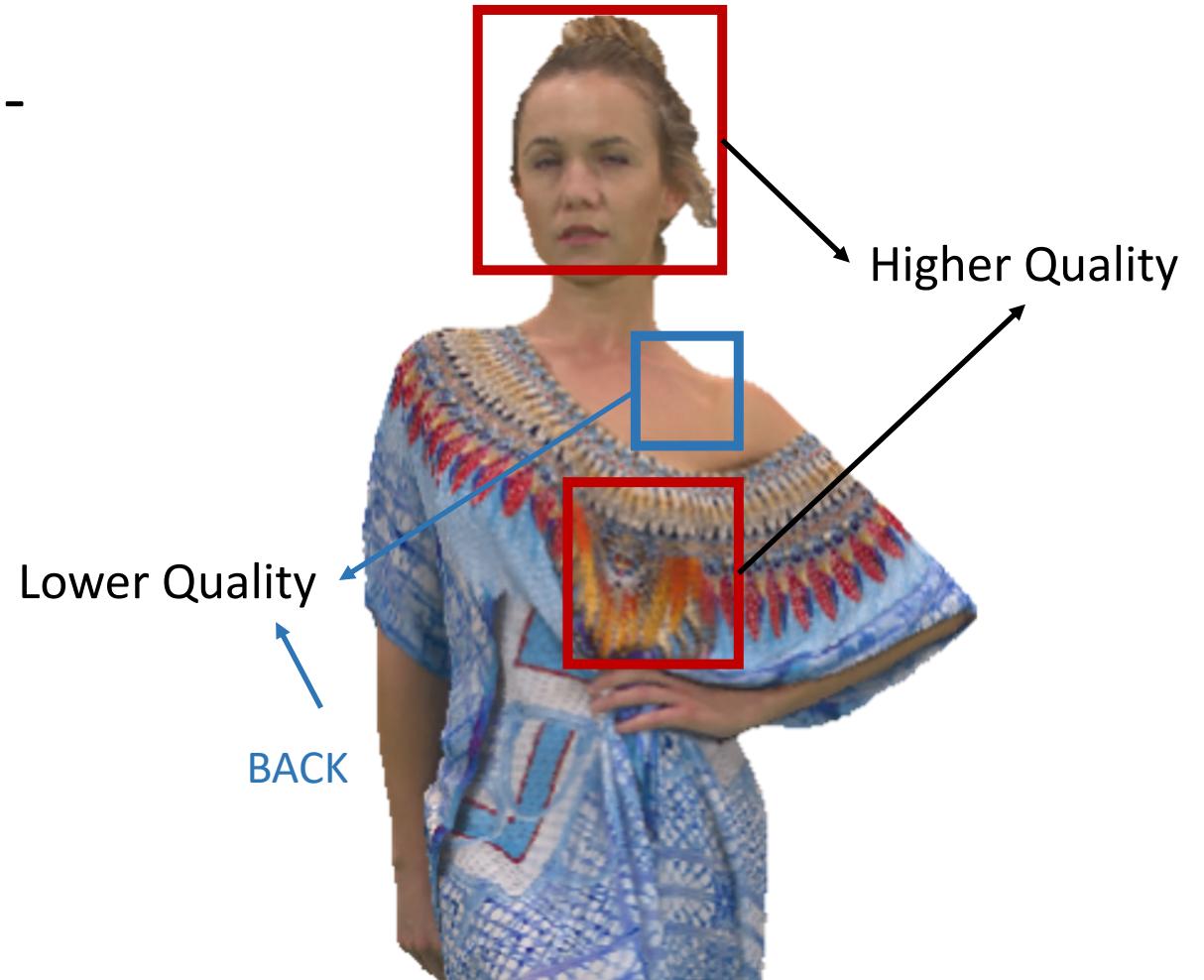
Limitation (Ongoing Works)

- **Ongoing Work 1:** Adaptive down-and up-sampling.



Limitation (Ongoing Works)

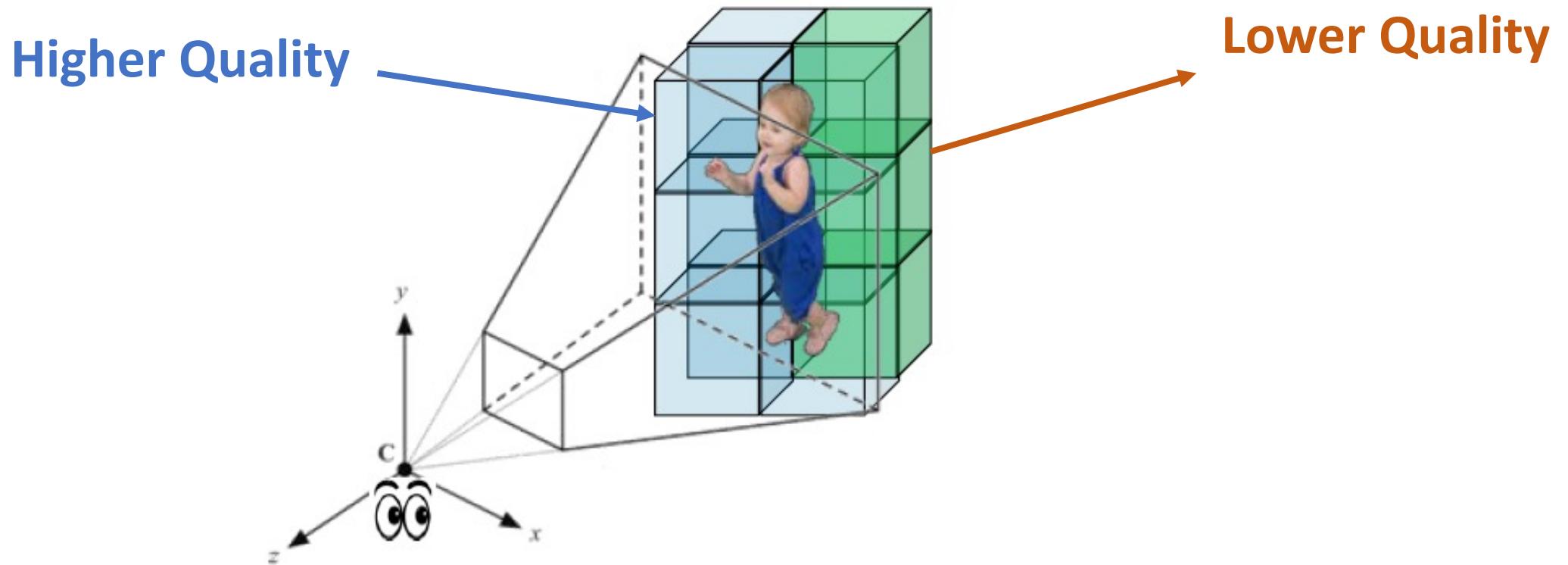
- **Ongoing Work 1:** Adaptive down-and up-sampling.
- **Ongoing Work 2:** Semantic Segmentation.



The Ultimate Question

How to stream such large video over the Internet while keeping the video quality high?

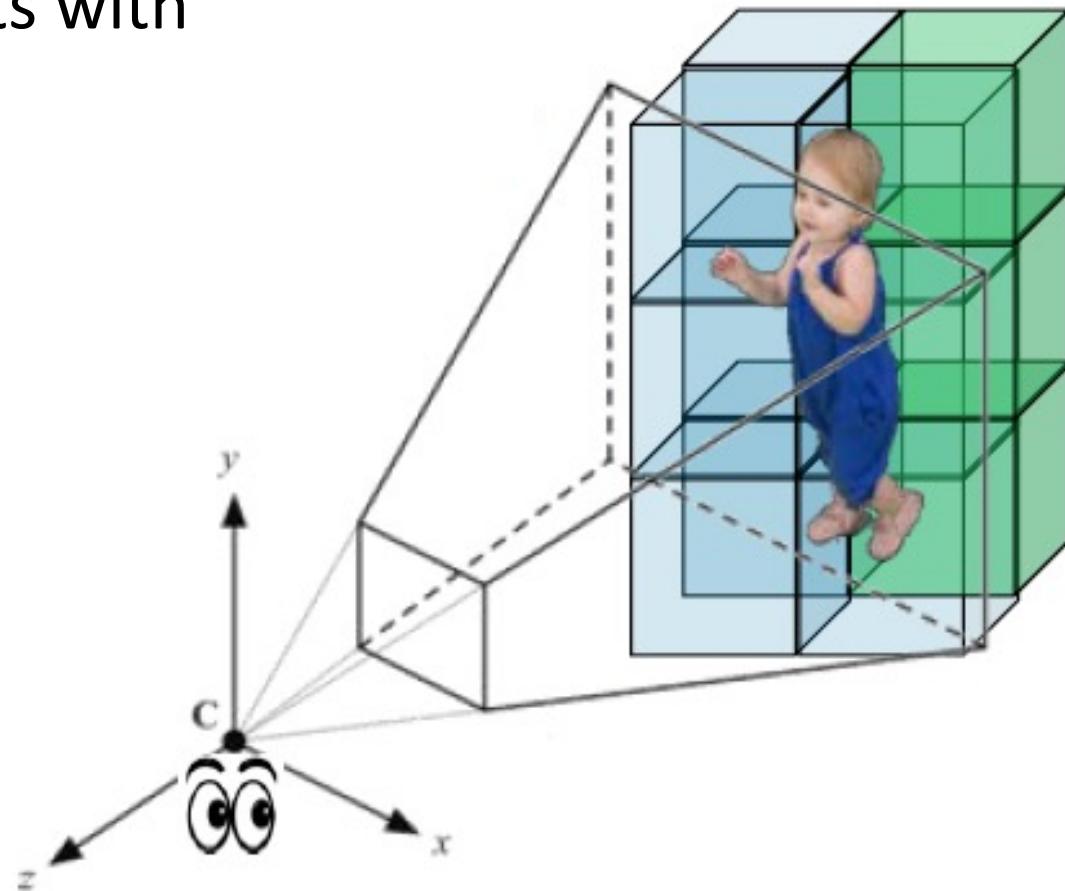
Viewpoint-Aware Optimization



Questions

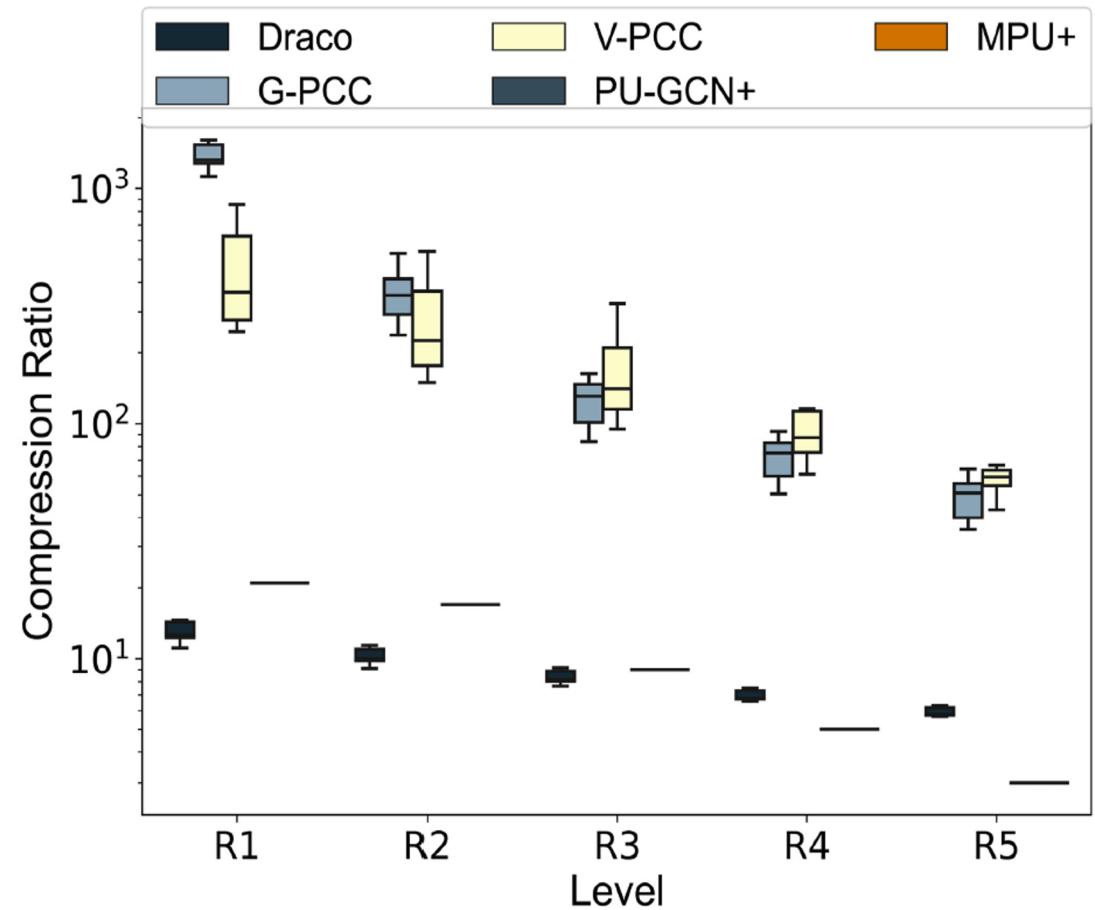
- How to compress the segments with V-PCC?

The segments are small and sparse



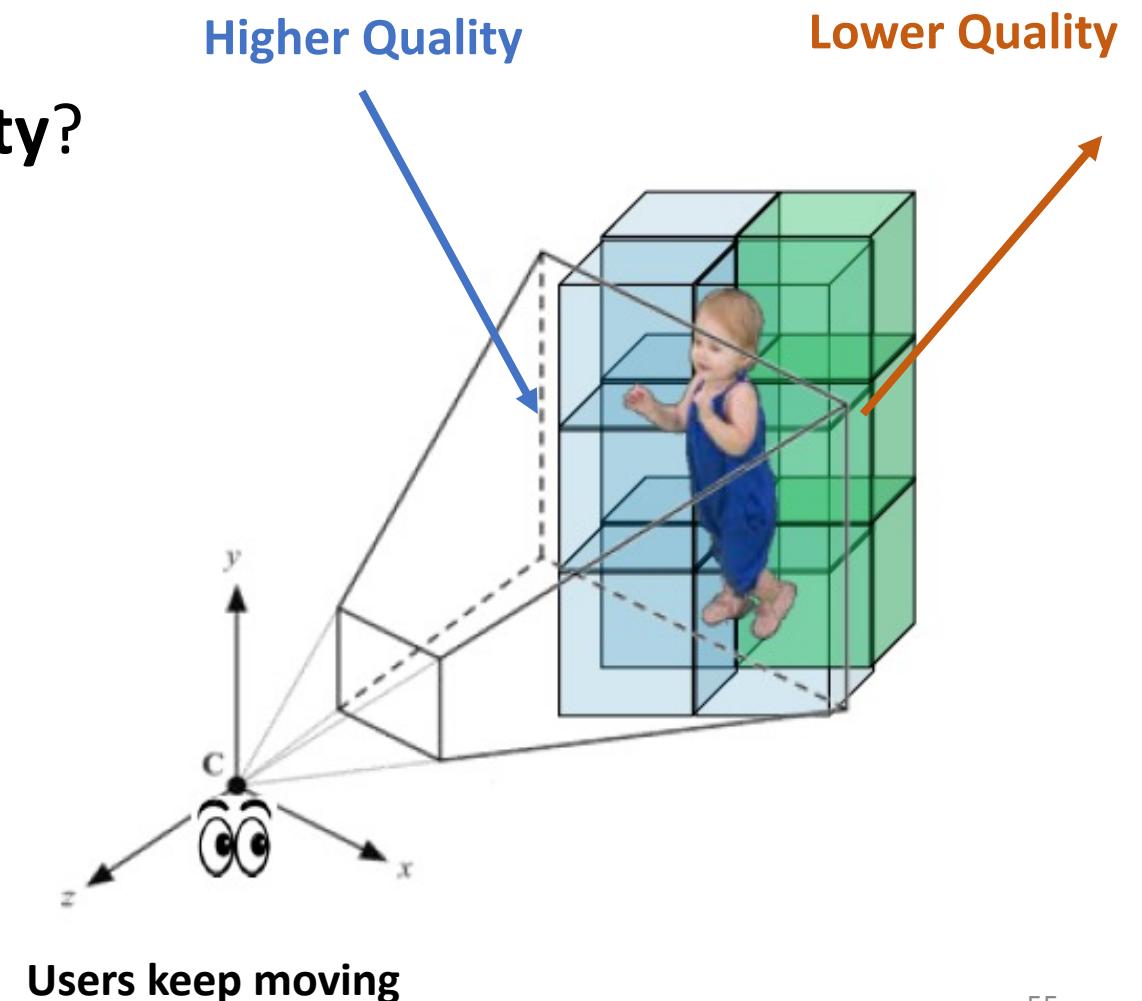
Point Cloud Compression (PCC)

- Video-based PCC (V-PCC)
 - However, V-PCC is terribly bad on sparse and small point cloud.



Questions

- How to define and measure **Quality**?

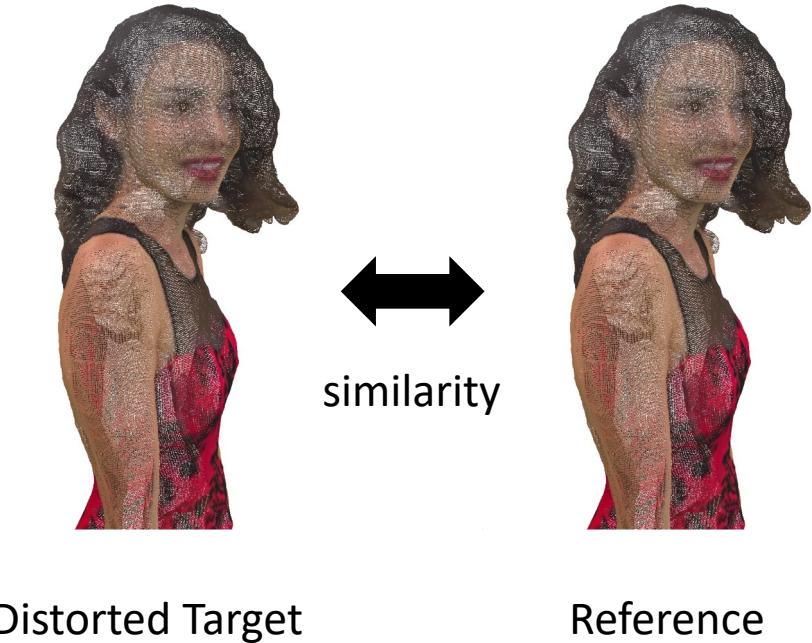


Quality Assessment



Objective Quality Assessment

- Reference based
 - **Core principle:** compare the similarity between the distorted target and its reference.

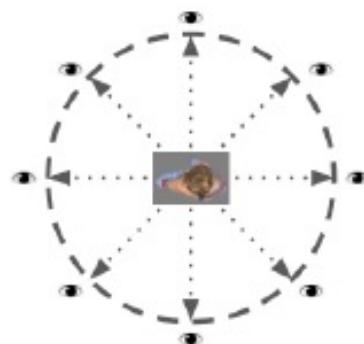
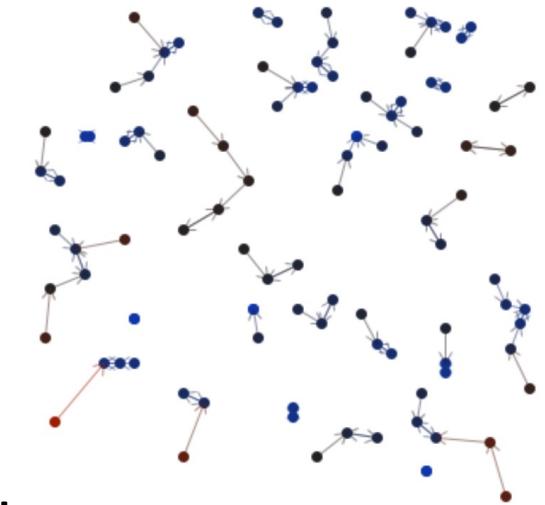


Objective Quality Assessment

- Reference based
 - 3D assessment
 - 2D assessment



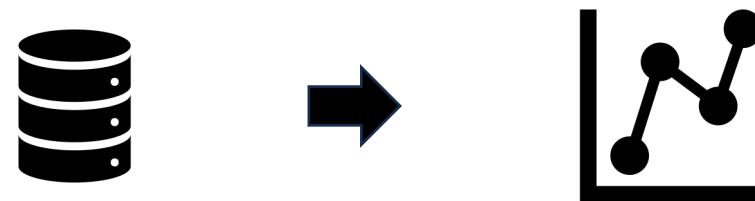
3D Assessment



2D Assessment

Objective Quality Assessment

- Non-reference based
 - **Core principle:** predict the quality of point cloud based on prior knowledge.

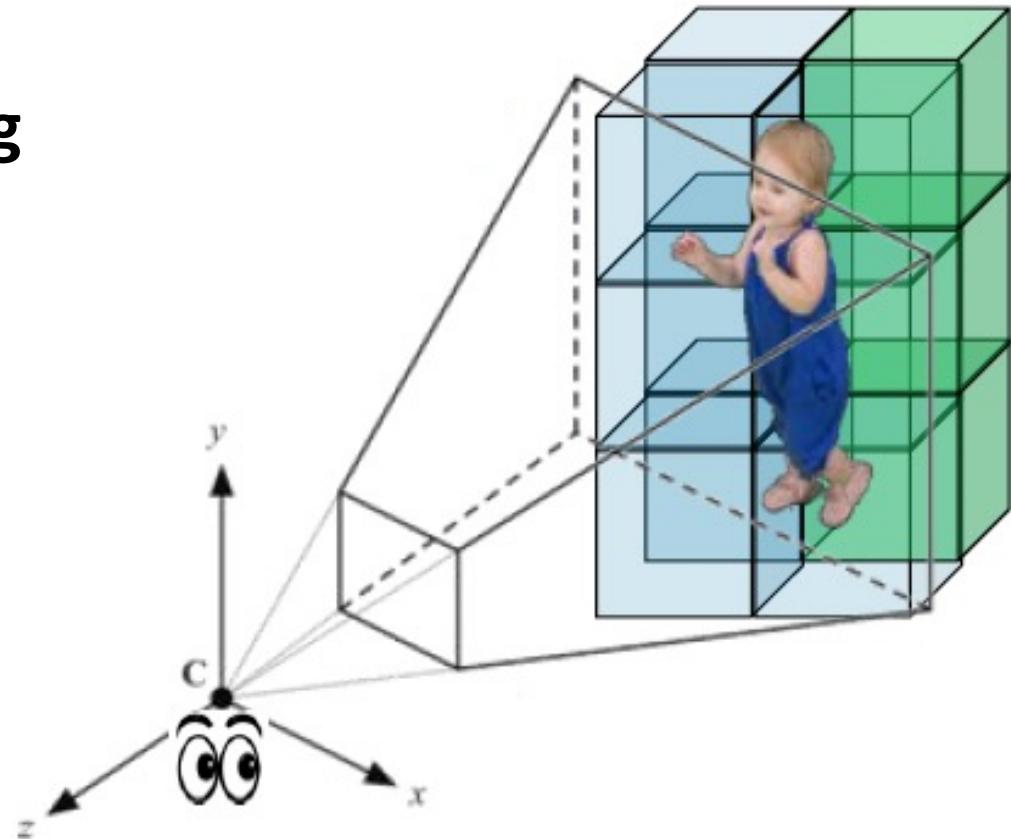


Quality Data / Rules
(prior knowledge)

Prediction Model

Problem

- Users keep **moving** and **interacting** with the video.



Problem

- Users keep moving and interacting with the video.
- Point cloud is different than pixels on a **regular grid**.

1 pixel shift



Distorted

Reference

Subjective Quality Assessment

- Interactive user study
- Non-interactive user study



Video Source: https://www.youtube.com/watch?v=aO3TAke7_MI

Subjective Quality Assessment

- Interactive user study
- Non-interactive user study



Image Source: Taken by myself

Subjective Quality Assessment

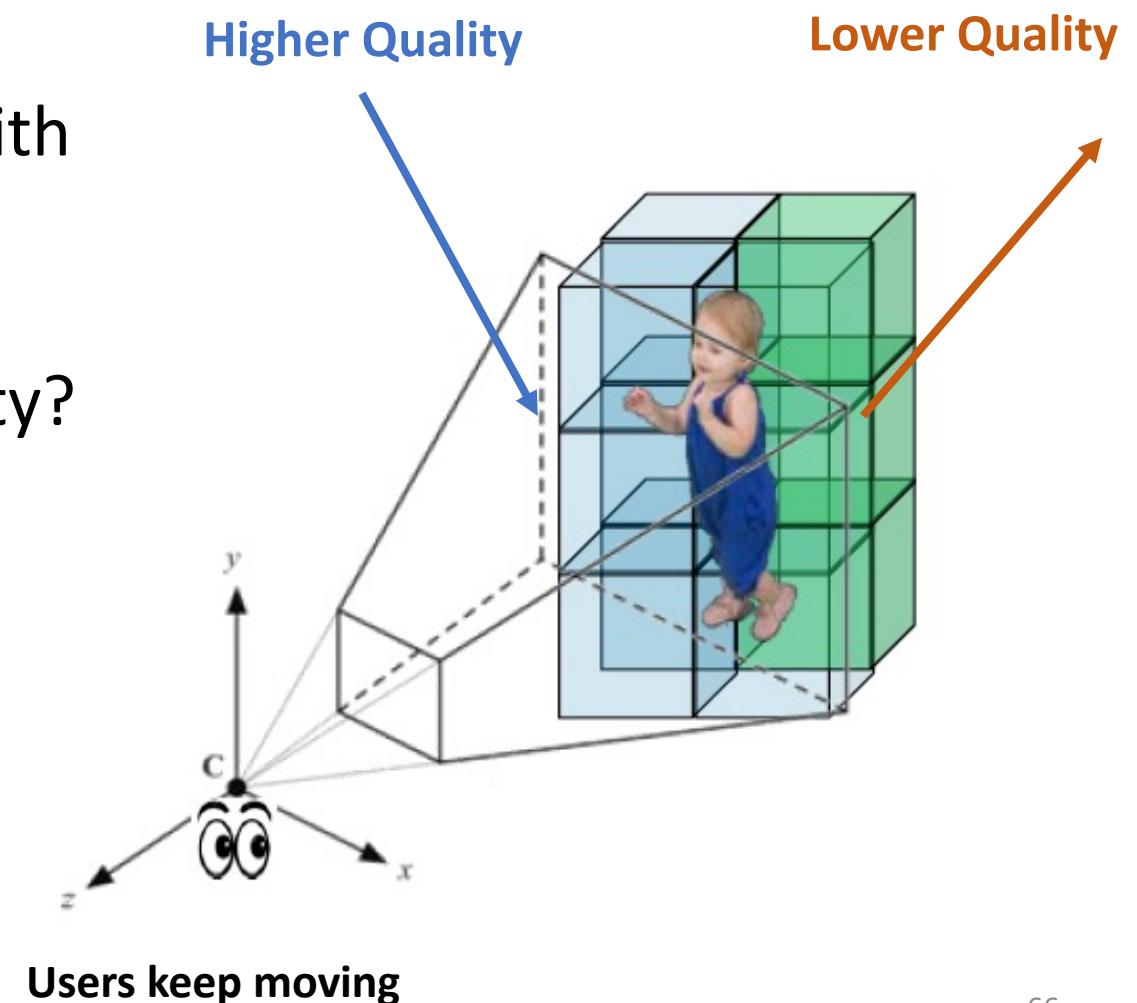
- Interactive user study
 - **Problem:** different interaction patterns.
- Non-interactive user study
 - **Problem:** ignore the nature of 3D contents.

Non-interactive User Study

- Existing Works
 - Focus on **static 3D object**.
 - Explore the effect of **impairments** on visual quality.
 - Many other factors.

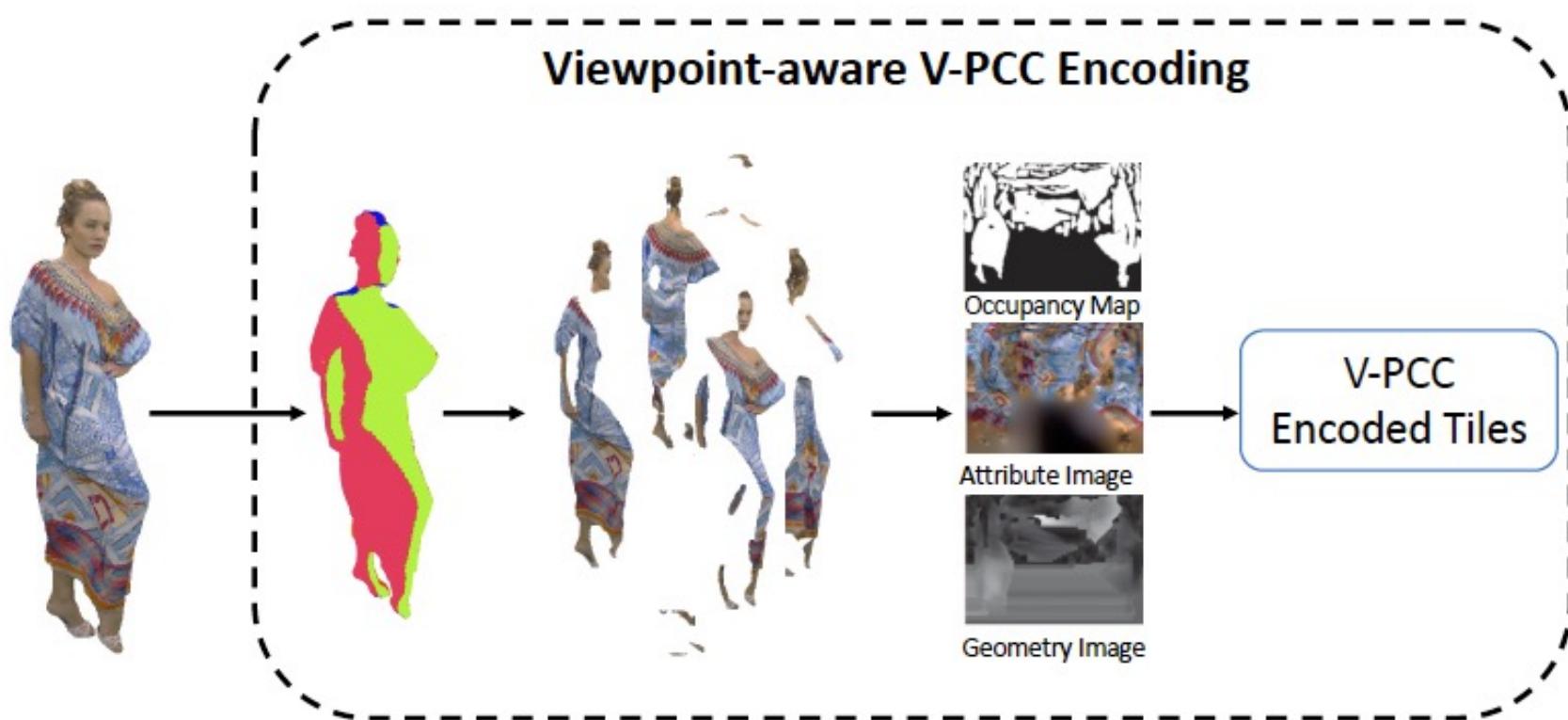
Questions

- How to compress the segments with V-PCC?
- How to define and measure Quality?



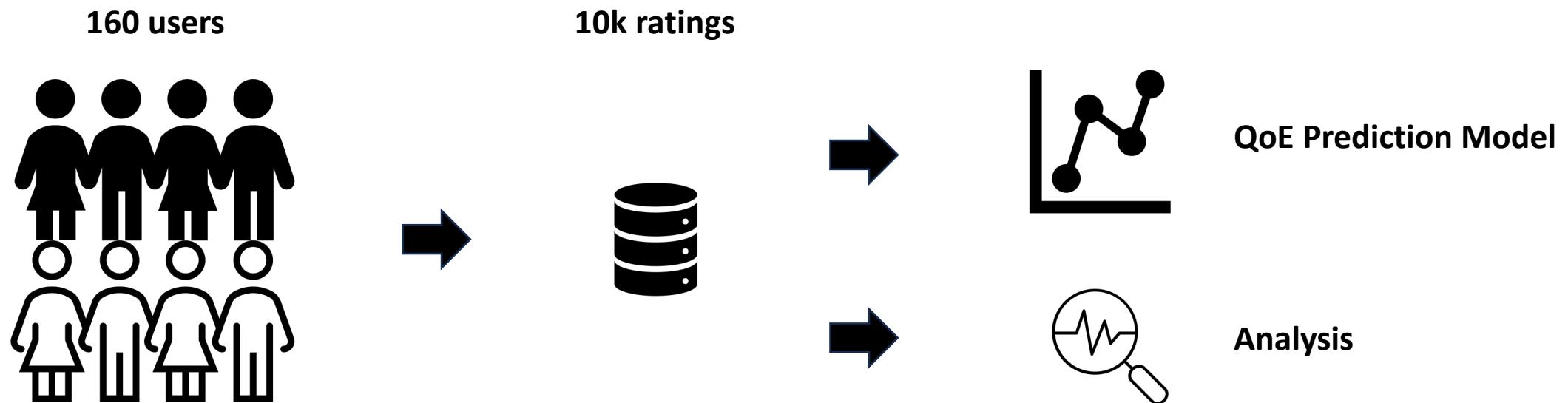
Our Contribution:

- Point cloud **segmentation technique** for V-PCC.



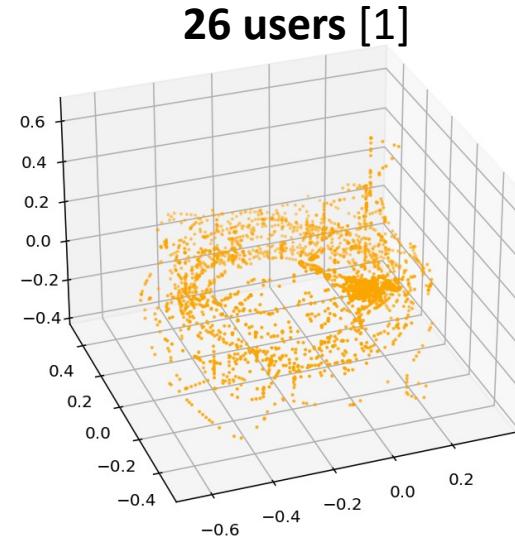
Our Contribution:

- Model to predict **user's quality of experience** (QoE).
- Detailed analysis of the impact on QoE.

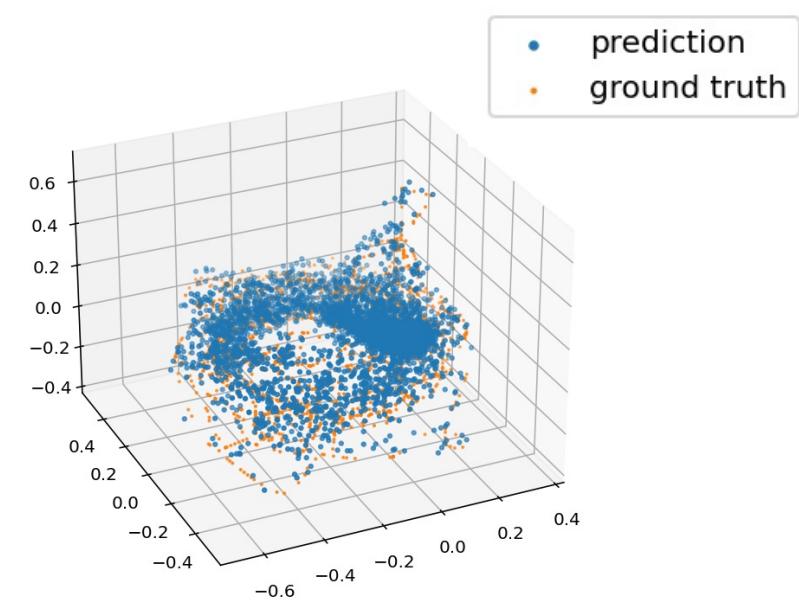
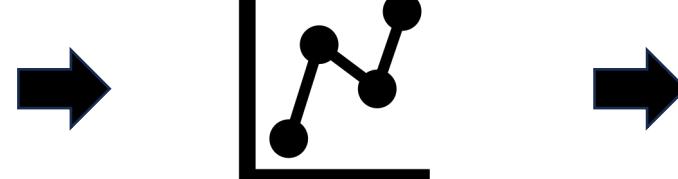


Our Contribution:

- Point cloud segmentation technique for V-PCC.
- Model to predict user's quality of experience (QoE).
- Investigation of **user's viewing behavior pattern and prediction**.

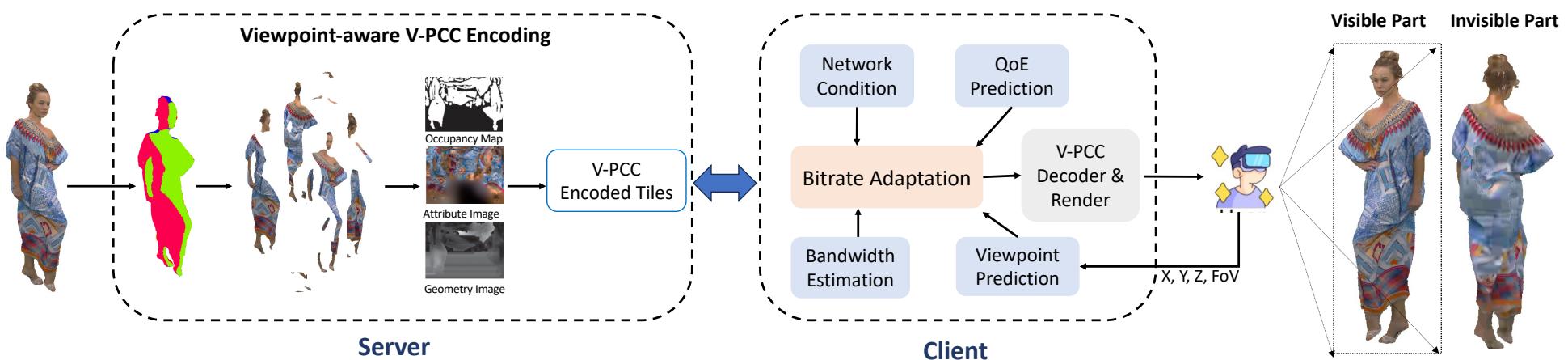


Viewing Behavior Model



Our Contribution: QV4

- Point cloud segmentation technique for V-PCC.
- Model to predict user's quality of experience (QoE).
- Investigation of user's viewing behavior pattern and prediction.
- **QV4:** a QoE-based streaming pipeline for Viewpoint-aware V-PCC-encoded Volumetric Video.



Our Contribution: QV4

- Point cloud segmentation technique for V-PCC.
- Model to predict user's quality of experience (QoE).
- Investigation of user's viewing behavior pattern and prediction.
- **QV4:** a QoE-based streaming pipeline for Viewpoint-aware V-PCC-encoded Volumetric Video.

Key Results:

Achieve up to 15% improvement on visual quality
with no drop in compression ratio of V-PCC

V-PCC Standard

VMAF: 55.753
Bandwidth: 7 Mbps



Our proposed QV4

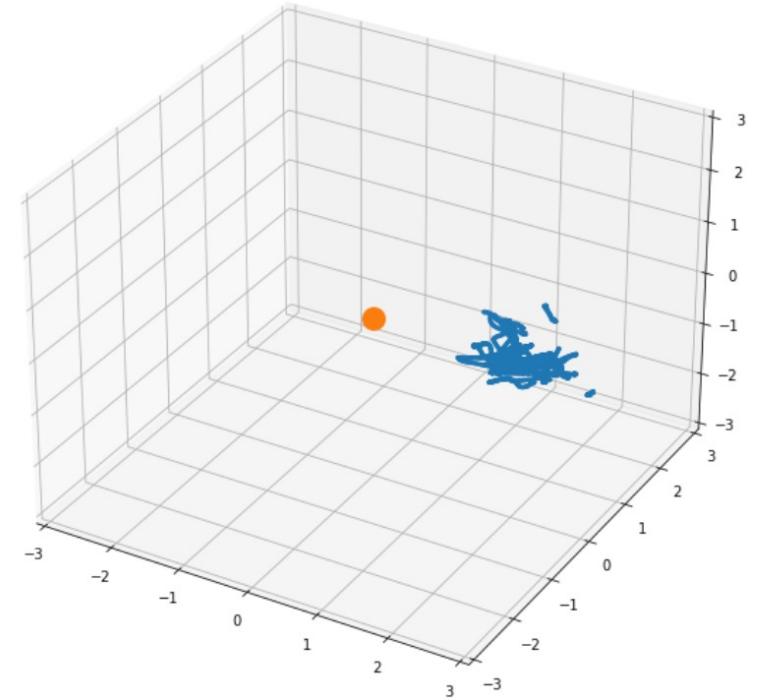
VMAF: 71.881
Bandwidth: 7 Mbps



Demo

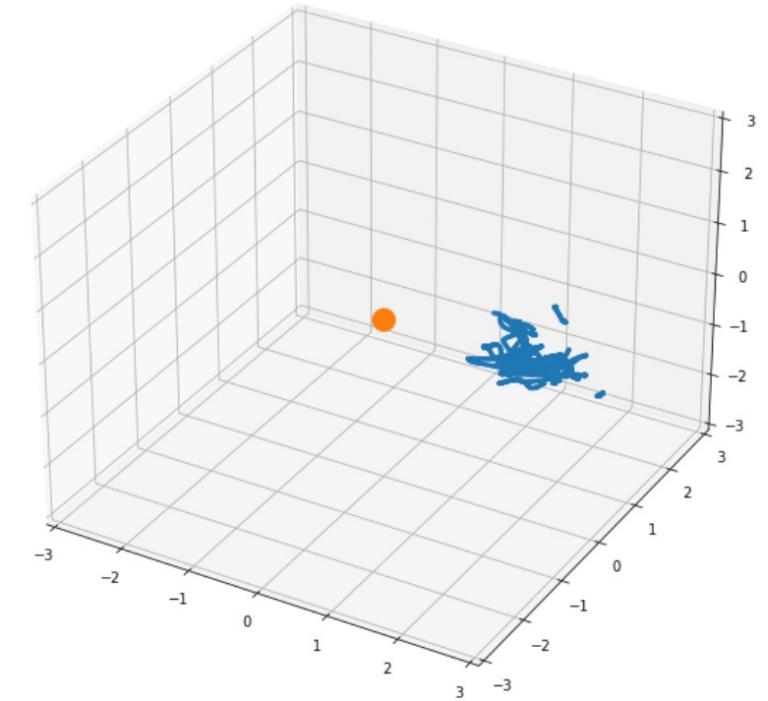
Side Product

- Looked toward the **frontal body** all the time.
- Looked at **face** always.



Side Product

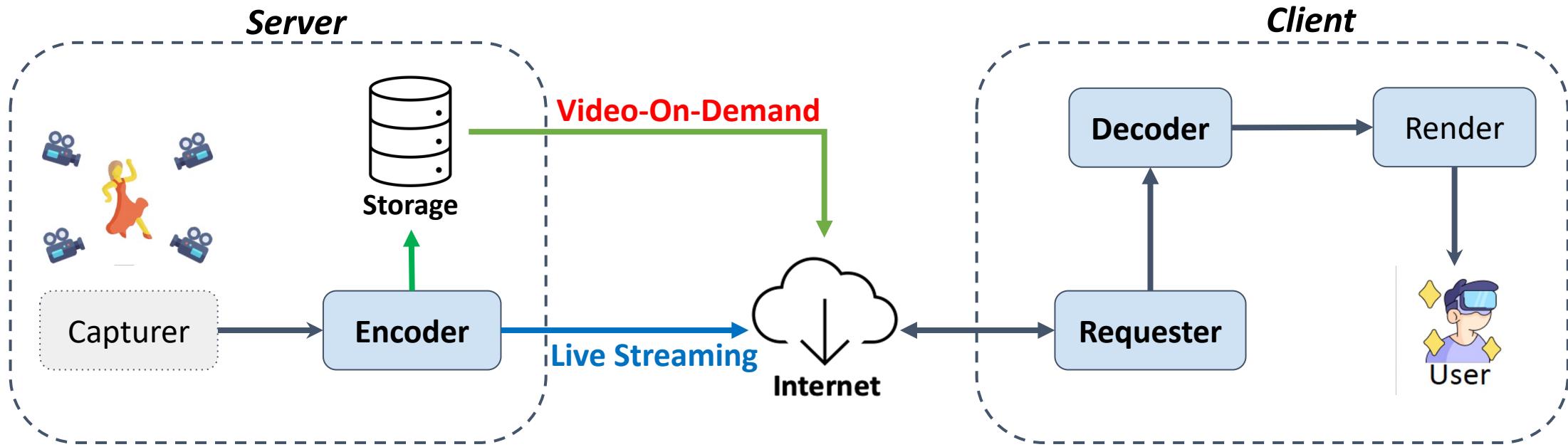
- Looked toward the **frontal body** all the time.
- Looked at **face** always.
- **Question:** What is the effect of face quality on user's viewing experience?





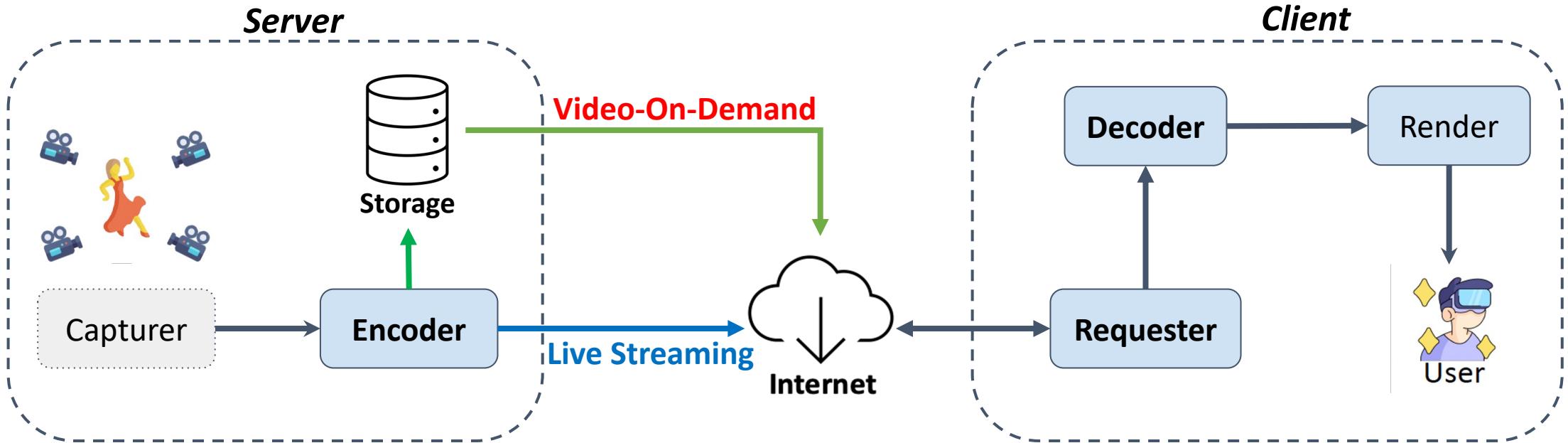
Viewing experience is
improved by up **to 15%**.

Summary – Our Contributions



- [1] **Shi, Yuang**, et al. "Enabling Low Bit-Rate MPEG V-PCC-encoded Volumetric Video Streaming with 3D Sub-sampling." *Proceedings of the 14th Conference on ACM Multimedia Systems*. 2023.
- [2] **Shi, Yuang**, et al. "QV4: QoE-based Viewpoint-Aware V-PCC-encoded Volumetric Video Streaming." *Under submission*.
- [3] **Shi, Yuang**, et al. "MPEG V-PCC Volumetric Video Quality Assessment and Modelling." *Under writing*.
- [4] Sun, Yuan-Chun, Huang, I-Chun, **Shi, Yuang**, et al. "A Dynamic 3D Point Cloud Dataset for Immersive Applications." *Proceedings of the 14th Conference on ACM Multimedia Systems*. 2023.
- [5] Huang, I-Chun, **Shi, Yuang**, et al. "Composing Error Concealment Pipelines for Dynamic 3D Point Cloud Streaming." *Under Submission*.

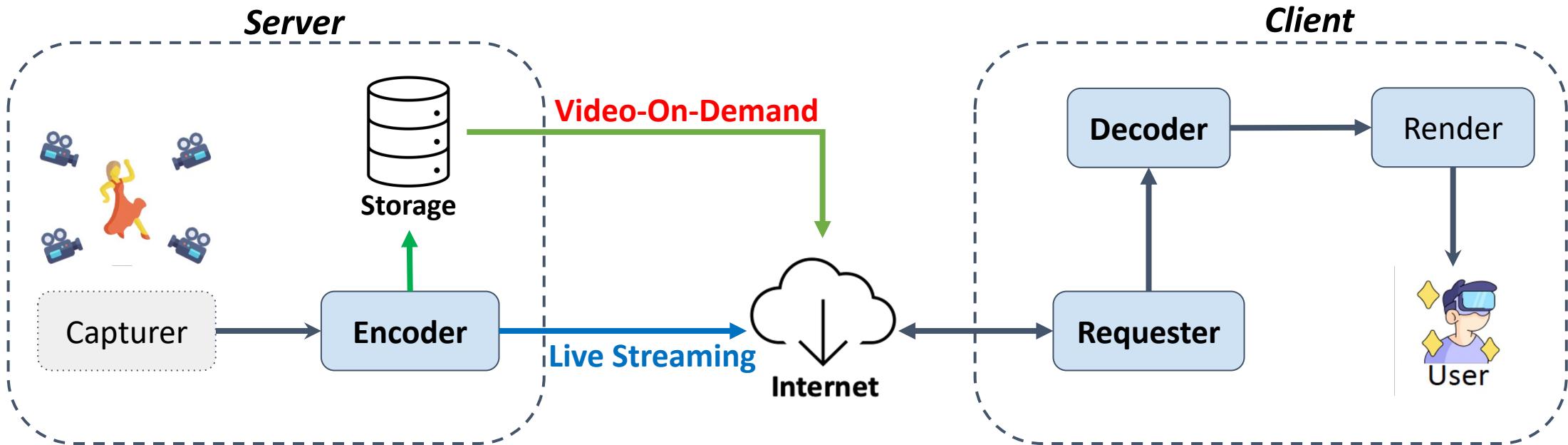
Summary – Ongoing Work



Ongoing Work:

- Semantic segmentation and adaptive down- and up-sampling.
- Neural-based Volumetric Video Streaming.

Summary — Future Work



Ultimate Goal:

- Build a streaming system which combines all my works together.

Thank you!