

Volumetric Video Technologies for Real-Time Immersive Communication

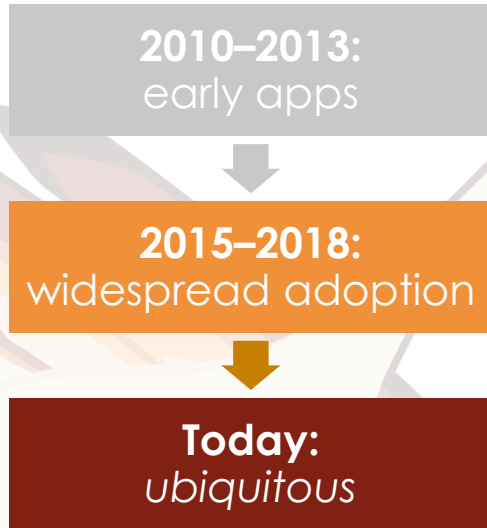


Asst. Prof. Alexandre Mercat
Post-doc. Guillaume Gautier

Apr. 22, 2026
CWI, Amsterdam

Volumetric Video Technologies for Real-Time Immersive Communication

Real-Time Video Communication

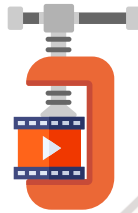


Video Communication Pipeline

Capture



Encoding



Transmission



Decoding



Playback



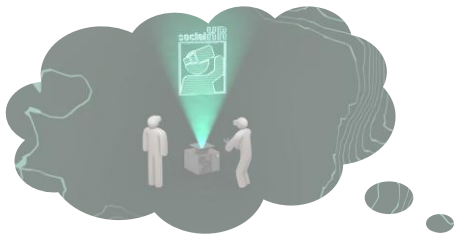
Requirements:



- Bandwidth
- Latency
- Real-time processing



ULTRA VIDEO GROUP



Volumetric Communication Pipeline

Capture

Encoding

Transmission

Decoding

Playback



Requirements:



- Bandwidth
- Latency
- Real-time processing



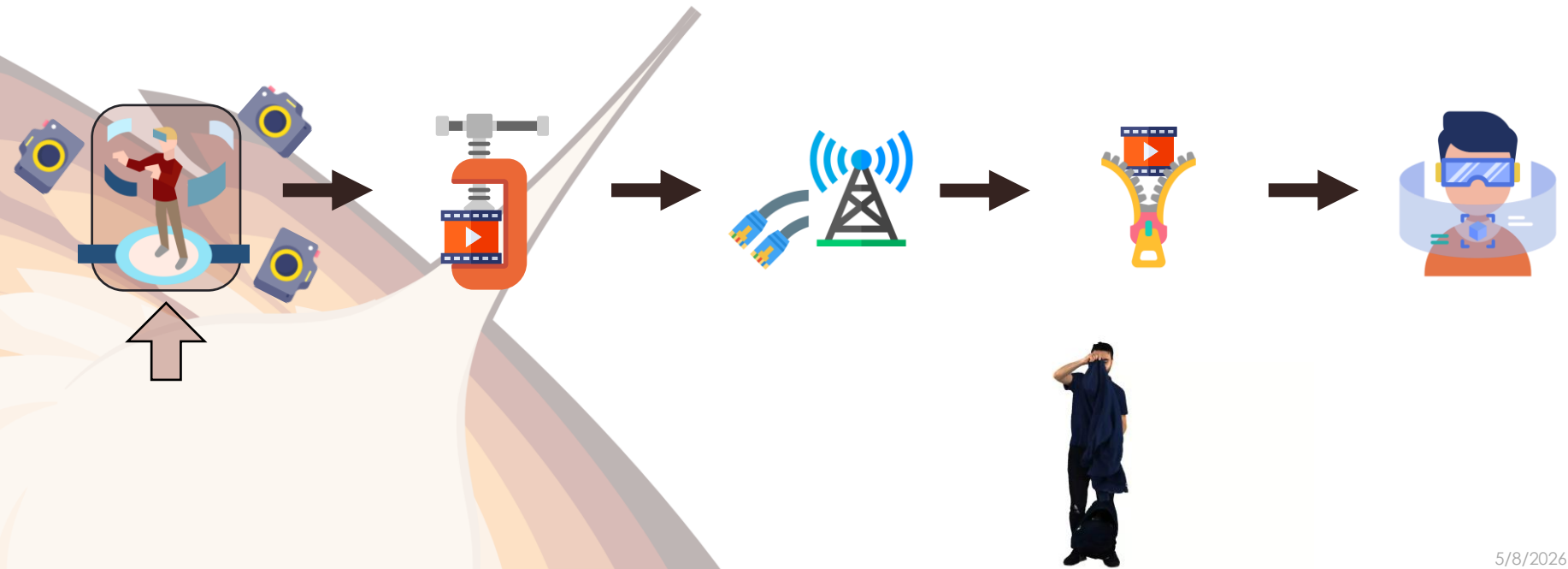
ULTRA VIDEO GROUP

Presentation Outline

1. Volumetric Data Representation
2. Dynamic Point Cloud Capture
3. Point Cloud Compression



Volumetric Data Representation



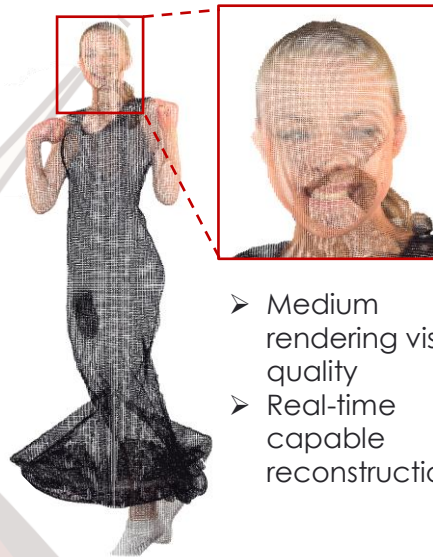
Comparing Volumetric Representations

Meshes



- High rendering visual quality
- Complex reconstruction

Point clouds



- Medium rendering visual quality
- Real-time capable reconstruction

Gaussian splats



Estátua do Leão - Parque Ibirapuera, São Paulo, Brazil
Posted by Lipi - [Link](#)

- Very high rendering visual quality
- No real-time reconstruction (yet)



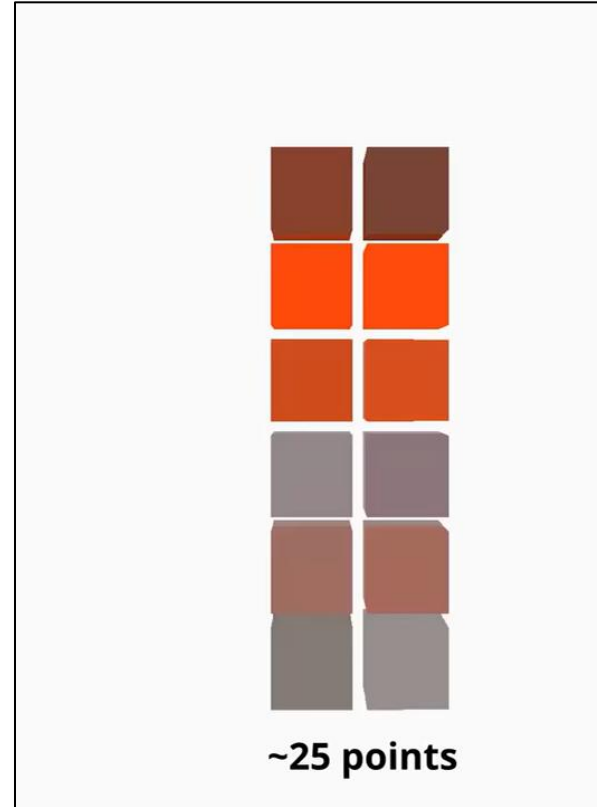
Point Clouds Representation

What is a point cloud

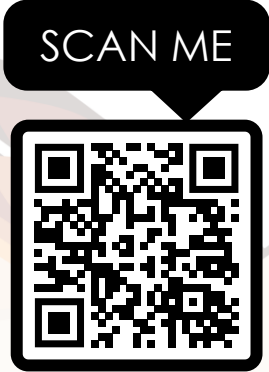
- Discrete set of points in 3D space
- Represents 3D shape or object
- Higher density → more detail

Data

- Geometry: (x, y, z)
- Attributes: color, transparency, normals, ...



Hands-On: Interactive Point Cloud Viewer



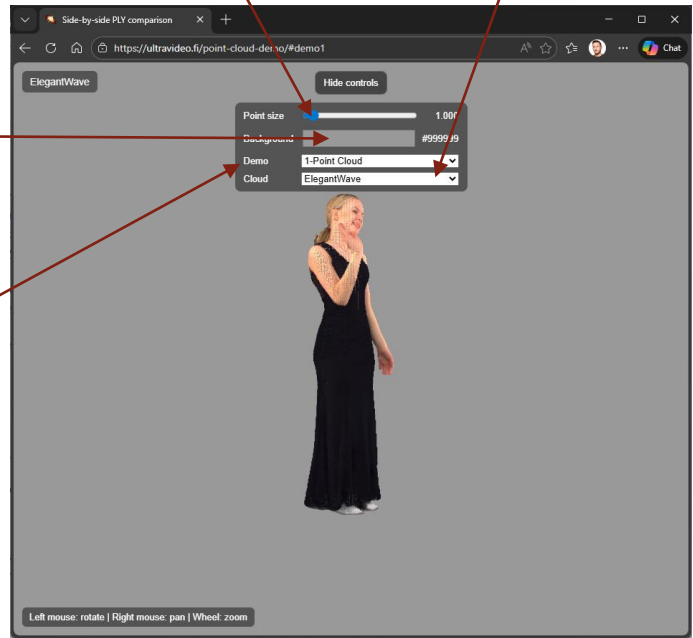
<https://ultravideo.fi/point-cloud-demo/>

Switch from demos

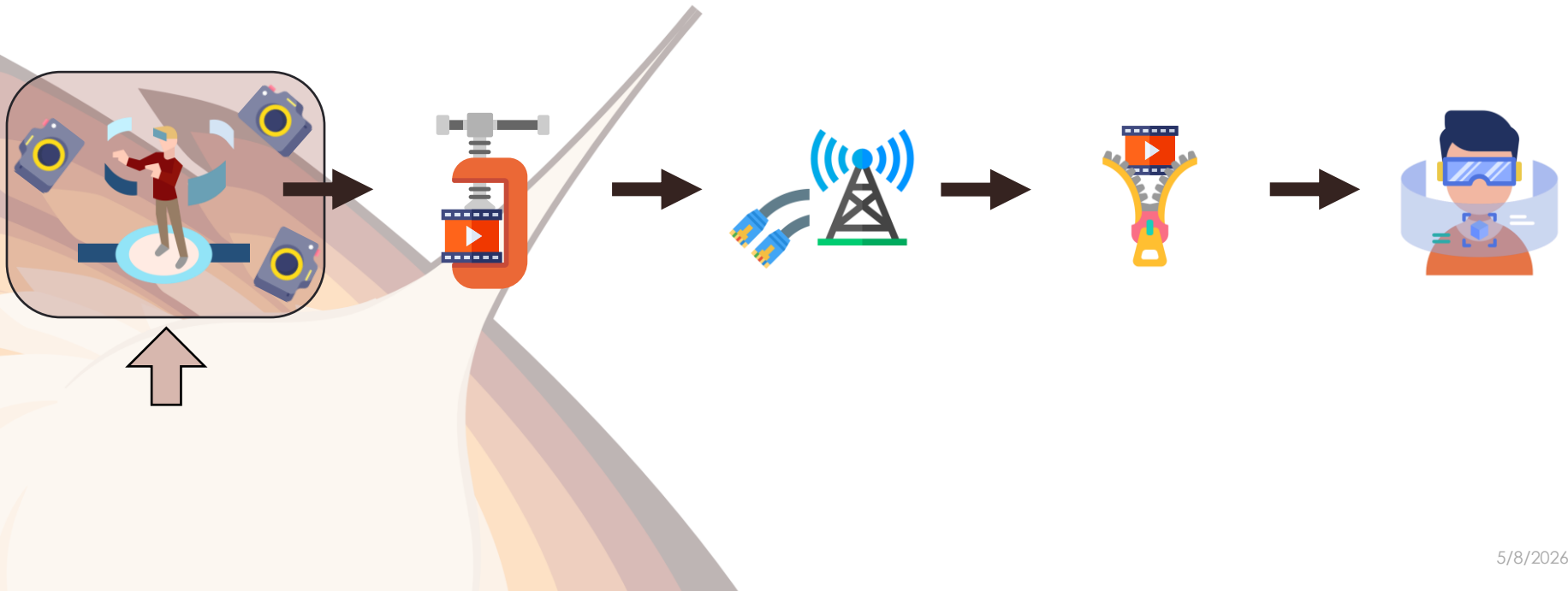
Change background color

Select point size

Change point clouds



Dynamic Point Cloud Capture

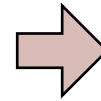


From Images to Point Clouds

Depth map



2D image (RGB)



3D point cloud



How do we get depth?

RGB-D sensor

Passive



RGB Left

RGB Right



Depth

Active



RGB



Depth

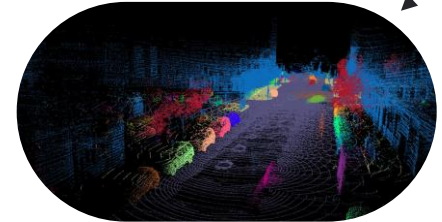


ULTRA VIDEO GROUP

LiDAR sensor



Point cloud



<https://segments.ai/blog/mastering-point-clouds/>

One Camera is Not Enough



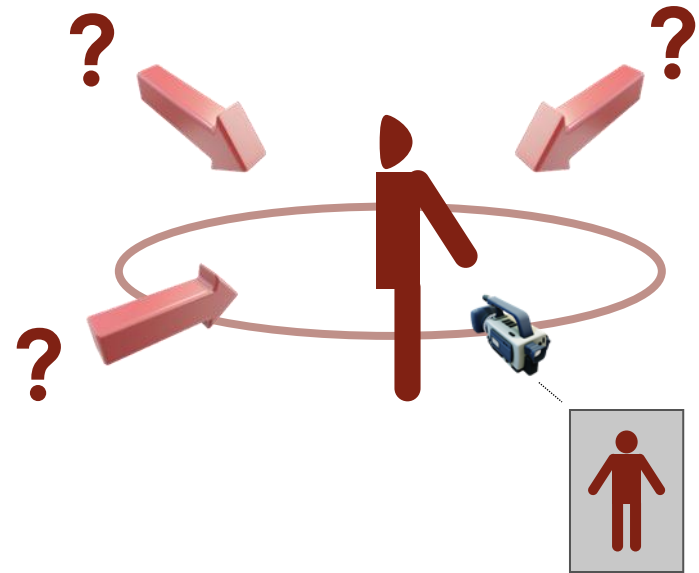
Challenge

- Partial view
- Occlusions
- Depth accuracy



Solution

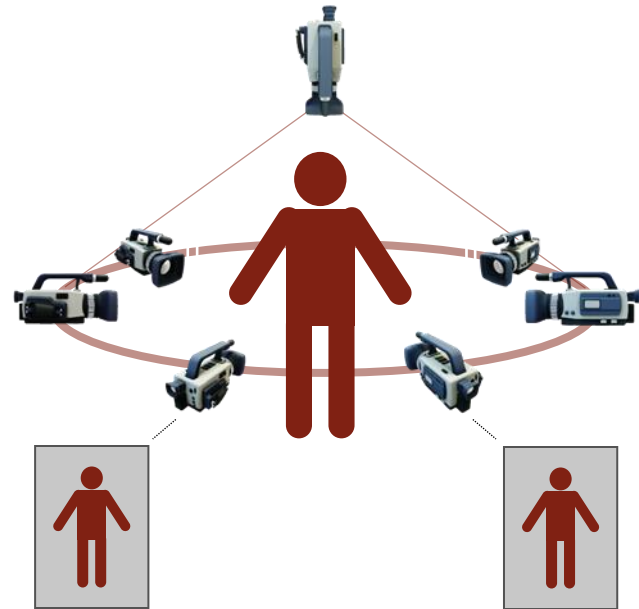
- Multiple viewpoints
- Camera placement
- View overlap



Multi-View Capture Systems



“What could go wrong?”



Camera Synchronization



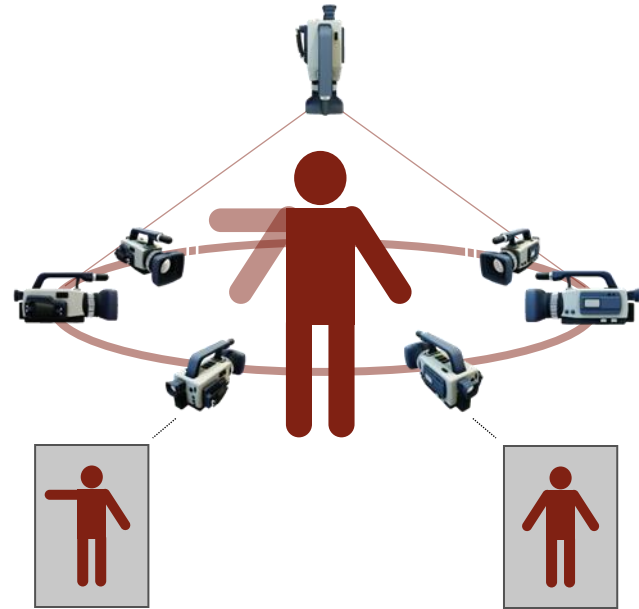
Challenge

- Time consistency between views



Solution

- *Hardware or software synchronization*



Camera Registration



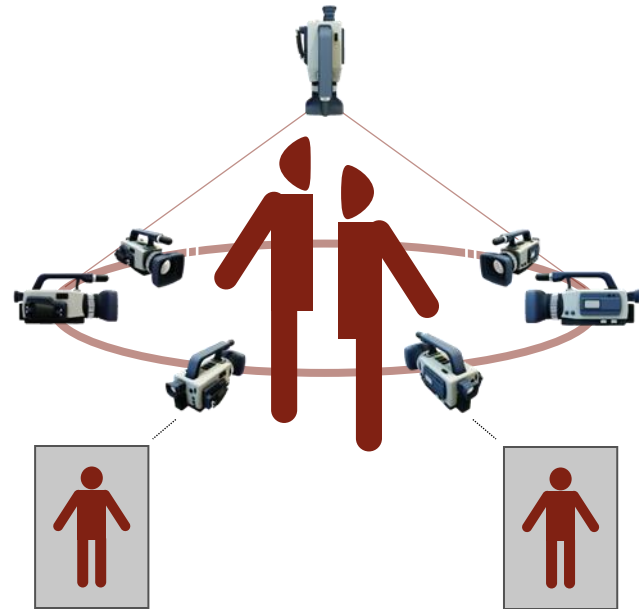
Challenge

- Align all views in a common 3D space



Solution

- Camera registration



Color Calibration



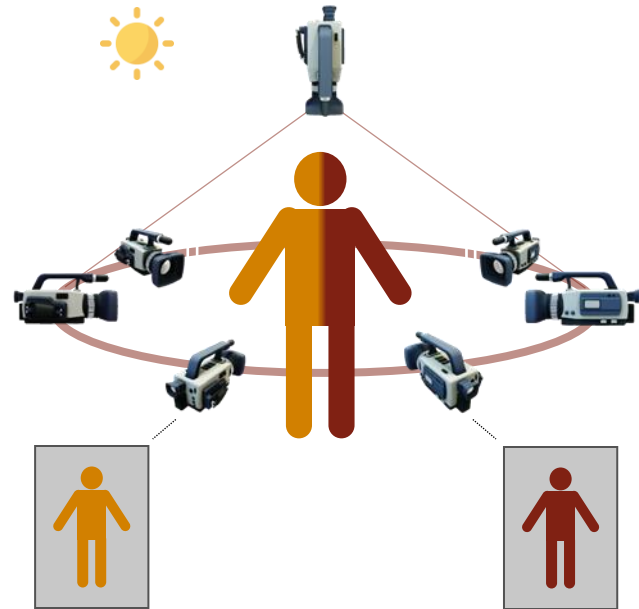
Challenge

- Inconsistent color between views

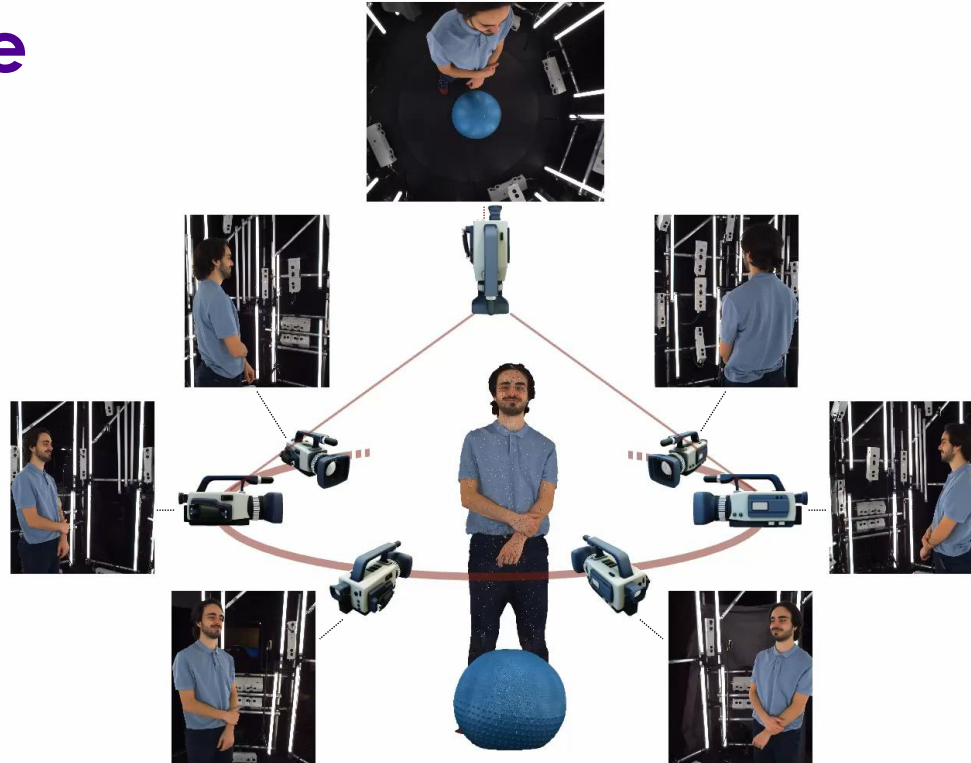


Solution

- Controlled lightning condition
- Color calibration

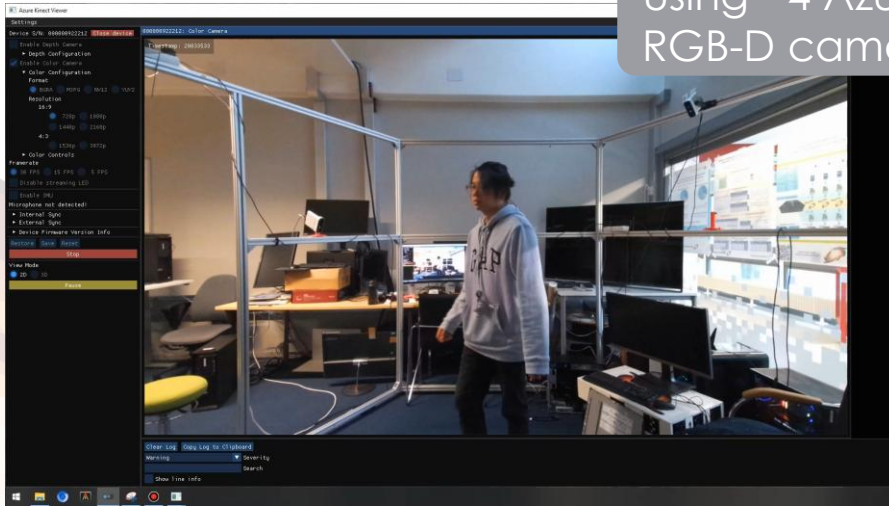


What We Aim to Capture



What Do We Actually Get?

Using ~4 Azure Kinect RGB-D cameras



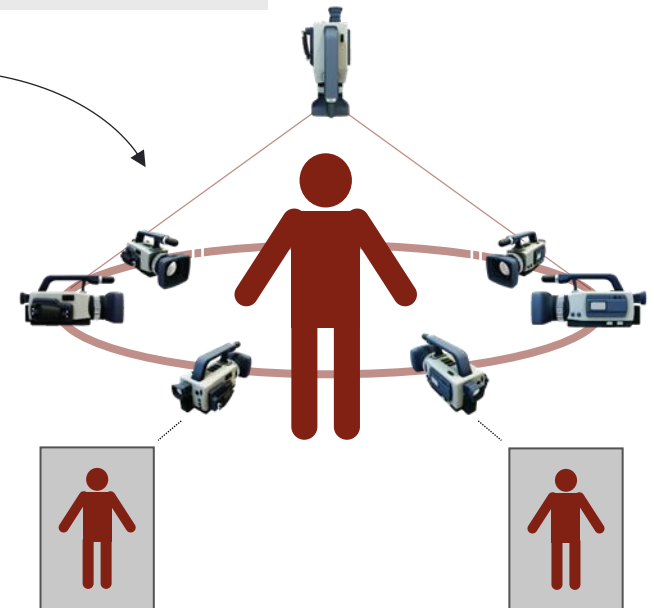
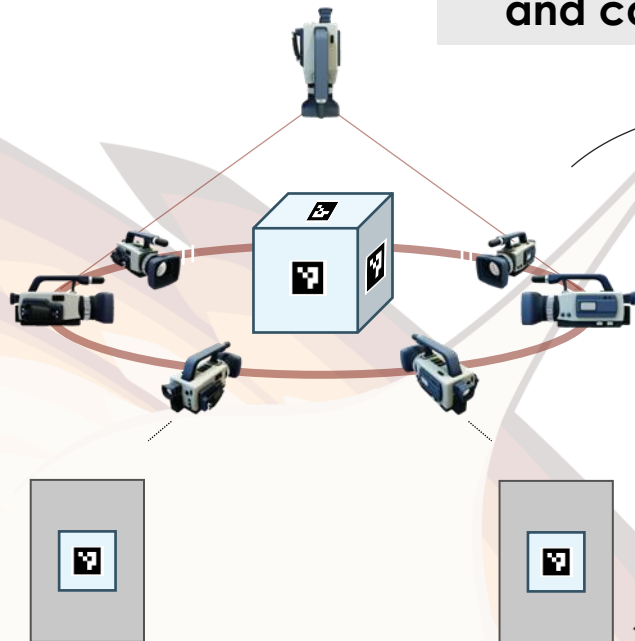
ULTRA VIDEO GROUP

Capture System Steps

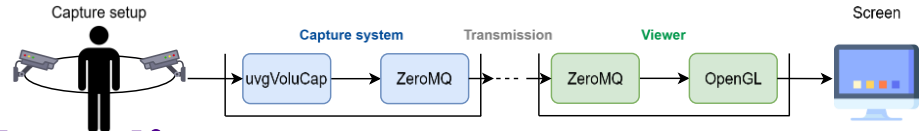
Offline registration
and calibration

Real-time online
reconstruction

Viewpoint
transform matrices

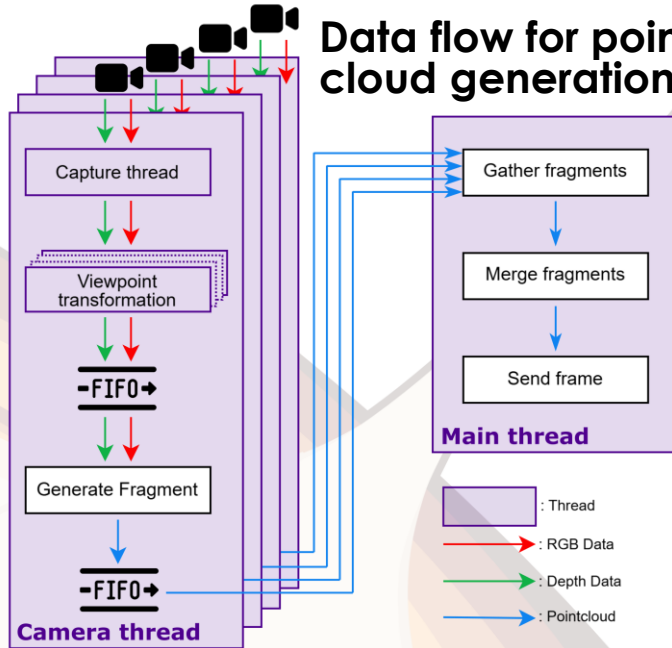


ULTRA VIDEO GROUP

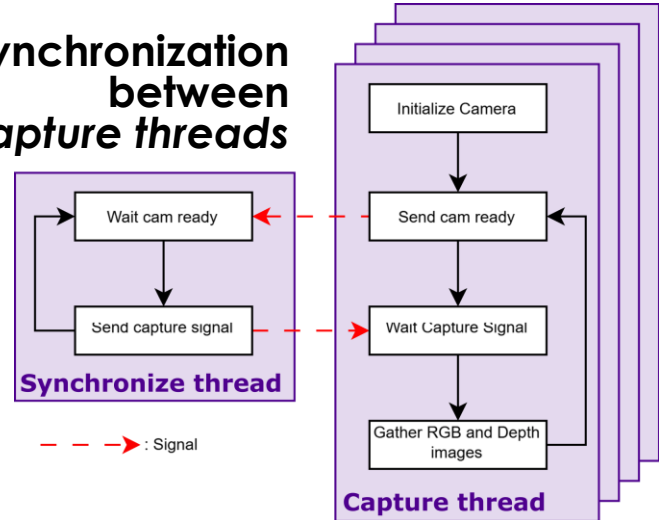


Real-time Online Reconstruction

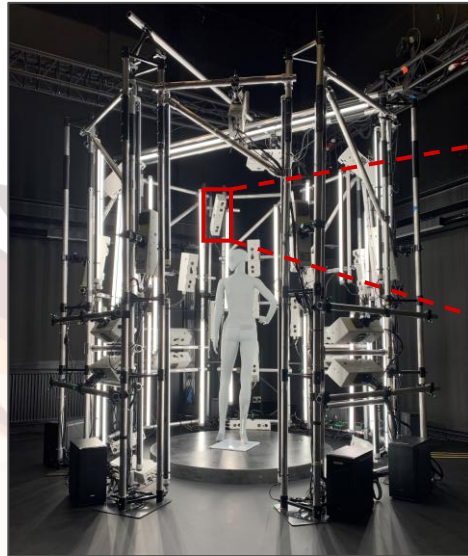
Data flow for point cloud generation



Synchronization between capture threads



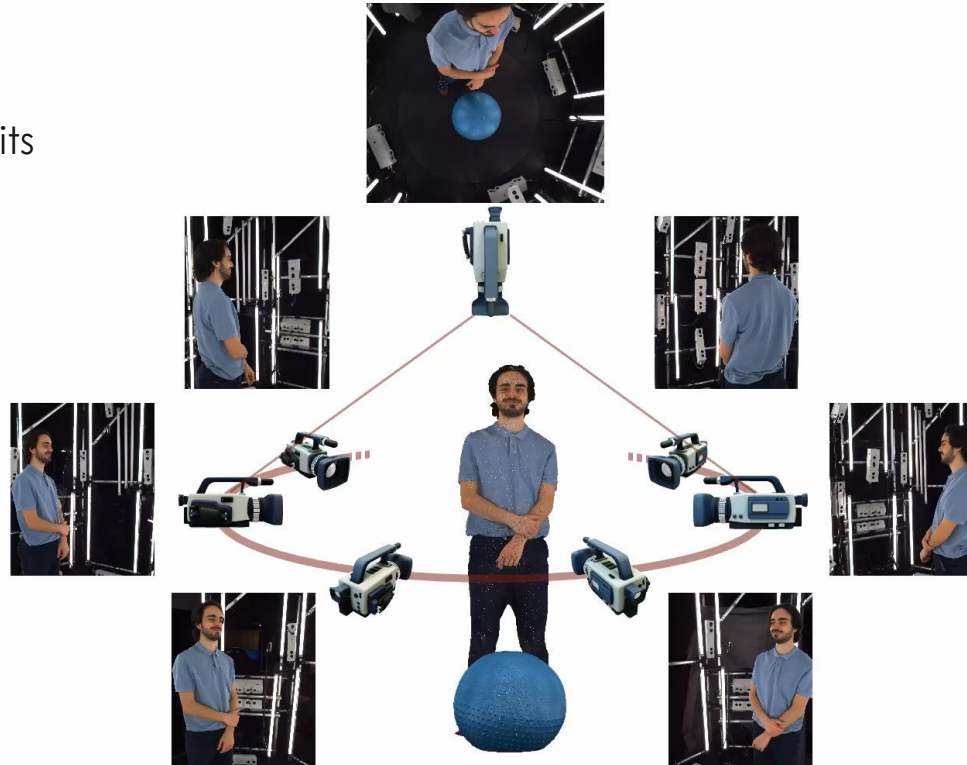
Professional Volumetric Capture Systems

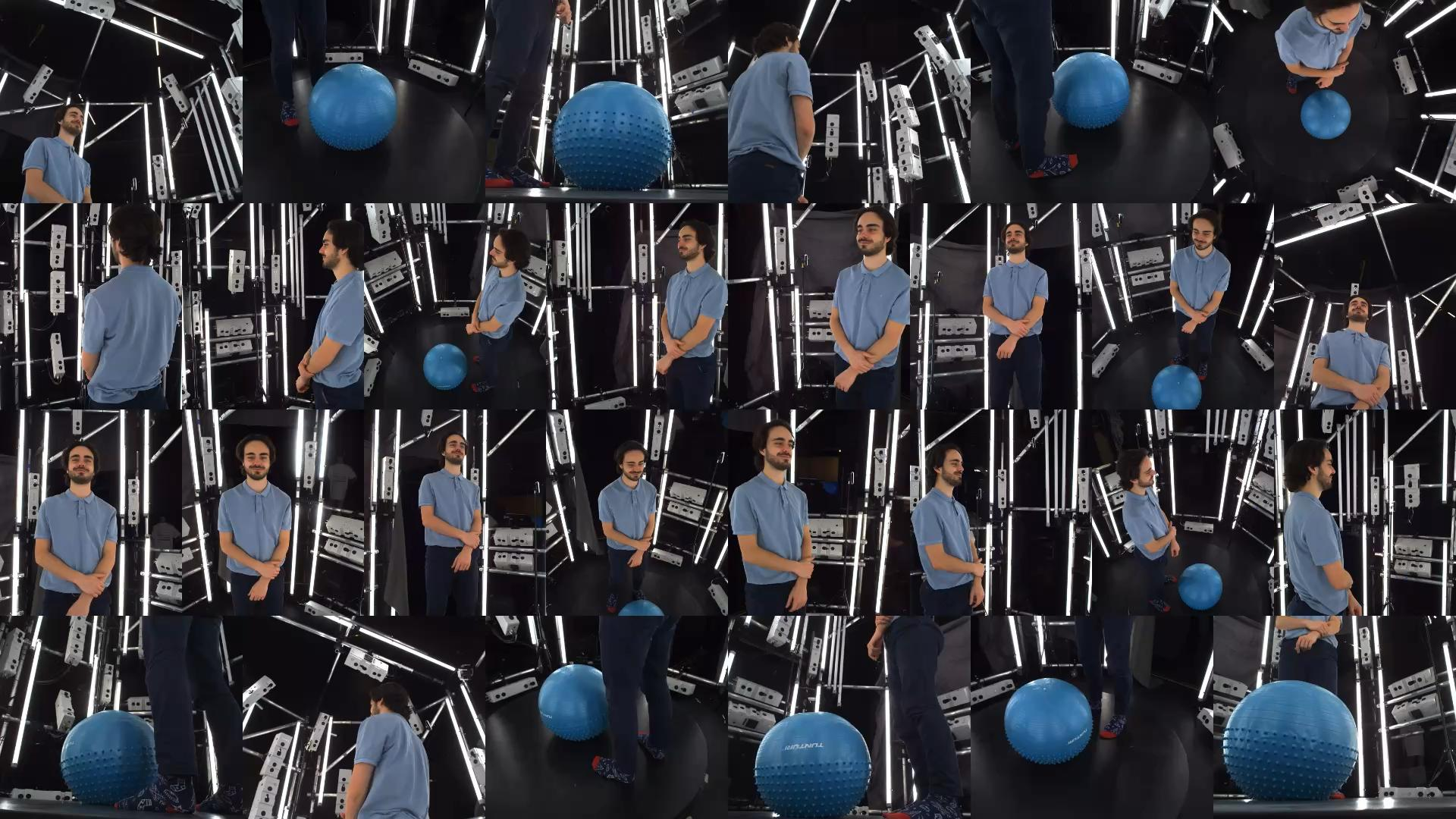


32 camera units



CIVIT Volumetric Studio
(Tampere University)





Hands-On: Raw and Post-process Voxelized Point Clouds

#demo2



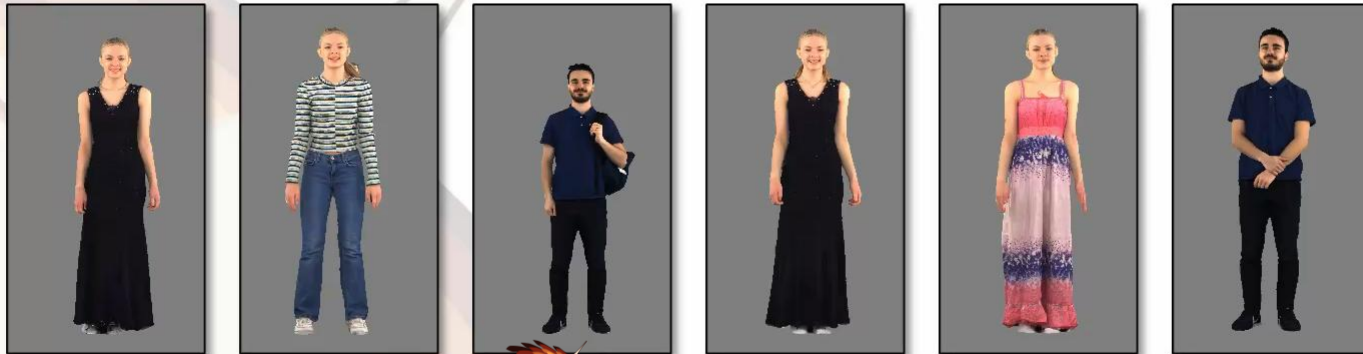
<https://ultravideo.fi/point-cloud-demo/>



UVG-VPC point cloud dataset



- 12 sequences
- 10s long
- 25 fps
- Meshes, GS_(soon) available



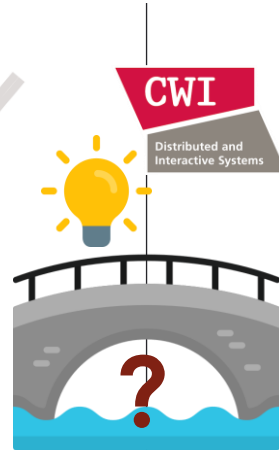
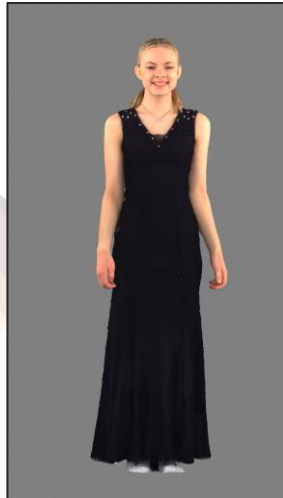
Link to dataset



ULTRA VIDEO GROUP

Current State-of-the-Art

Professional capture system



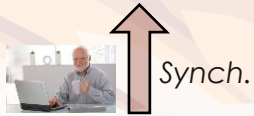
Consumer-grade capture system



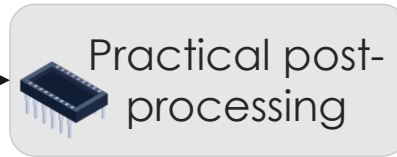
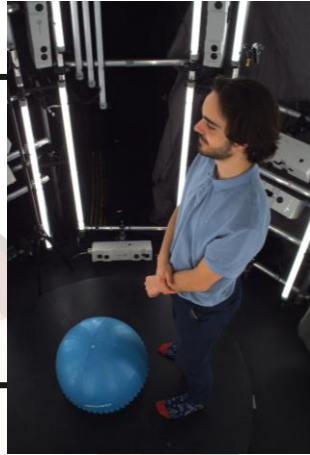
ULTRA VIDEO GROUP

Dual-Capture System for Point Clouds

High-end capture system



Consumer-grade capture system



ULTRA VIDEO GROUP



Hands-On: Dual Quality Point Clouds

#demo3



<https://ultravideo.fi/point-cloud-demo/>



ULTRA VIDEO GROUP

CWI

Distributed and
Interactive Systems

UVG-CWI-DQPC

Dual-Quality Point Cloud Dataset

Dataset in brief:

- 12 dual sequences
- 157–201 frames per sequences
- Intermediary data available

Support development and benchmarking of:

- Point cloud enhancement
- Point cloud compression
- Quality assessment



Grand challenge

- Team Registration : 24-April-2026
- Software Submission : 13-May-2026
- Leaderboard Reveal : 18-May-2026
- Paper Submission : 28-May-2026



ACM
Multimedia 2026
Rio de Janeiro, Brazil

Welcome to the UVG-CWI-DQPC Challenge 2026!

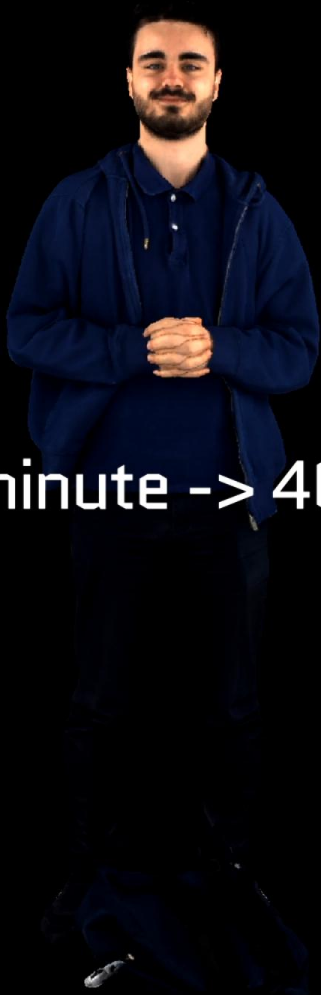
Enhancing Dynamic Point Clouds in the Wild: A Grand Challenge on Real-World 4D Volumetric Data

The goal of the Grand Challenge is to advance algorithms that enhance real-world, dynamic (4D) color point clouds by producing accurate, temporally consistent, and visually faithful reconstructions from consumer-grade captures.



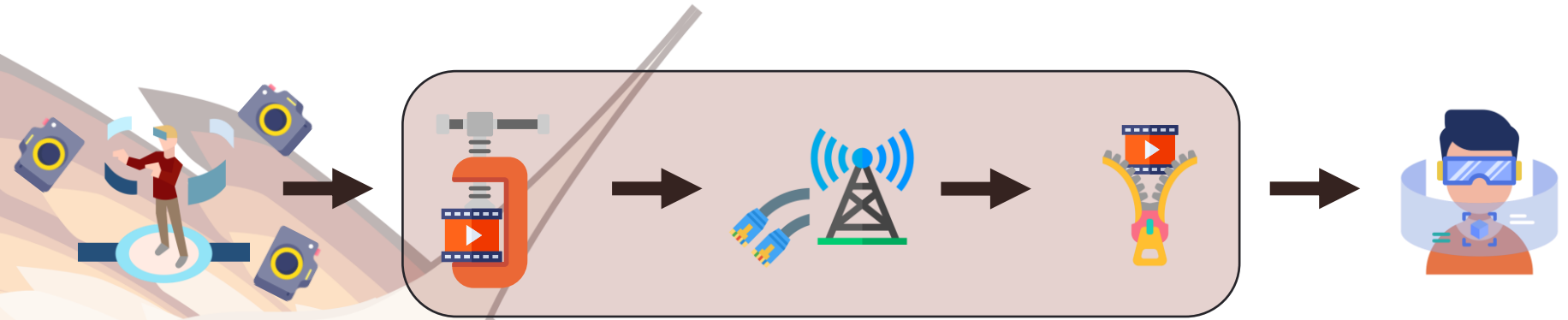
ULTRA VIDEO GROUP



A man with a beard and mustache, wearing a dark blue zip-up jacket over a dark blue polo shirt and dark pants, stands with his hands clasped in front of him. He is centered against a solid black background. The text "1 minute -> 40GB" is overlaid on the image in white, bold, sans-serif font.

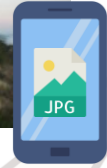
1 minute -> 40GB

Point Cloud Compression



You Compress Data Every Day

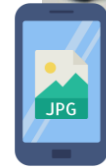
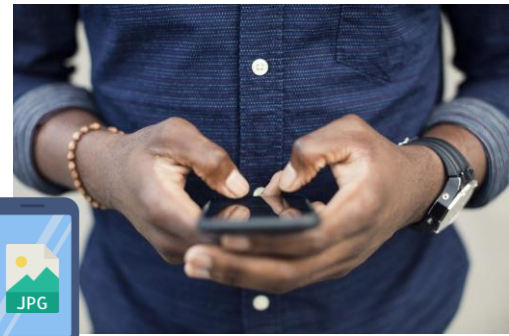
Alice



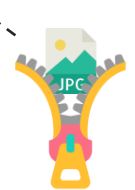
Bitstream
0110101



Bob



ULTRA VIDEO GROUP



Decoder

Lossless vs Lossy Compression



Lossless compression:

Compression that reduces data size while allowing the original signal to be reconstructed exactly, with no loss of information.



<https://pomodo.io/tech-archive/jpeg-definitive-guide/>

Lossy compression:

Compression that reduces data size by allowing controlled loss of information, trading perfect reconstruction for higher compression efficiency.



JPEG Compression Quality vs Size



Quality 10 - 49.1 kB

Quality 40 - 156 kB

Quality 70 - 261 kB

Quality 100 - 2.2 MB

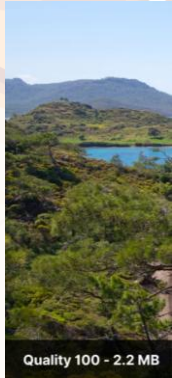
Introduction to Rate-distortion Paradigm



"Reducing bitrate comes at the cost of visual quality."



"Improving visual quality requires a higher bitrate."



Trade-off



Visual quality (distortion):

The perceived fidelity of the reconstructed signal compared to the original.

Bitrate (rate):

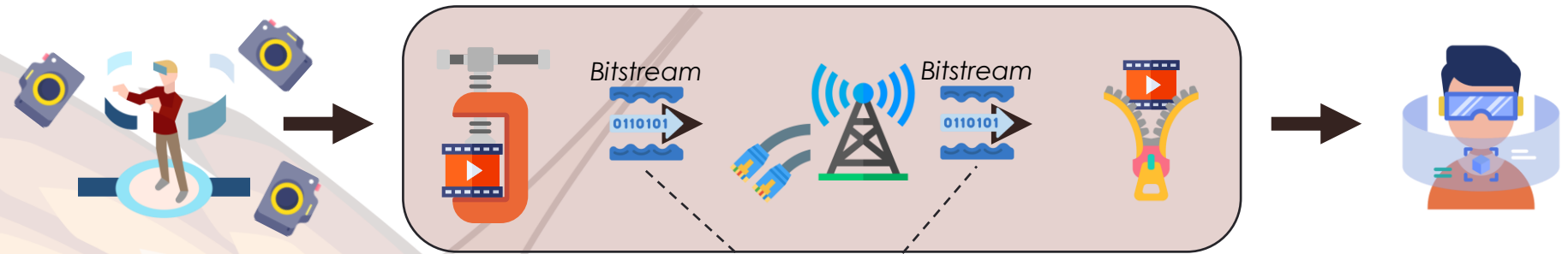
The amount of data used to represent a signal per unit of time or per sample.



ULTRA VIDEO GROUP

“To ensure interoperability between systems.”

Why Coding Standards Matter

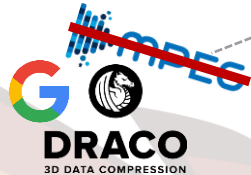


**Coding standard
(codec)**



Volumetric Compression Standards Overview

Geometric-based



G-PCC

Geometry-based Point
Cloud Compression

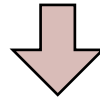


ULTRA VIDEO GROUP

Video-based

V3C

Visual Volumetric
Video-based Coding



V-PCC

Video-based Point
Cloud Compression



V-DMC

Video-based Dynamic
Mesh Coding

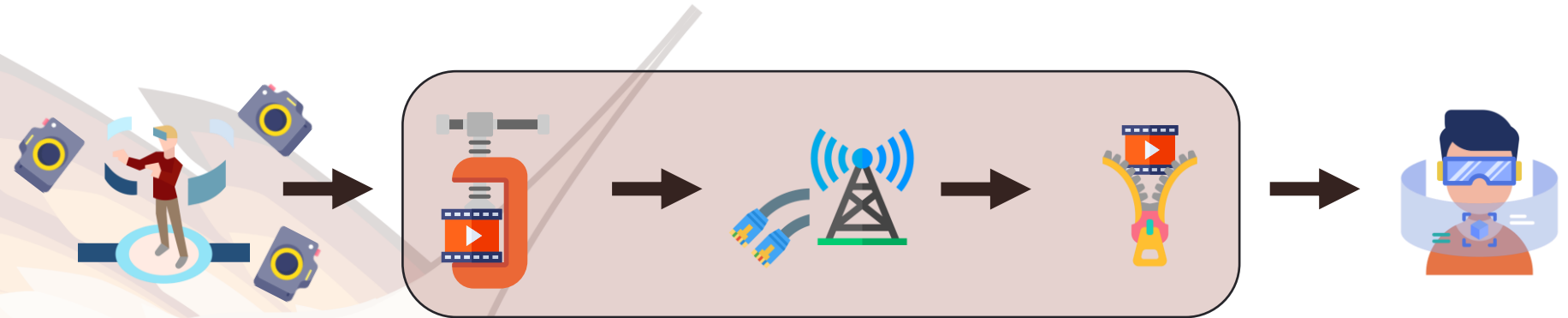


MIV

MPEG Immersive
video

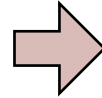
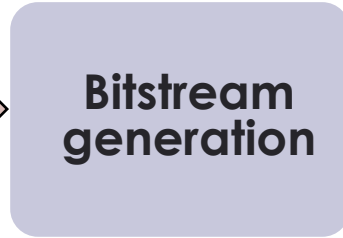
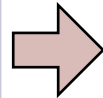
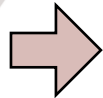


Volumetric Communication Pipeline

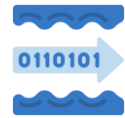


V-PCC Encoding Overview

Input point cloud frame



Output V3C bitstream

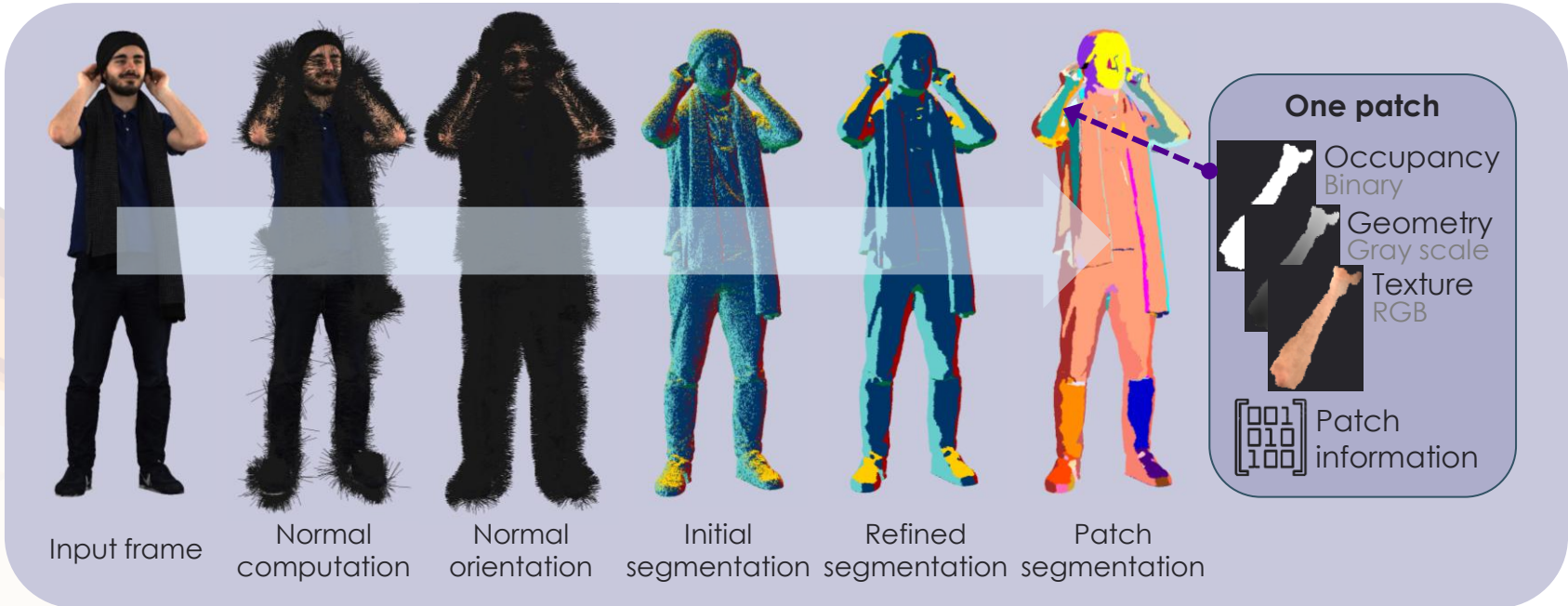


Projection plane mapping





V-PCC Patch Generation





Attribute image



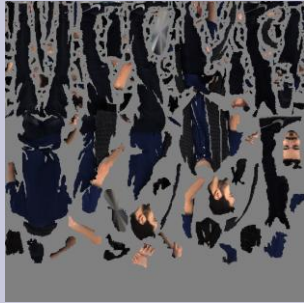
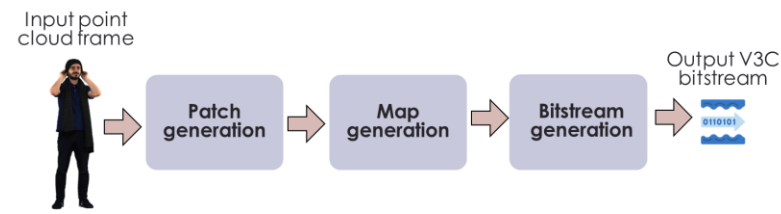
Geometry image



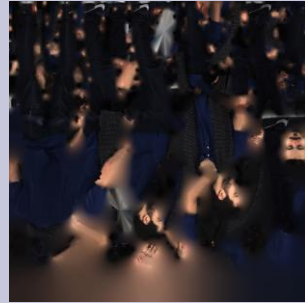
Occupancy image



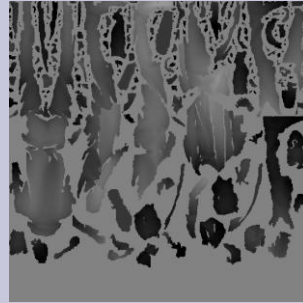
V-PCC Map Generation



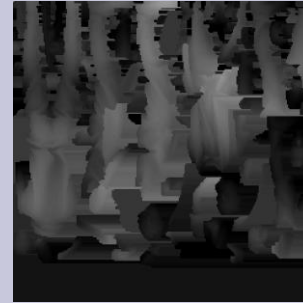
Attribute map



Attribute map with background filling



Geometry map



Geometry map with background filling



Occupancy map





Traditional 2D
video encoder

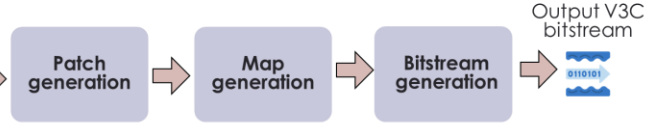


Traditional 2D
video encoder

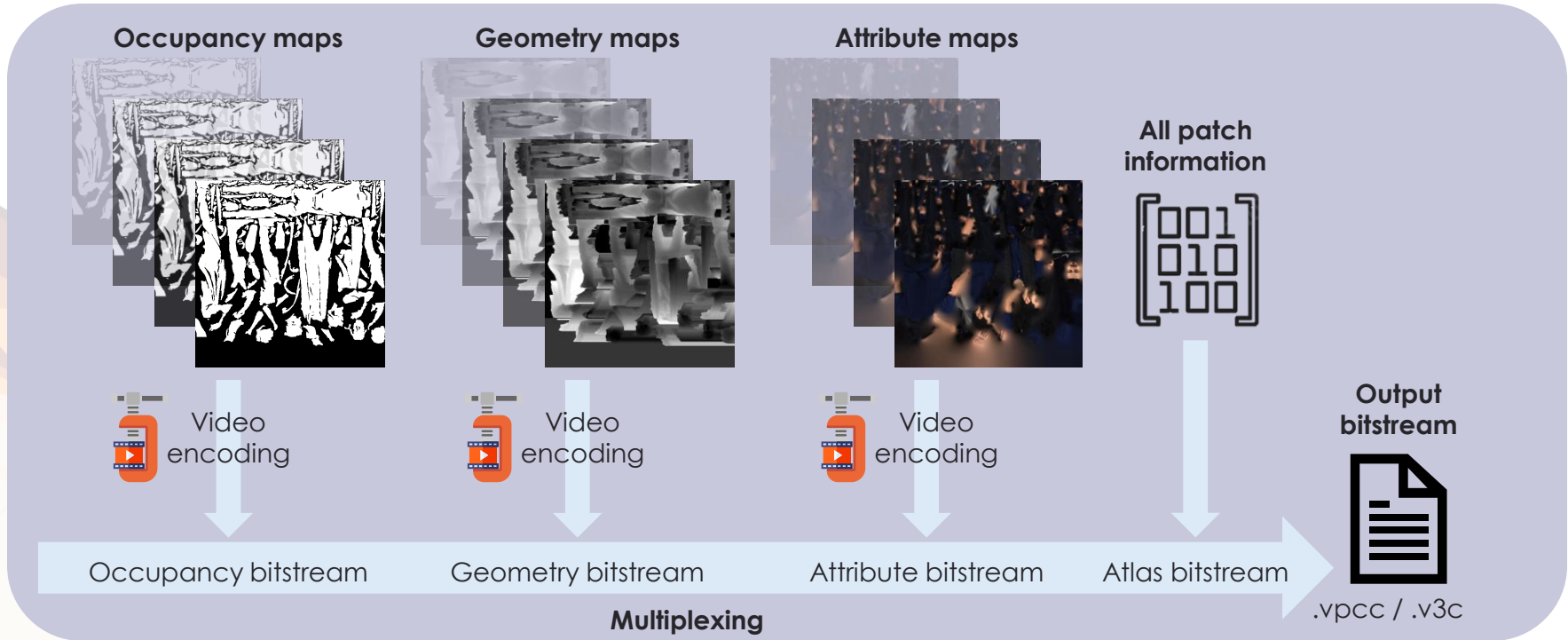


Traditional 2D
video encoder

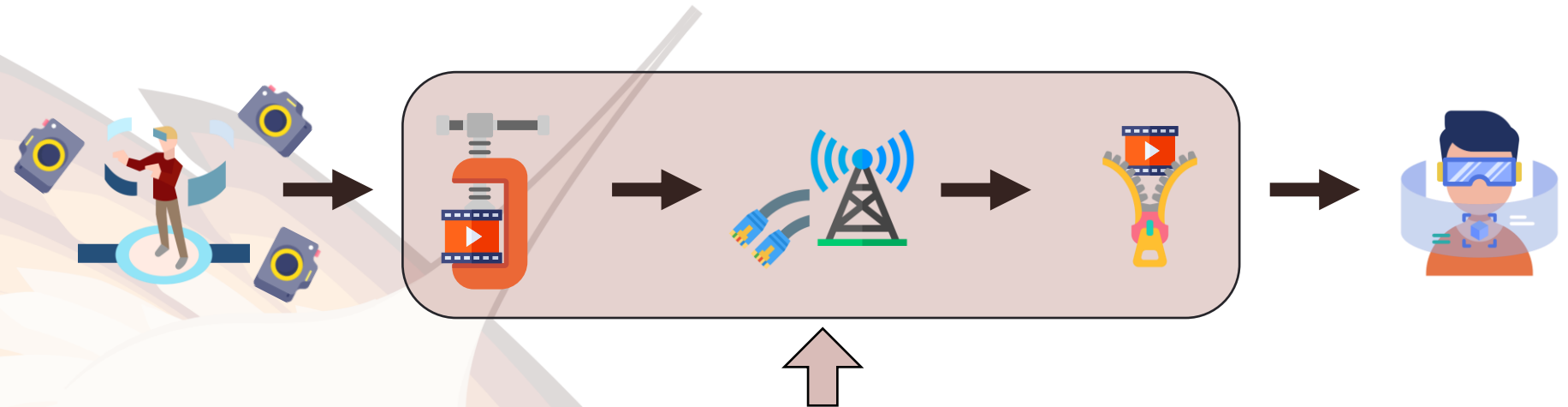
V3C Bitstream



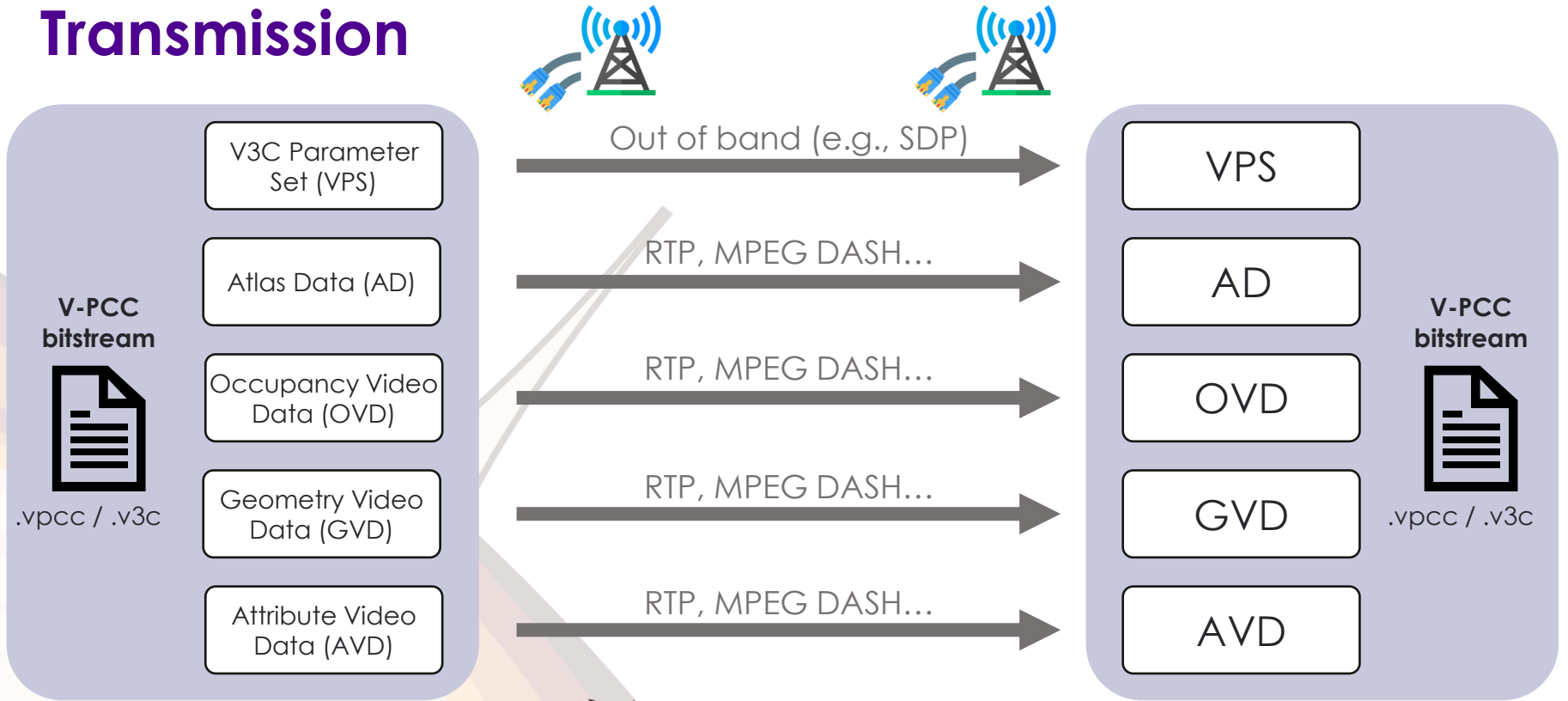
V-PCC Bitstream Generation



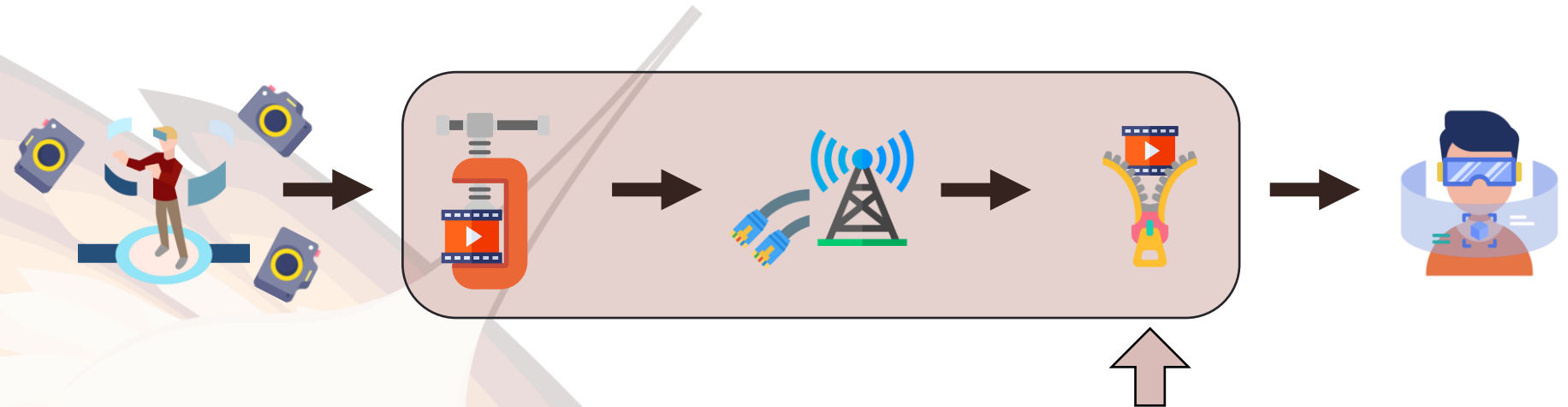
Volumetric Communication Pipeline



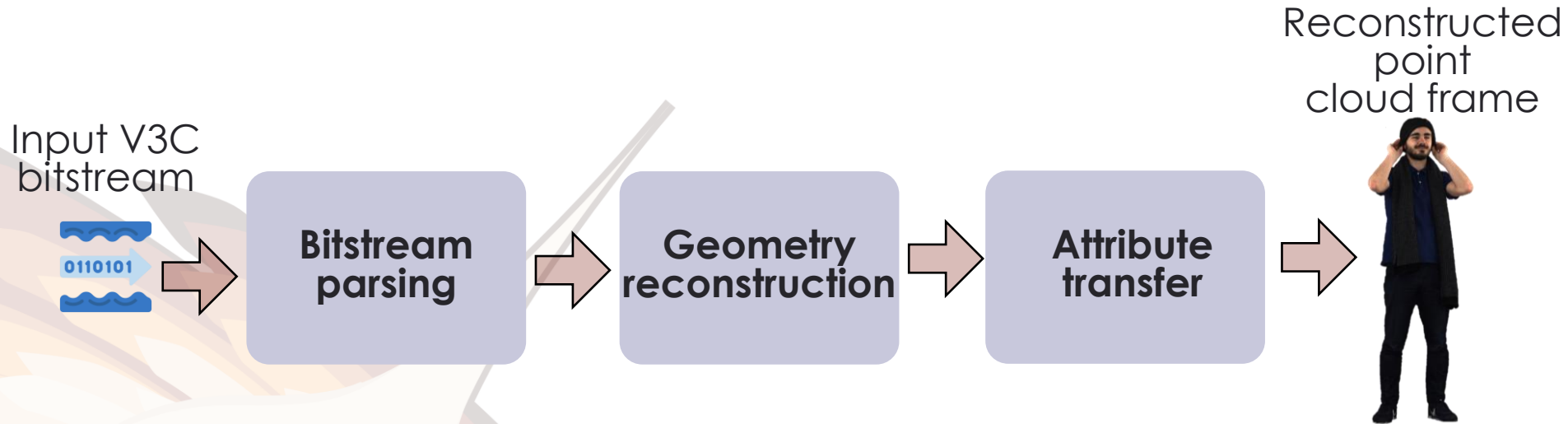
Transmission



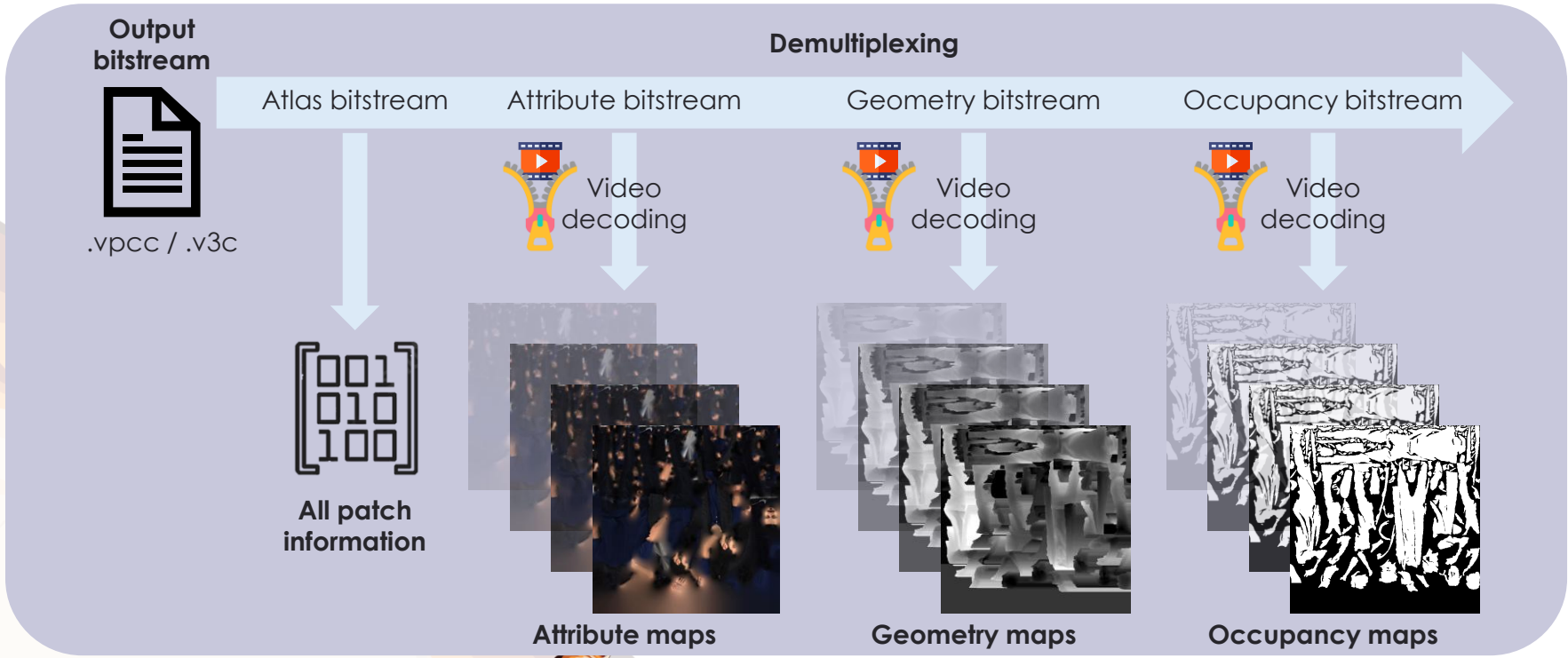
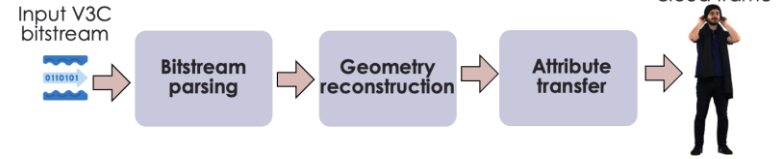
Volumetric Communication Pipeline



V-PCC Decoding Overview



V-PCC Bitstream Parsing

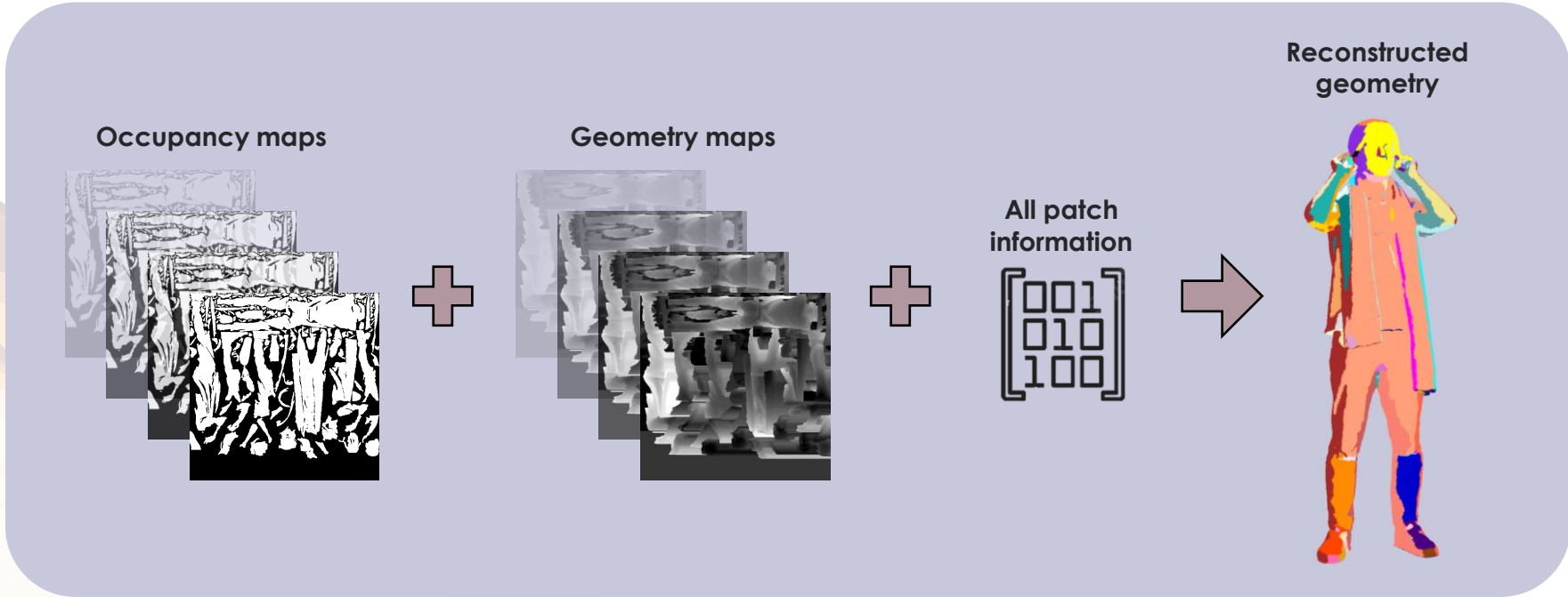


V-PCC Geometry Reconstruction

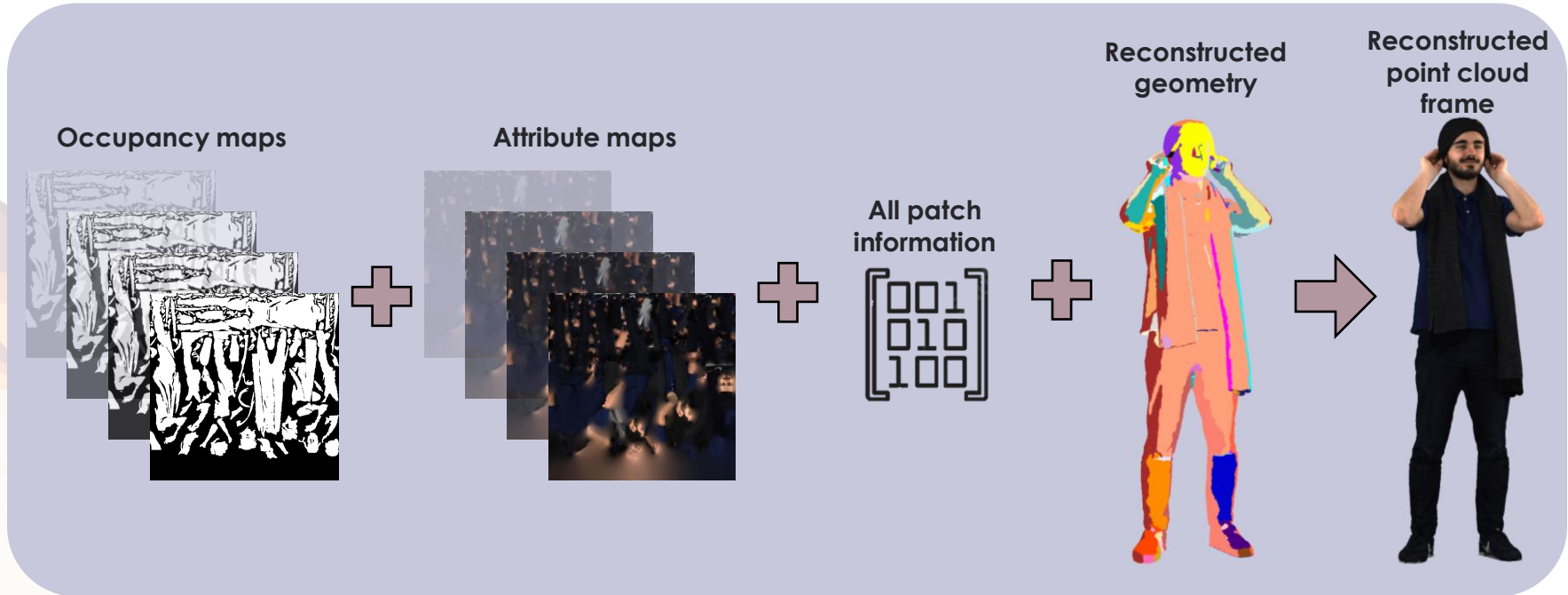
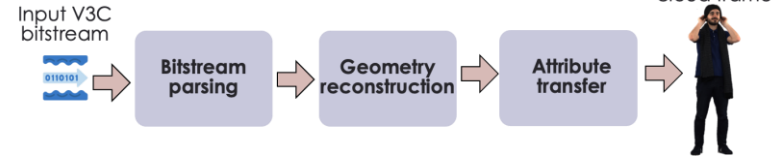
Input V3C bitstream



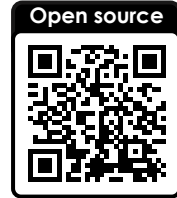
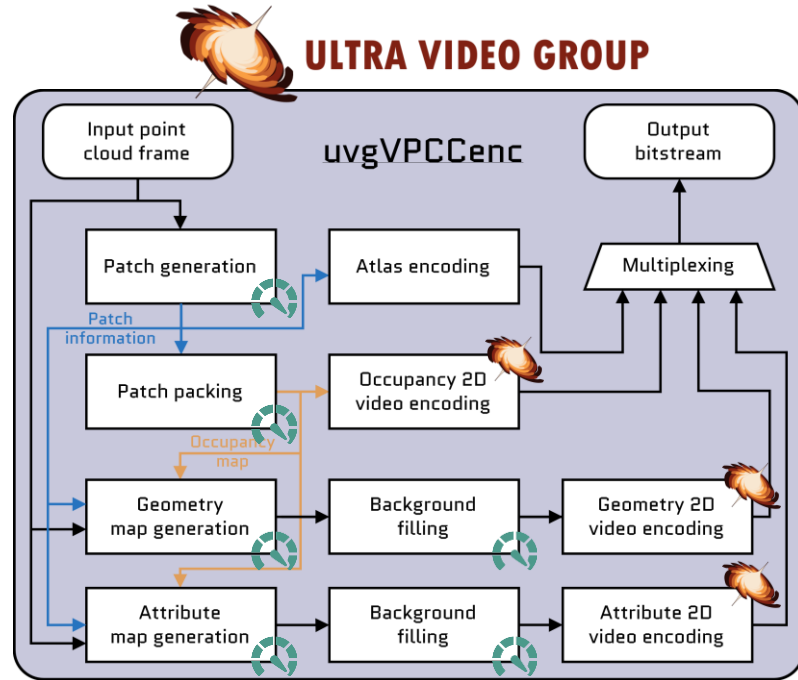
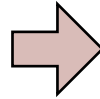
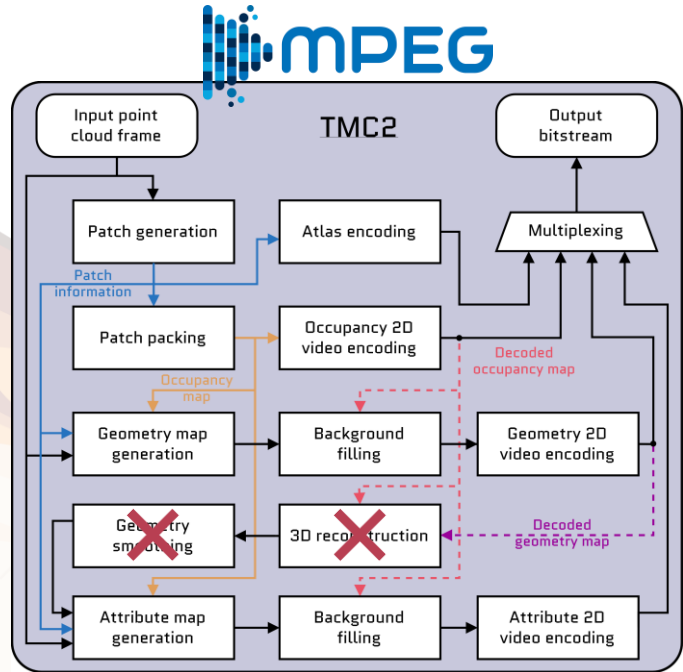
Reconstructed point cloud frame



V-PCC Attribute Transfer



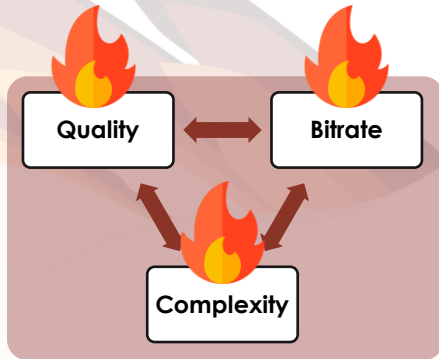
From V-PCC Standard to Practical Implementation



uvgVPCenc Performance

Encoder preset:

Predefined configurations of encoding tools for different trade-offs between encoding complexity and compression efficiency (quality/bitrate).



- **Slow preset:** High compression efficiency at the cost of encoding complexity
- **Fast preset:** : Low encoding complexity at the cost of compression efficiency

#demo4



<https://ultravideo.fi/point-cloud-demo/>



ULTRA VIDEO GROUP

Uncompressed



TMC2



uvgVPCcenc Preset slow



uvgVPCcenc Preset fast



91x speedup
over TMC2

729x speedup
over TMC2

Voxel 10

~709k points

3.6 Gbps

ULTRA VIDEO GROUP

62.3 sec/frame

4.68 Mbps

1.4 fps

6.87 Mbps

11.6 fps

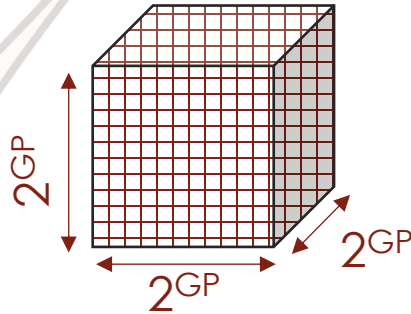
10.01 Mbps

uvgVPCenc Performance

Voxelization:

Sets the geometry precision (GP), i.e., spatial resolution, of the point cloud by quantizing 3D space into a regular voxel grid.

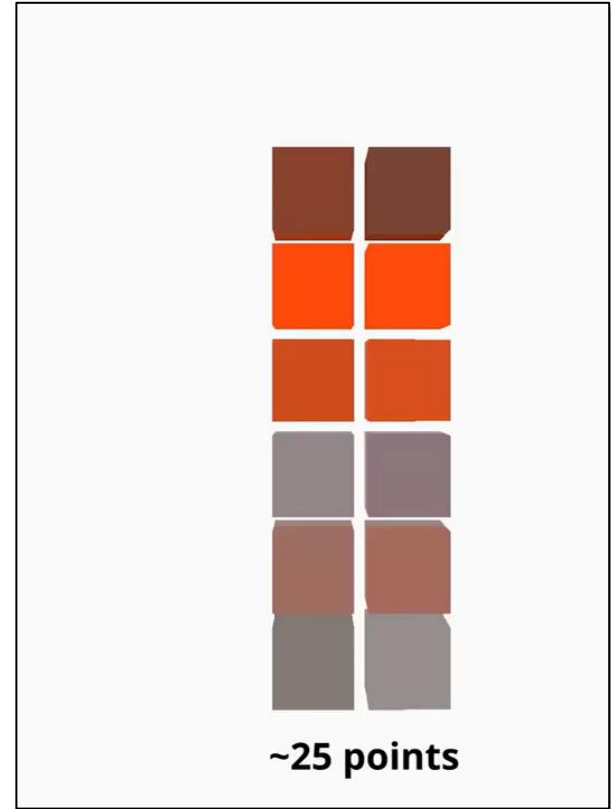
#demo5



<https://ultravideo.fi/point-cloud-demo/>



ULTRA VIDEO GROUP



Uncompressed

uvgVPCcenc - Slow

uvgVPCcenc - Fast

Voxel 9
~185k points



Voxel 10
~709k points



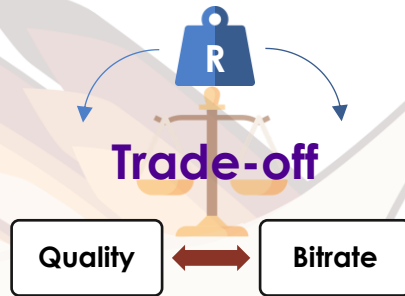
Voxel 11
~2.4M points



uvgVPCenc Performance

Rate (R1-5):

A control parameter of the encoder that sets the trade-off between bitrate and visual quality.



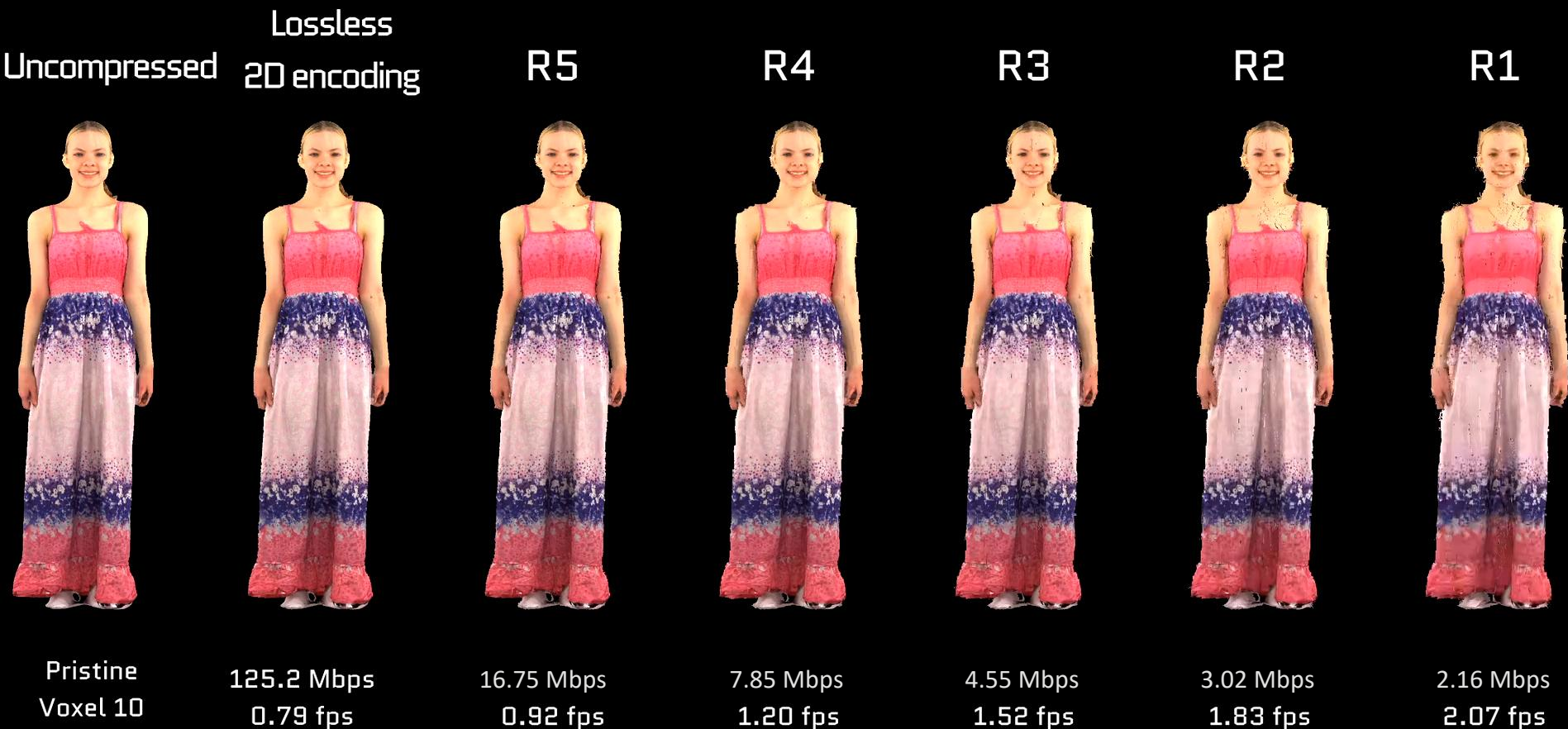
- $R \downarrow \rightarrow \text{bitrate} \downarrow, \text{quality} \downarrow$
- $R \uparrow \rightarrow \text{quality} \uparrow, \text{bitrate} \uparrow,$



<https://ultravideo.fi/point-cloud-demo/>



ULTRA VIDEO GROUP



~709k points

UVG Practical Volumetric Communication Pipeline

