Immersive Video Compression in the Learning Era

Giuseppe Valenzise

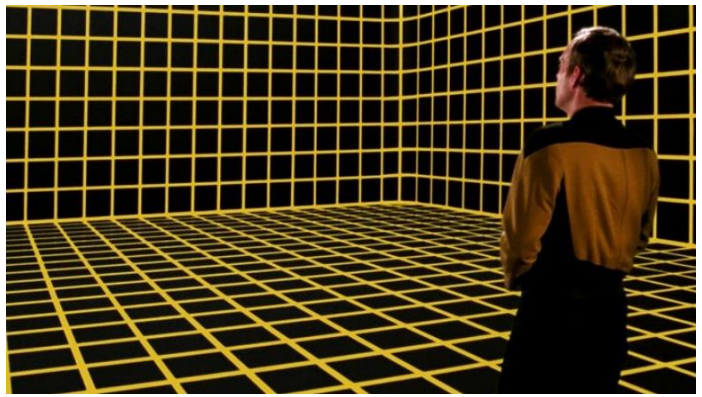
Laboratoire des Signaux et Systemes

Université Paris-Saclay, CNRS, CentraleSupelec, France

NEM Summit, Zagreb, May 25, 2023

Immersion and Telepresence

- Telepresence: term coined by M. Minsky in 1980
  - To develop a sense of being physically present at a remote location through interaction with the system's user interface
  - Emphasizes the importance of high-quality sensory feedback, such as vision, sound, and touch



Star Trek's Holodeck

realism and interaction

New video applications

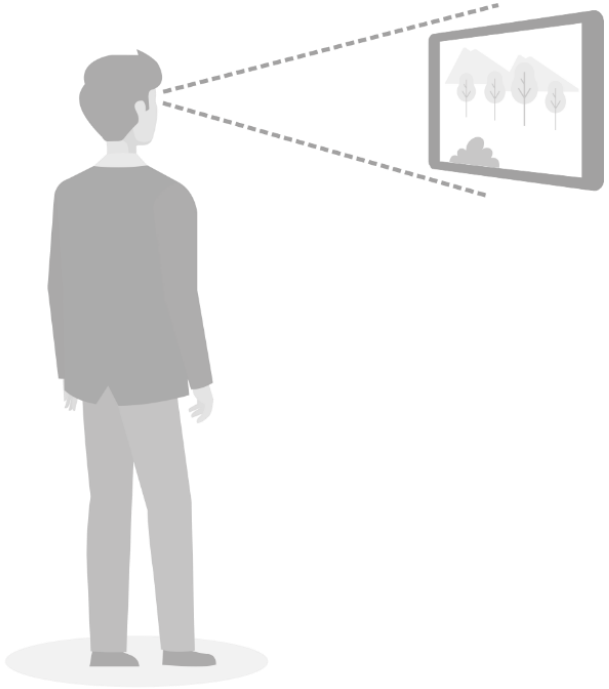
- Telepresence
- Collaborative working
- Telemedicine and remote surgery
- Storytelling and creative arts
- **Metaverse**
- …

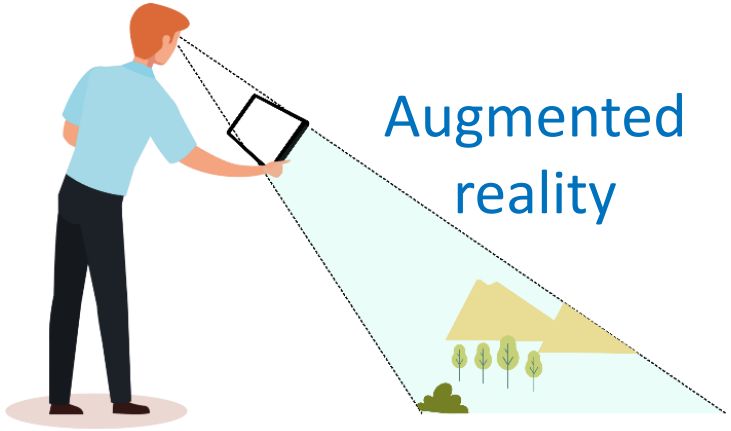
Immersive visual experience!

Immersive Video Technologies

360 degrees video

Traditional display



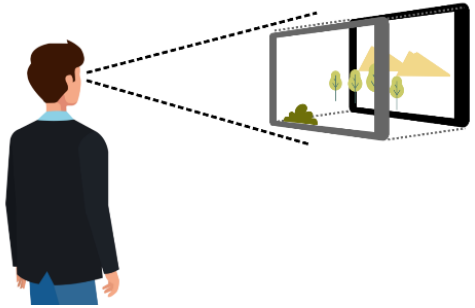


Augmented reality

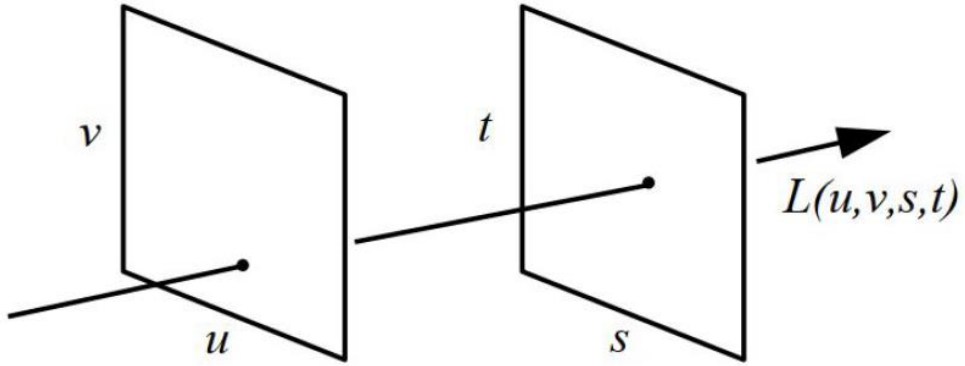
Volumetric video

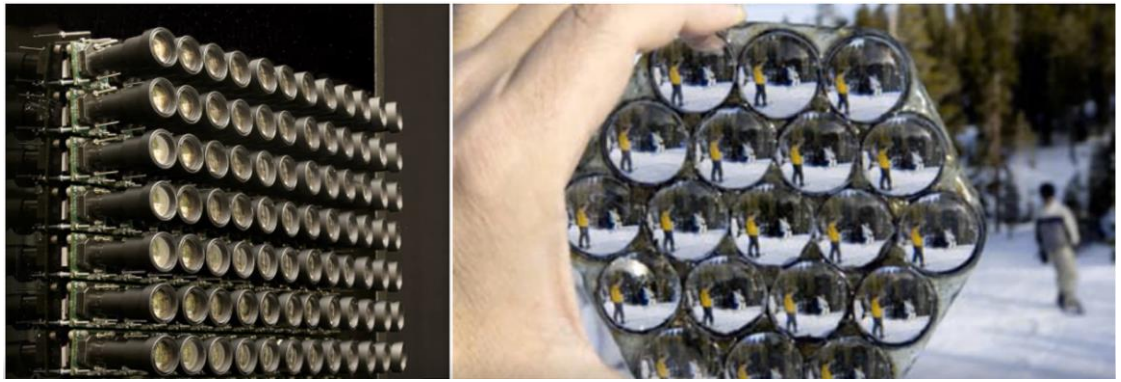
Virtual reality

Light fields

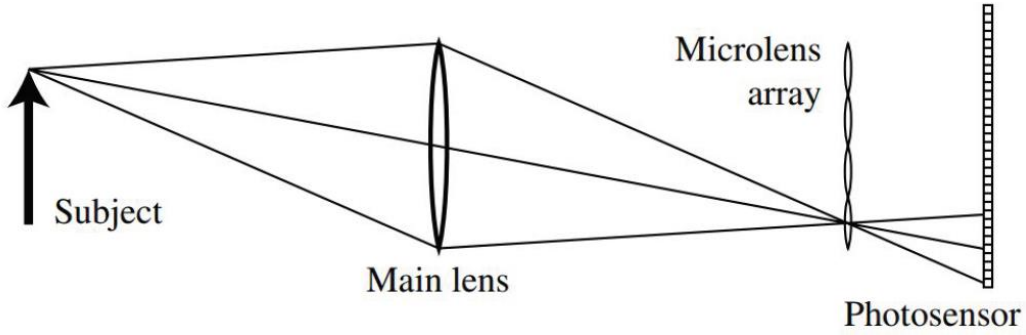


Light field





- 4D light field (lumigraph)
  - Constant radiance along its path
  - 2-planes parametrization
  - $L(u,v,s,t)$
- 2 kinds of acquisition:
  - Camera arrays
  - Microlens cameras (e.g., Lytro, Raytrix)

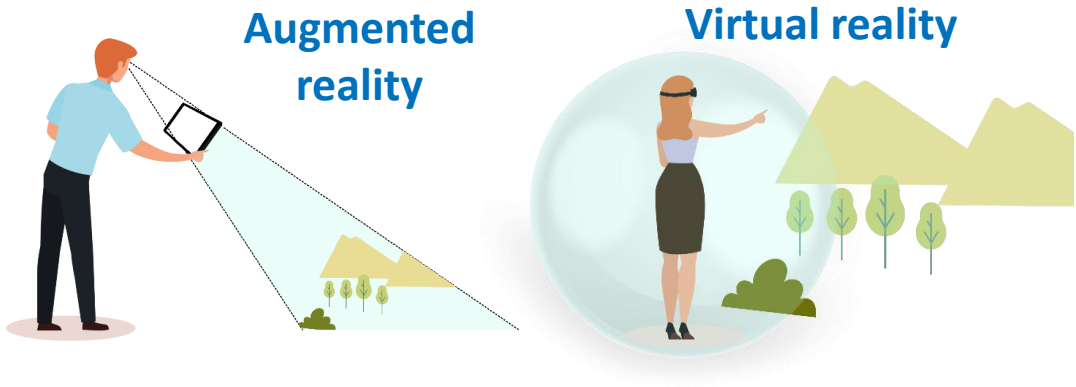


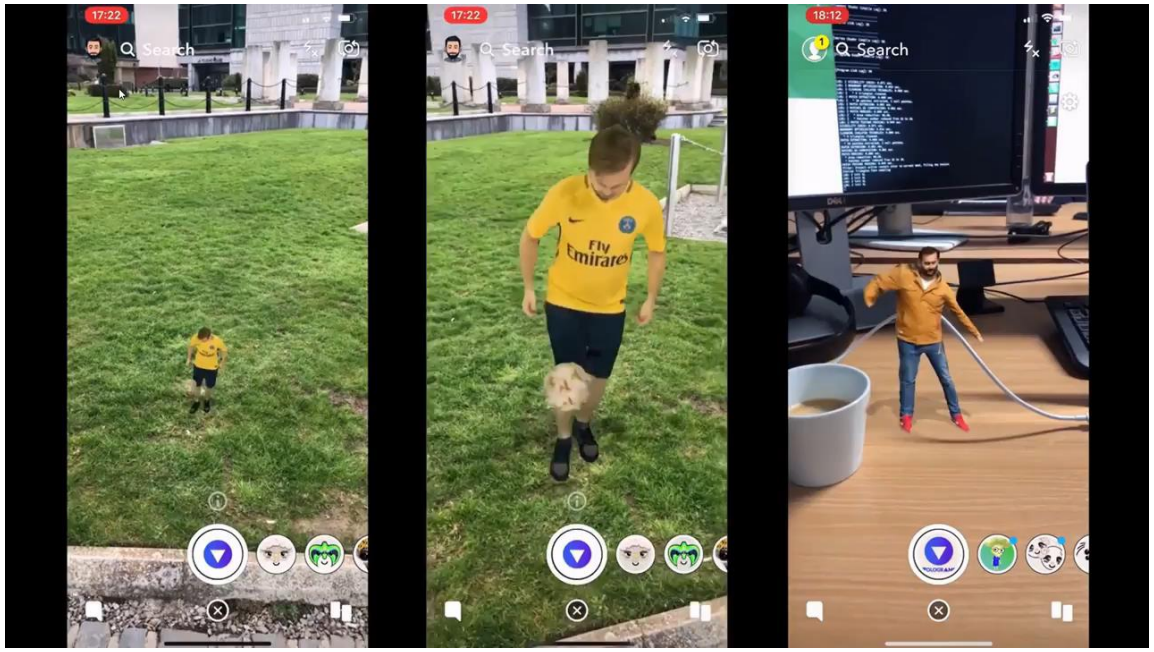
Light field

- Features
  - Viewpoint change (to some extend)
  - Refocusing

Volumetric video

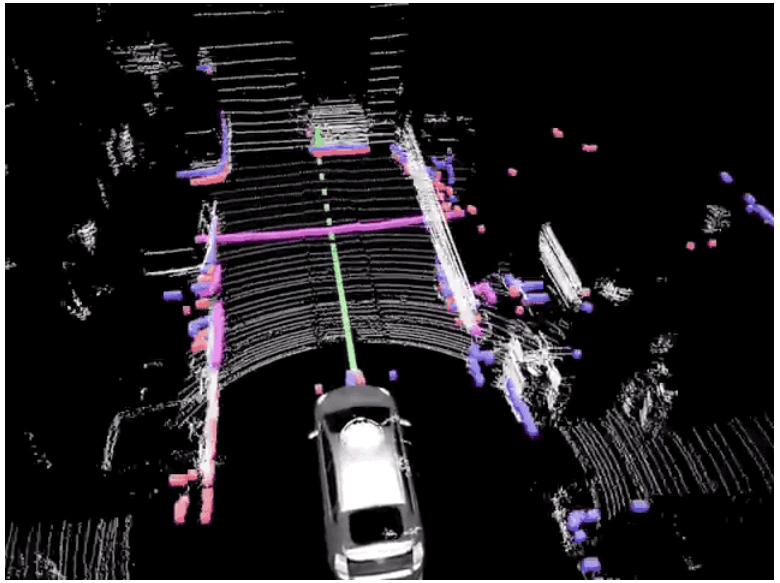
- 3D point clouds
- 3D meshes
- 6 degrees of freedom





Why compression is essential

- Typical point cloud video size:
  - 1 million points per frame
  - 30 frames/second
  - 32 bit geometry, 8 bits color -> ~ 3.6 Gbps

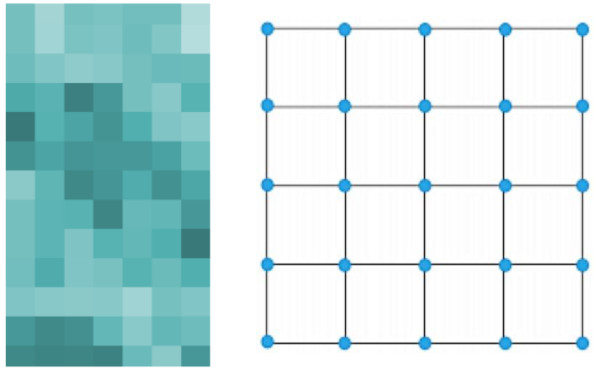


- Example: Velodyne HDL-64 LiDAR sensor
  - Over 100k points per sweep
  - 3 billion points per hour

Point Cloud Coding

Why coding point clouds is difficult?

Regular, dense



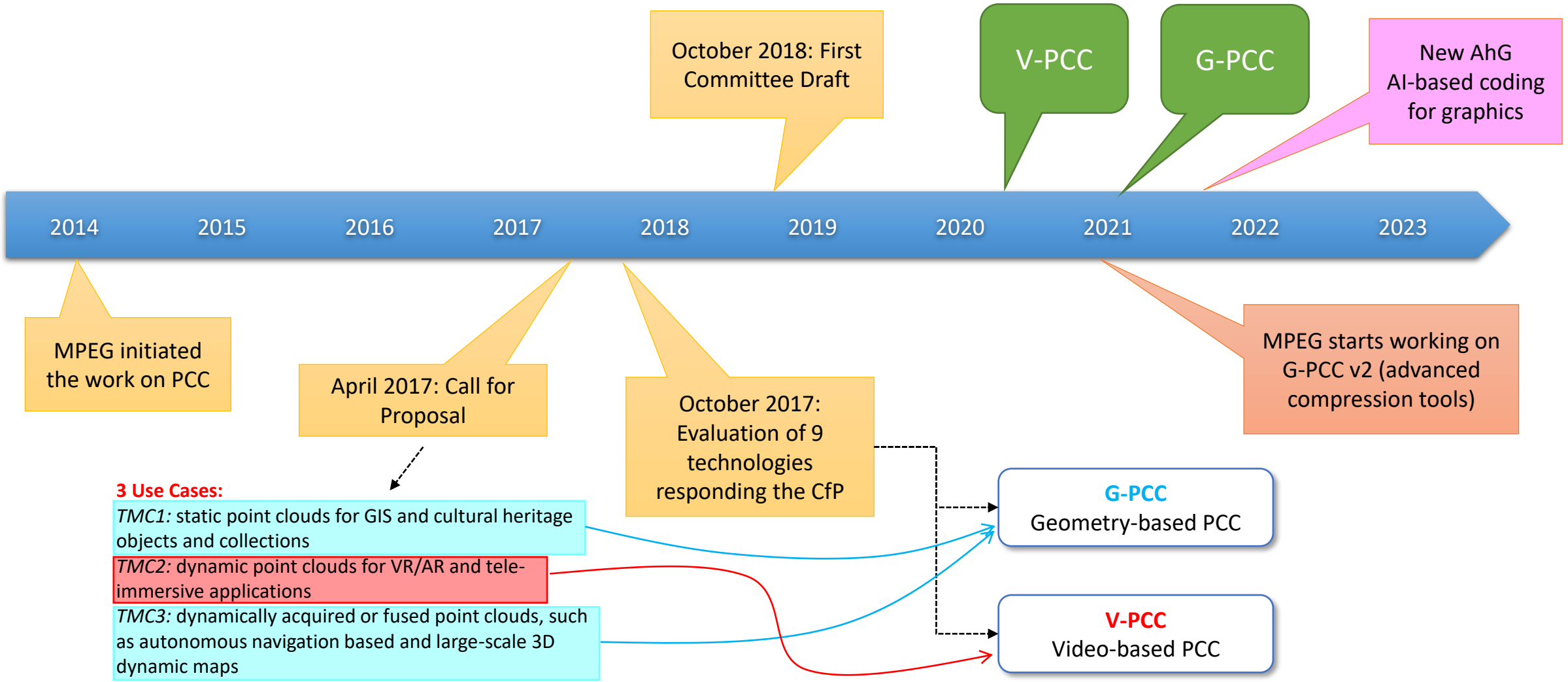
2D image

Irregular, sparse

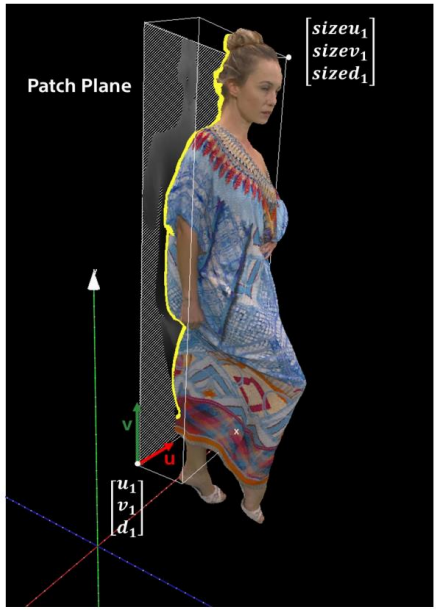
Point Cloud

Irregular sampling grid, **non-uniform** density, **sparsity**

MPEG standardization timeline



VPCC: Projection-based coding principle



Source code:
https://github.com/MPEGGroup/mpeg-pcc-tmc2

Occupancy map

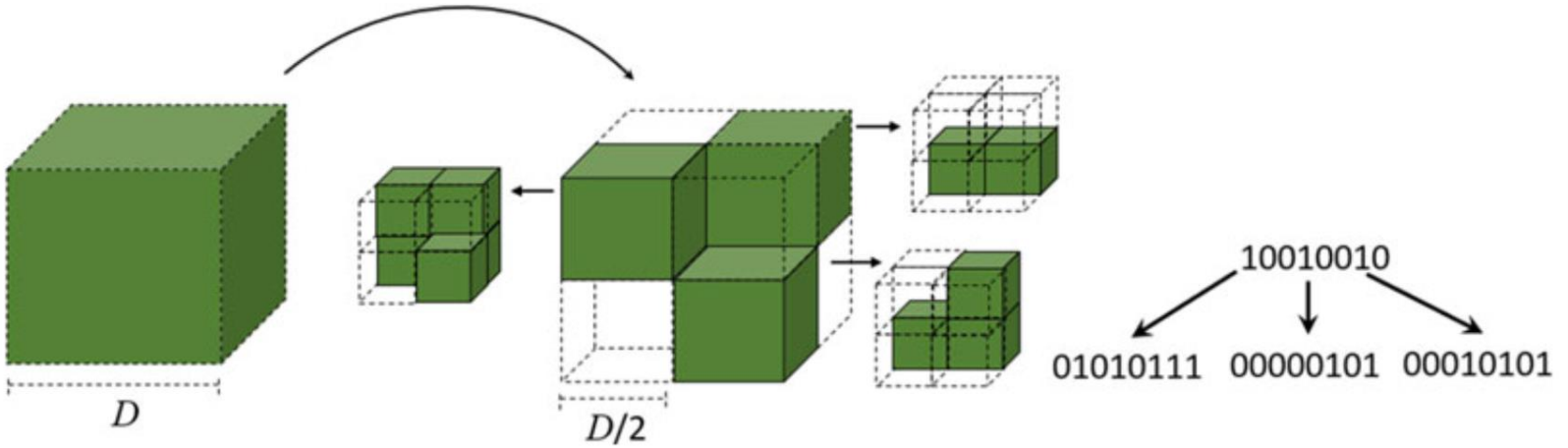
Geometry

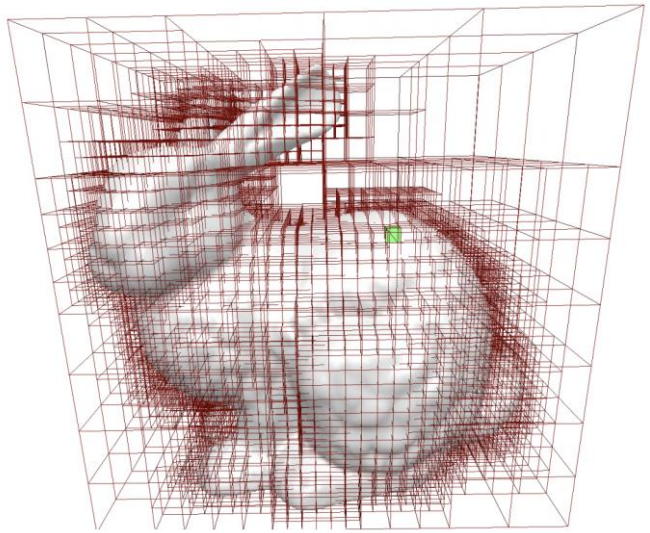
Attribute images

Graziosi, D., Nakagami, O., Kuma, S., Zaghetto, A., Suzuki, T. and Tabatabai, A., 2020. An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC). APSIPA Transactions on Signal and Information Processing, 9.

GPCC: Octree-based + wavelet transform

- Separate geometry and attribute coding
- Geometry: octree-based + trisoup, planar & angular modes
- Attributes: wavelet transform over tree





MPEG 3DG, V-PCC codec description, ISO/IEC JTC 1/SC 29/WG 7 N 0099, April 2021

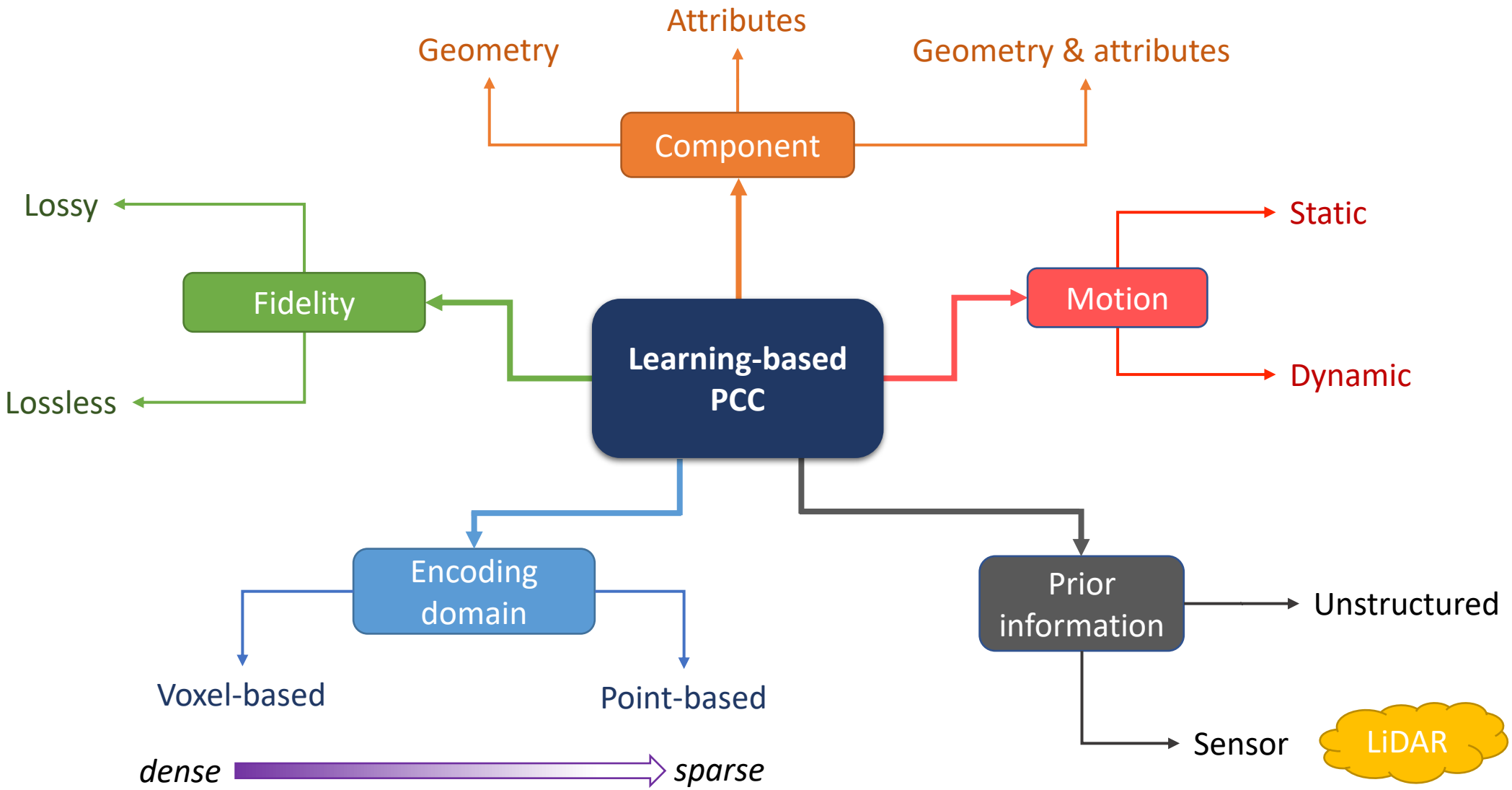
MPEG PCC standards

- V-PCC:
  - 2D projection-based
  - Dense PC
  - Dynamic content
  - AR/VR applications
- G-PCC:
  - Mostly **octree-based** + many optimizations
  - Static content
  - Low-to-high density
  - Wide range of applications: AR/VR, cultural heritage, LiDAR (fused and scans), etc.

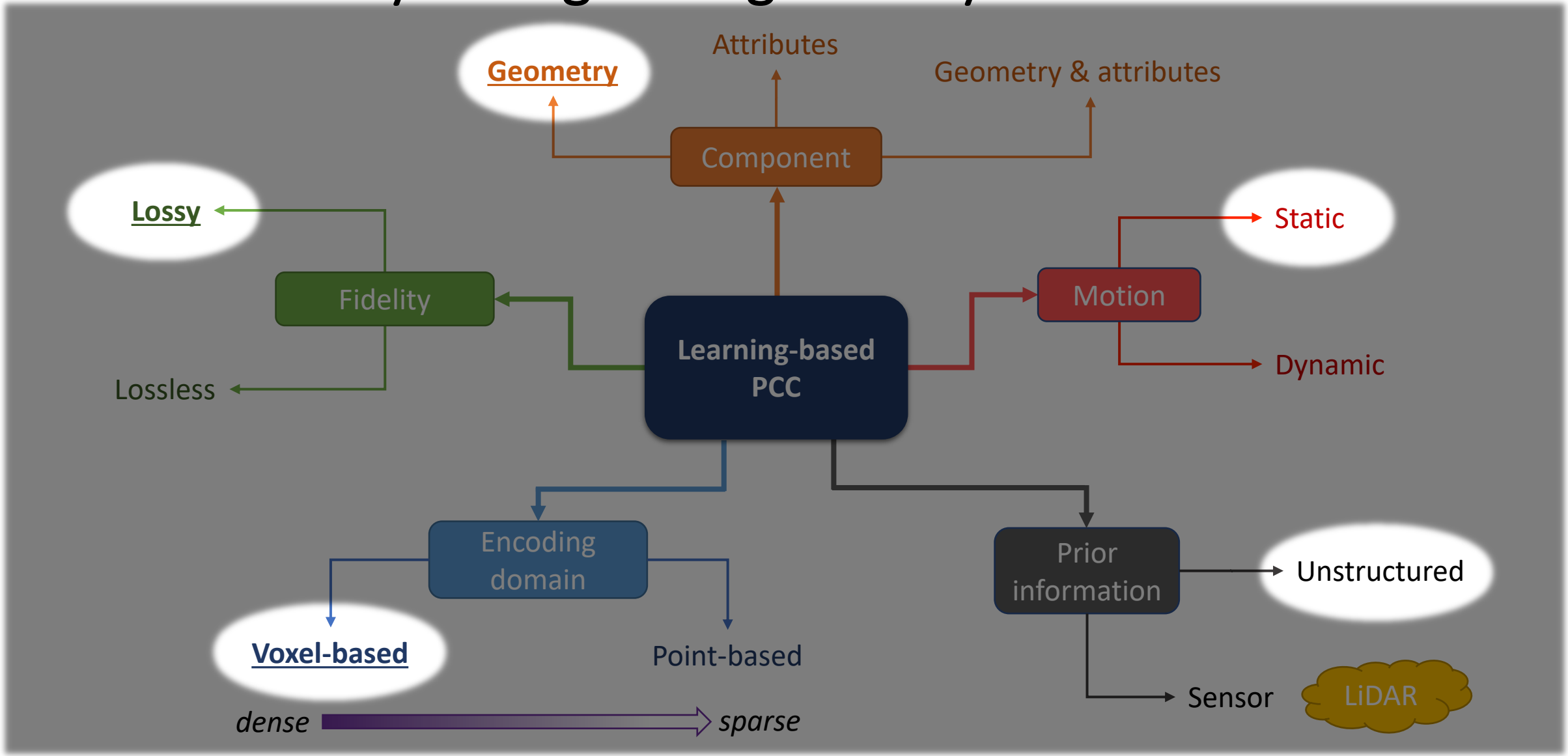
Learning-based Point Cloud Coding

Modeling **complex signal dependencies** using machine learning tools

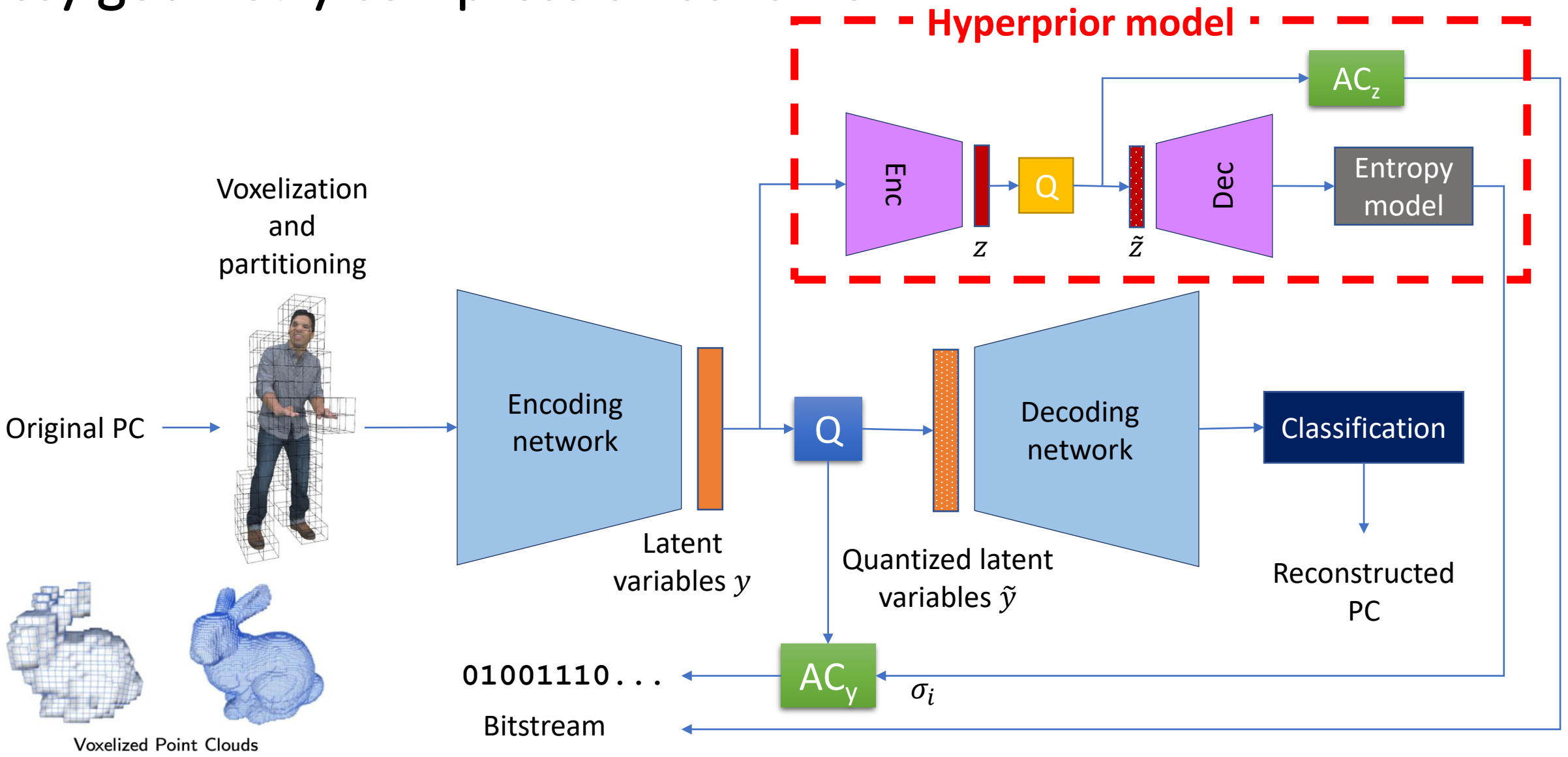
Taxonomy [Quach et al. 2022]



Voxel-based lossy coding of PC geometry



Lossy geometry compression scheme



Qualitative results

- Comparison to basic octree coding (MPEG GPCC anchor)

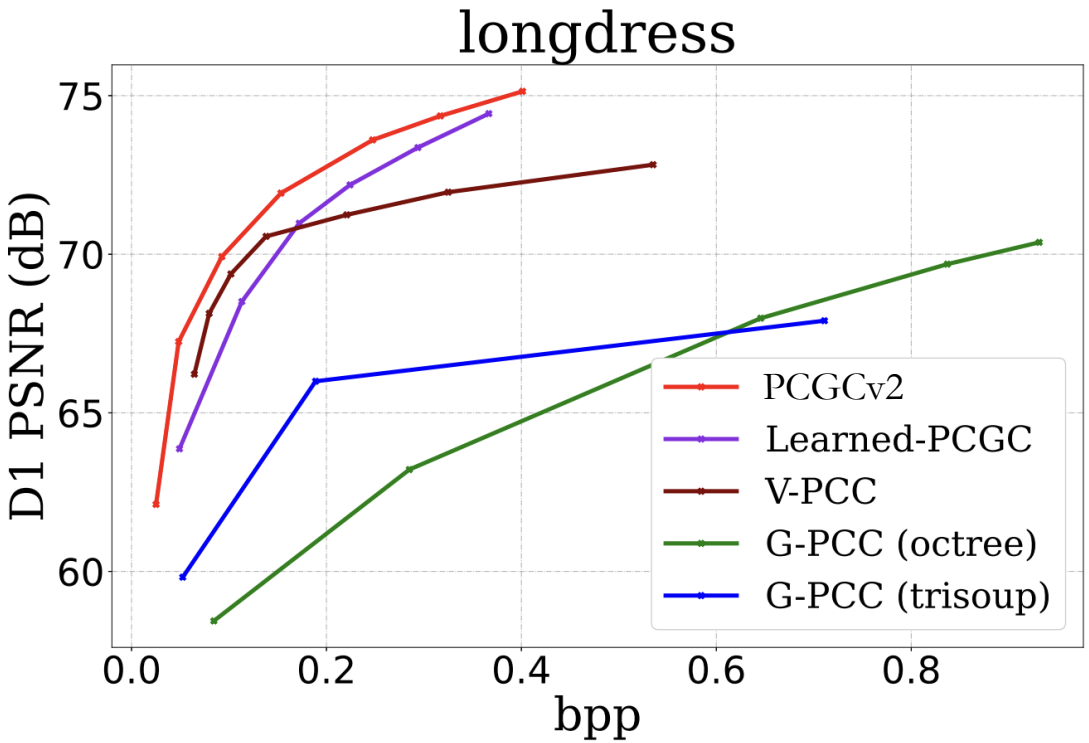
(a) Original

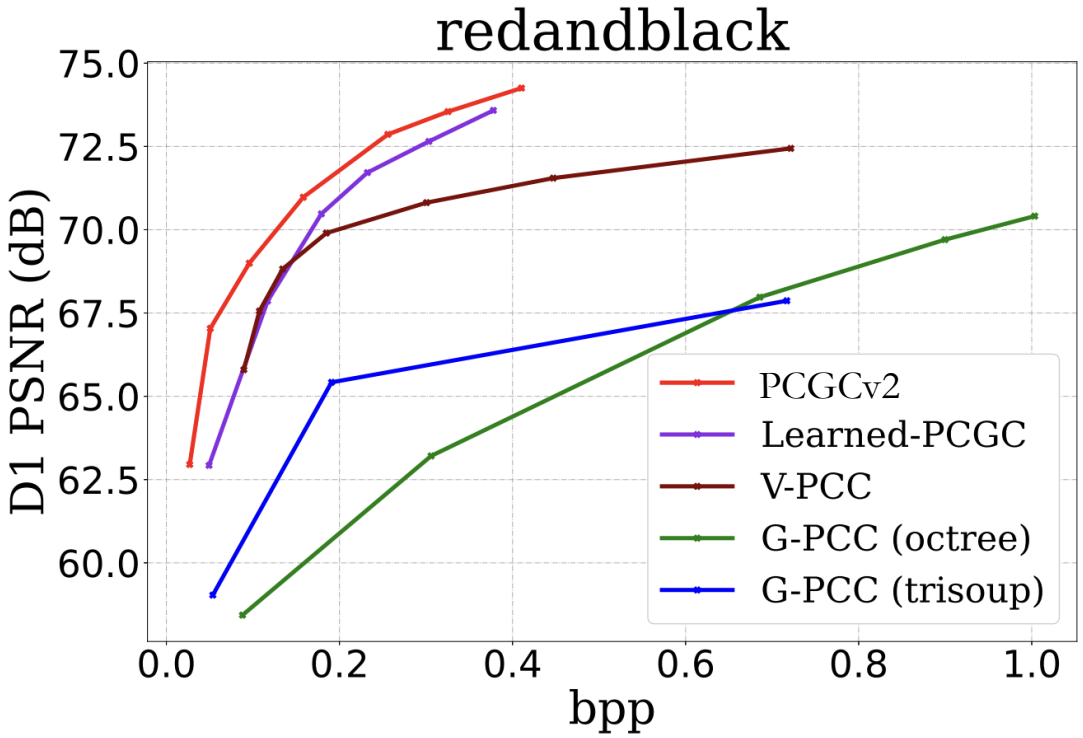
(c) MPEG Anchor

(b) Proposed method

M. Quach, G. Valenzise, F. Dufaux. “Learning Convolutional Transforms for Lossy Point Cloud Geometry Compression.” *ICIP 2019*

Coding performance





Wang, J., Ding, D., Li, Z. and Ma, Z., 2021, March. Multiscale point cloud geometry compression. In *2021 Data Compression Conference (DCC)* (pp. 73-82). IEEE.

What's next?

- Dynamic point clouds
- Attribute coding / joint geometry+attribute
- Call for proposals MPEG AhG on Learning-based PCC in September 2023
- Other representations (Neural Radiance Fields)

Further readings

- Maurice Quach, Jiahao Pang, Dong Tian, Giuseppe Valenzise, Frédéric Dufaux. *Survey on Deep Learning-based Point Cloud Compression.* Frontiers in Signal Processing, Frontiers, 2022, 2

- G. Valenzise, M. Alain, E. Zerman, C. Ozcinar. *Immersive Video Technologies.* Elsevier (September 2022)

- Contact: giuseppe.valenzise@l2s.centralesupelec.fr