Perceptual Quality Assessment of Point Clouds

Evangelos Alexiou

15/03/2023

Outline

1. Introduction
2. Point cloud subjective quality assessment
3. Point cloud objective quality assessment

Outline

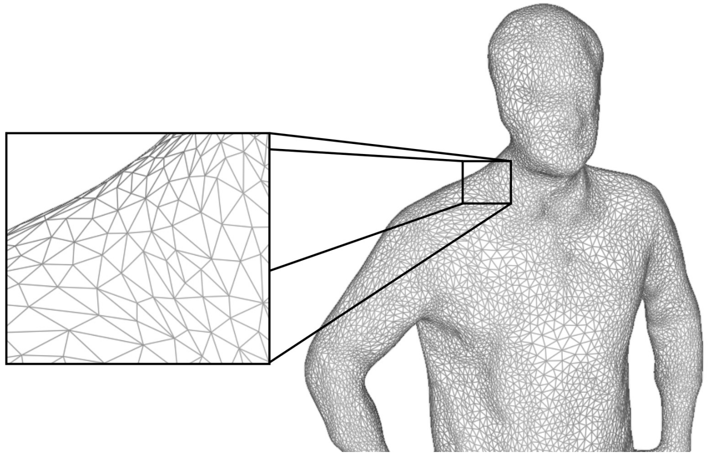
1. Introduction
2. Point cloud subjective quality assessment
3. Point cloud objective quality assessment

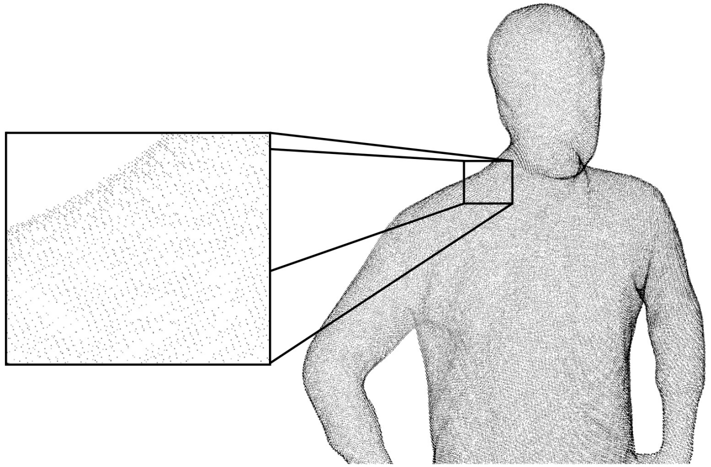
XR and 3D representations

- Emerging applications in **eXtended Reality**
  - **Entertainment**: Social XR, Immersive video, Gaming, Travel experiences
  - **Healthcare**: Consultation, Training, Visualization
  - **Education**: Interactive ways of learning, Engaging experiences
  - **Manufacturing**: Computer-aided design, Development, Maintenance
  - **Sales**: E-commerce, Retail, Real estate
- XR systems make use of **3D contents**
  - Better simulate scenes and objects
  - Allow higher degrees of interactivity
  - Increase user engagement

3D representations

- Point clouds and meshes are the most common types of 3D representations
- *Meshes*
  - **Advantages:**
    - Computer graphics pipeline are tailored to their usage
    - More compact representation of 3D content; watertight; higher visual quality
  - **Disadvantages**
    - Additional complexity due to connectivity information especially for dynamic content
- *Point clouds*
  - **Advantages**:
    - Characterized by their simplicity in acquisition and compression
    - Enable photorealistic reproduction of 3D content
    - Absence of connectivity information leads to lower complexity in manipulating them
  - **Disadvantages:**
    - Scattered representation with holes
    - Not friendly to graphics pipeline





E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

Point cloud imaging

- *Past*
  - Existing type of 3D representation since 80s
  - **Kinect** v1 release in 2010 was a break-through in capturing technologies
  - Research community and industry saw the potentials
  - Standardization efforts in MPEG and JPEG since 2017
- *Present*
  - Wide integration of depth sensors for **acquisition**
  - Wide integration of GPUs for **processing** and **rendering**
  - Release of MPEG V-PCC and G-PCC **compression** standards
- *Future*
  - **Open technical challenges** from acquisition to rendering
  - Ongoing standardization efforts in MPEG and JPEG

Definition of a point cloud:
A set of points in 3D space with x, y, z coordinates defining topology, and optional attributes characterizing properties of the underlying surface (e.g., color, normals, curvature, reflectance)

E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

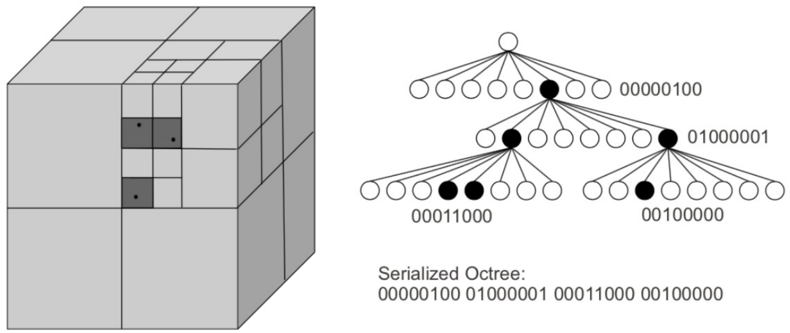
Point cloud acquisition

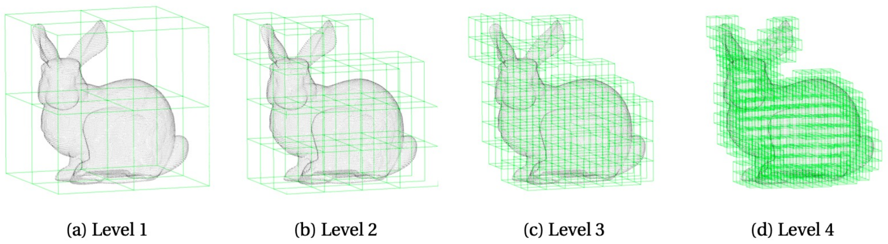


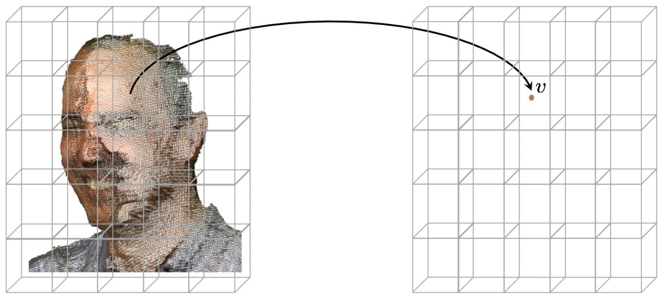

- Different types of sensors
  - Time-of-flight
  - Structured light
  - Laser scanning
  - Photogrammetry
- Different types of acquisition noise
- Reference quality highly varies

Point cloud compression

- Point clouds carry a **vast amount of data**
- Efficient **data structures** are required
  - Octrees are very popular
  - Voxel grids discretize content
- Efficient **compression** solutions are required
  - Lossy methods lead to signal distortions
  - Compression can be applied on geometry and/or color

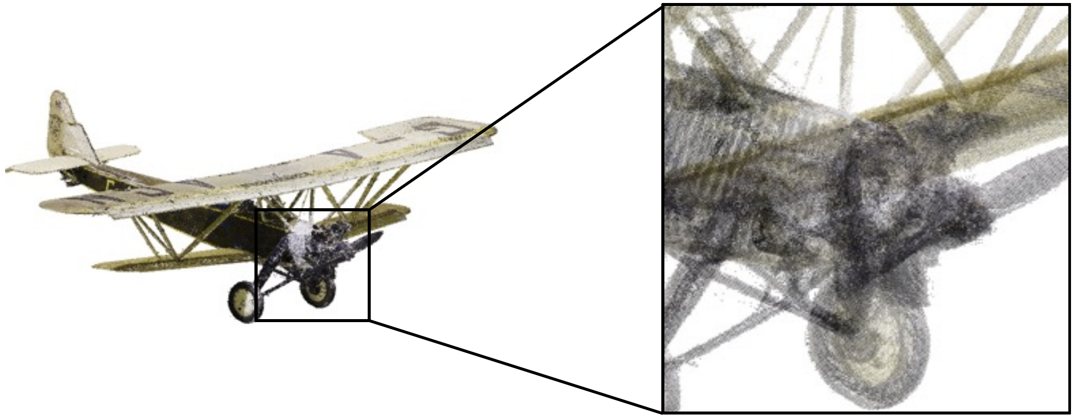






E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

Point cloud rendering



- Rendering points without any size leads to **holes**
- Rendering primitives or shaders are used to assign **spatial attributes**
  - Which primitive?
- How big should be the **point size**?
  - Commonly, fixed point size across a model
  - However, the sparsity of a model might be non-uniform
  - Adaptive point size selection based on local sparsity

E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

Perceptual quality assessment

- Perceptual quality is an important aspect of the overall Quality of Experience (**QoE**) of a user
  - Relates to the level of accuracy a content is reproduced
  - Relates to human visual model and perceptual characteristics
- Perceptual quality assessment is the process of **determining the visual quality** of a content
- Perceptual quality can be affected by **different types of distortions** introduced during processing, compression, transmission and rendering
- Perceptual quality can be assessed using:
  - Subjective methods
  - Objective methods

Definition of Quality of Experience (QoE):
The degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state.

Outline

1. Introduction
2. Point cloud subjective quality assessment
3. Point cloud objective quality assessment

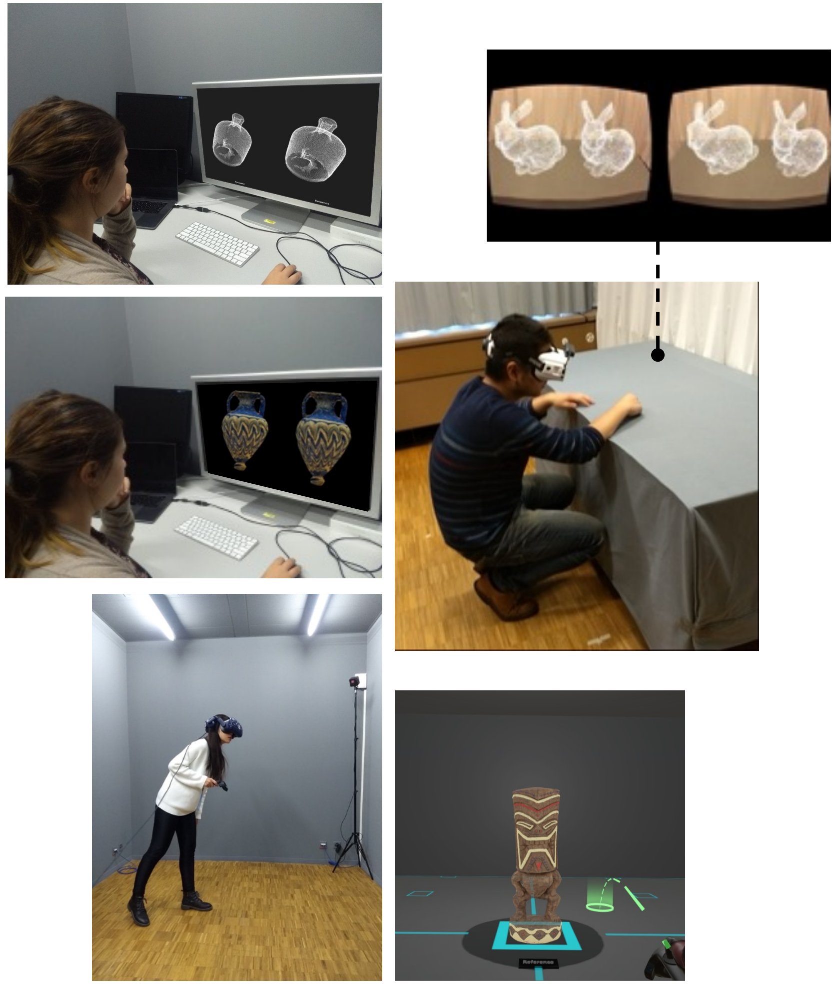
Subjective quality assessment

- Methods that rely on **human subjects** submitting their opinion regarding the visual quality of a content they consume
- **Advantages**:
  - Provide ground truth ratings by end consumers
- **Disadvantages:**
  - Costly and time-consuming
- **Challenges:**
  - Rigorousness and reproducibility
- Recommendations from ITU and experts groups
  - ITU-R BT500.13 [1]
  - ITU-T P.910 [2]
  - Provide guidelines regarding the design of user studies: test methods, experimental setups, lighting conditions, monitors, distance of viewer, etc

[1] ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," International Telecommunications Union, Jan 2012
[2] ITU-T P.910, "Subjective video quality assessment methods for multimedia applications," International Telecommunications Union, Apr 2008

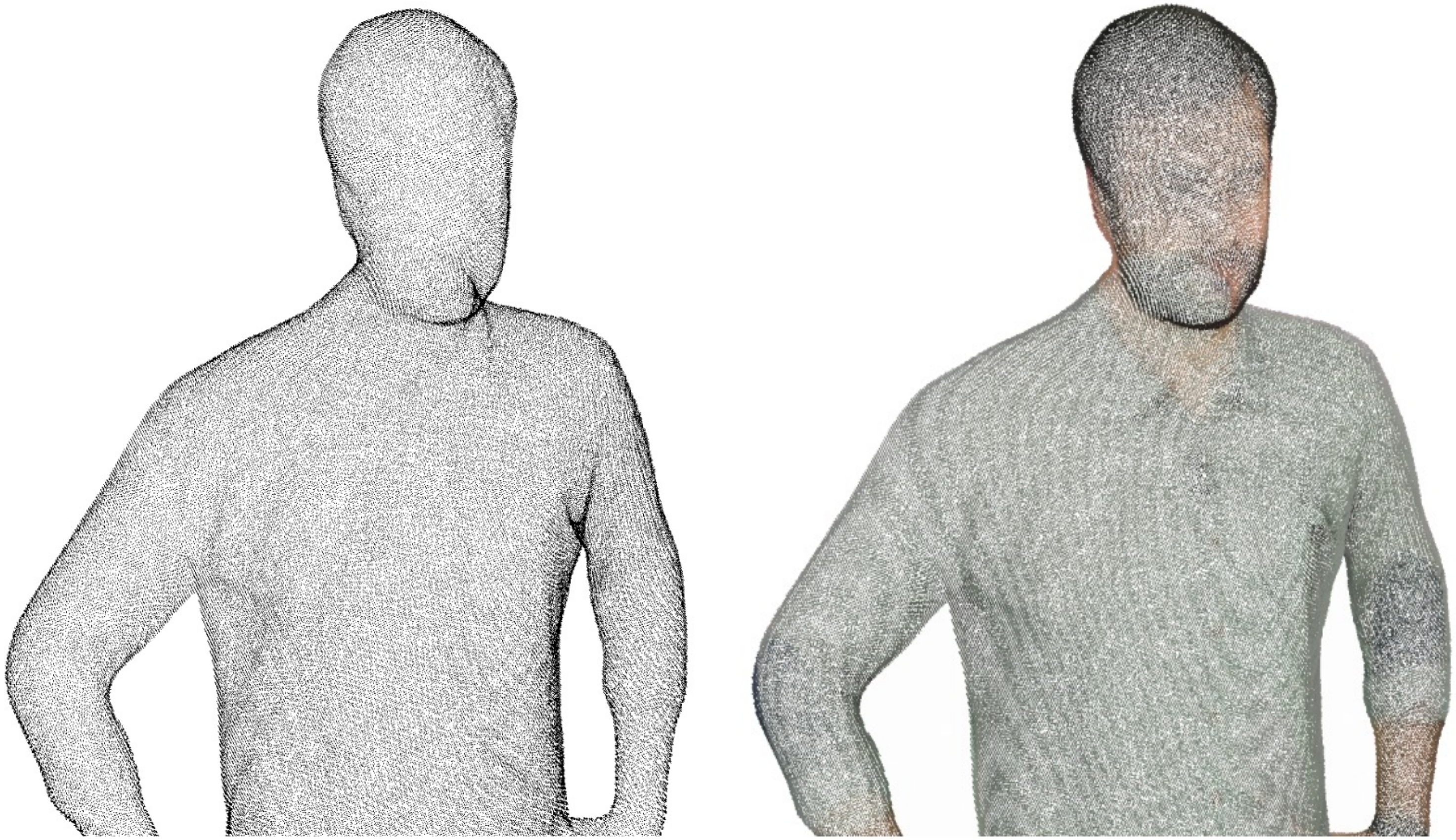
Subjective quality assessment for point clouds

- **Challenges**:
  - Inspection with 6DoF leads to partial consumption
  - Content variations
    - Reference quality – resolution, noise
    - Small objects to Large scenes
- Important aspects for the design of subjective user studies
  - **Temporal variation**: Static, Dynamic
  - **Attributes**: Colorless, Colored
  - **Mode of inspection**: Passive, Interactive
  - **Display devices**: 2D/3D monitors, HMDs, Smartphones
  - **Rendering**: Point size, Mesh conversion, Lighting, Background
  - **Methodology**: Single stim., Double stim., Pairwise comparison
  - **Distortion type**: Noise, compression, transmission error



E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

Effect of ***attributes***

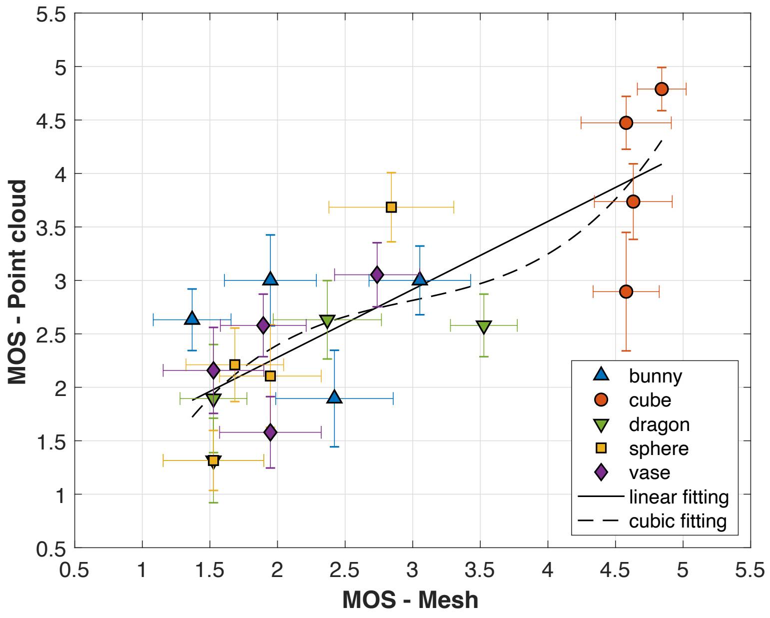


E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

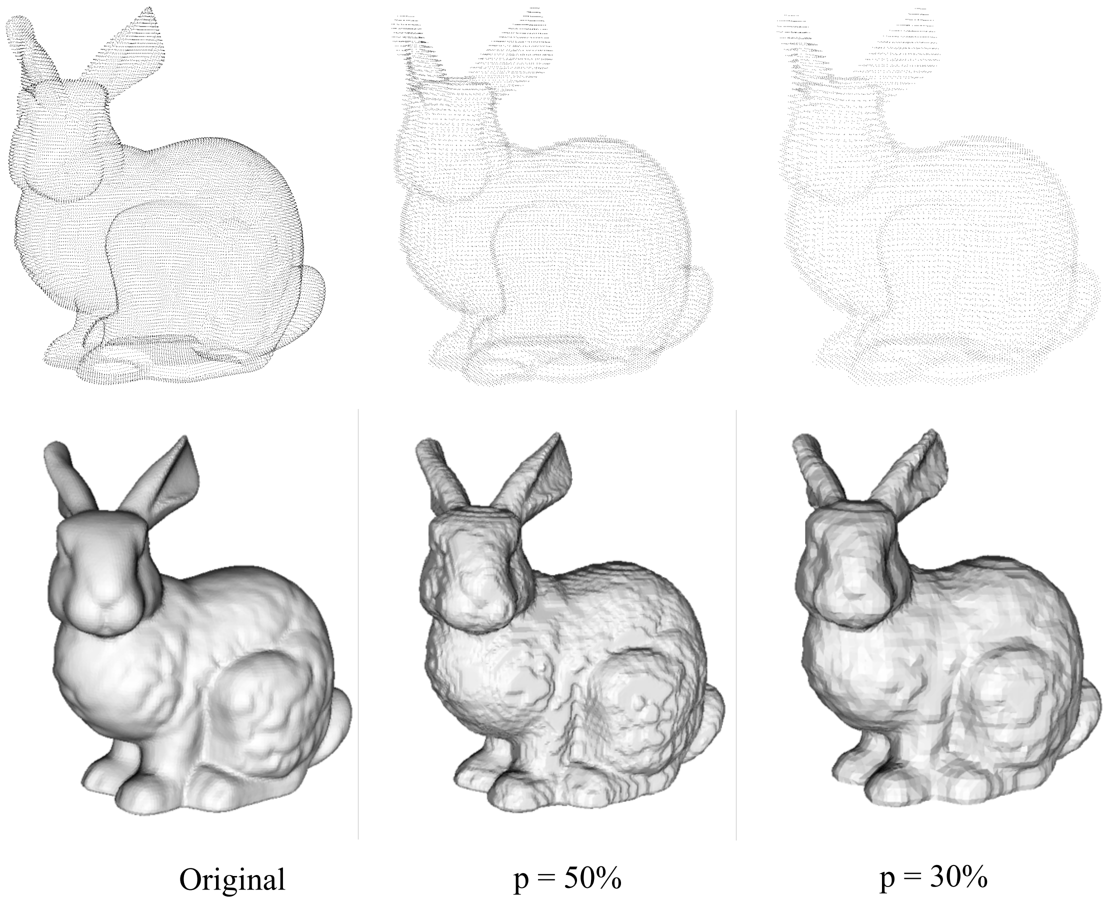
Effect of ***point size (rendering parameter)***

E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

Effect of *rendering as mesh*

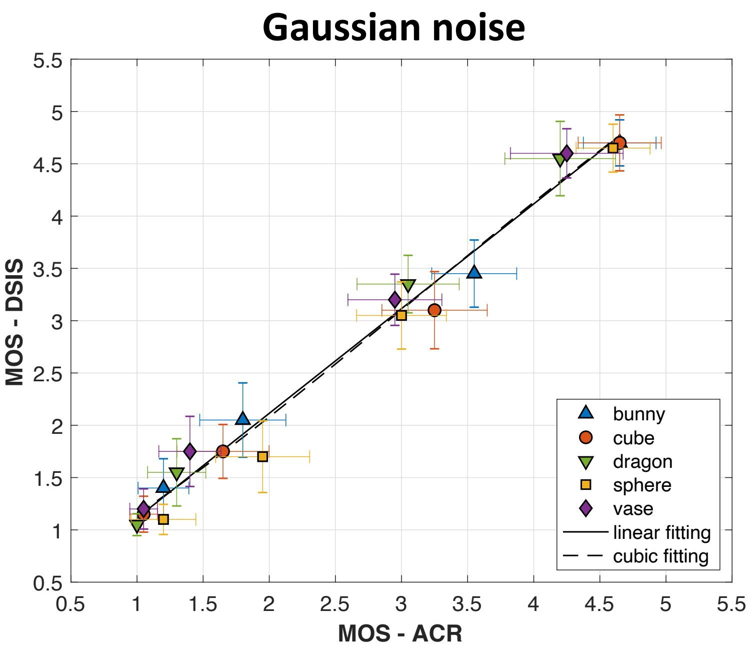


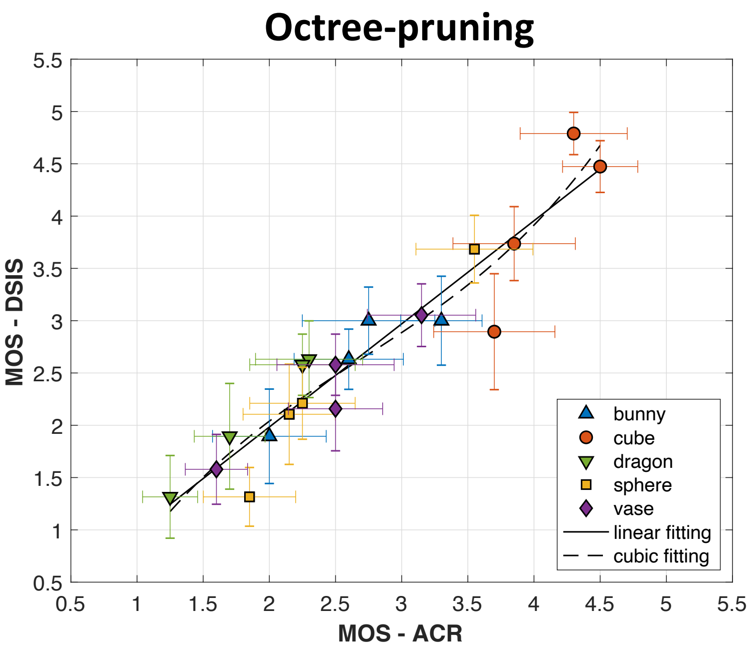
- **Different artifacts**
- **Depend on the reconstruction**

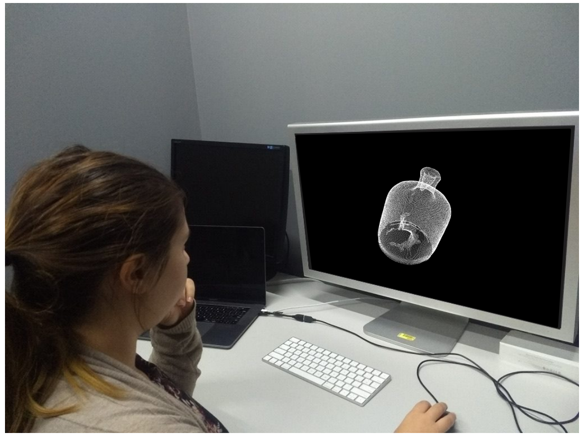


E. Alexiou, T. Ebrahimi, et al. "Point cloud subjective evaluation methodology based on 2D rendering." In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, 2018

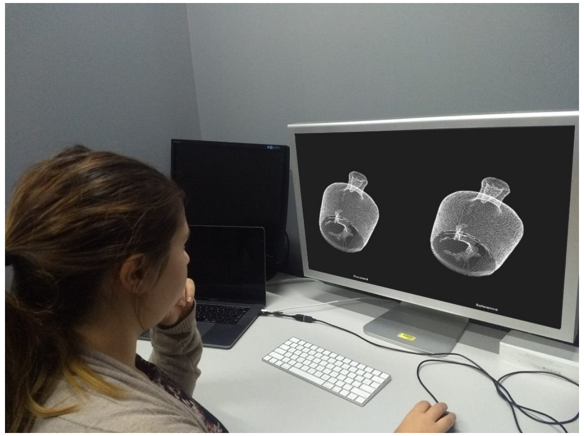
Effect of *test method*







Desktop - ACR



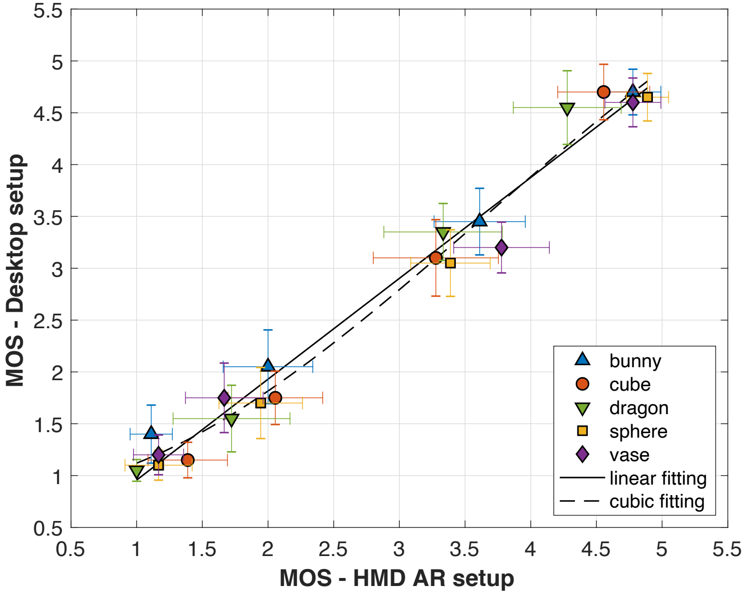
Desktop - DSIS

- **DSIS is more consistent for identification of impairments**

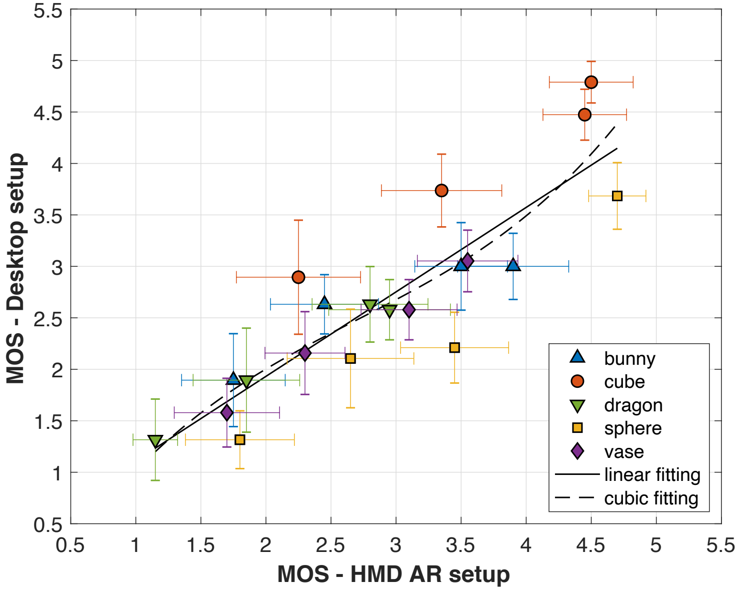
E. Alexiou, and T. Ebrahimi. "On the performance of metrics to predict quality in point cloud representations." In *Applications of digital image processing XL*, vol. 10396. SPIE, 2017

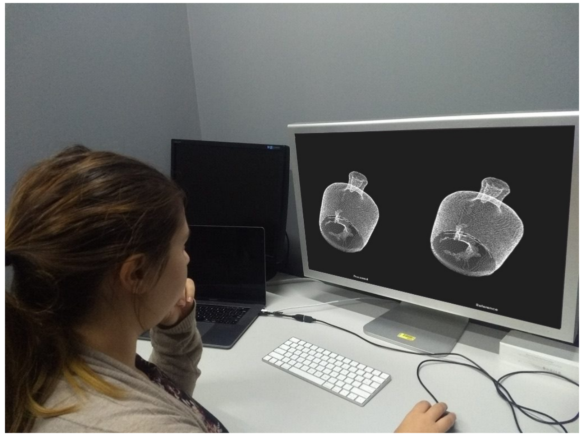
Effect of *display devices*

Gaussian noise

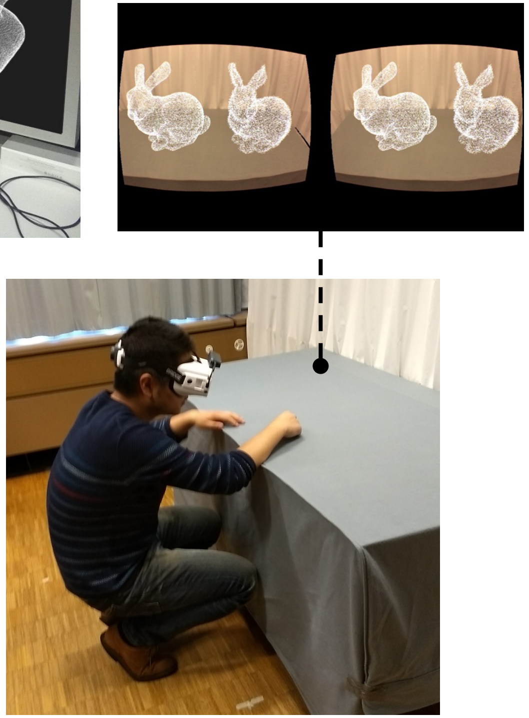


Octree-pruning





Desktop



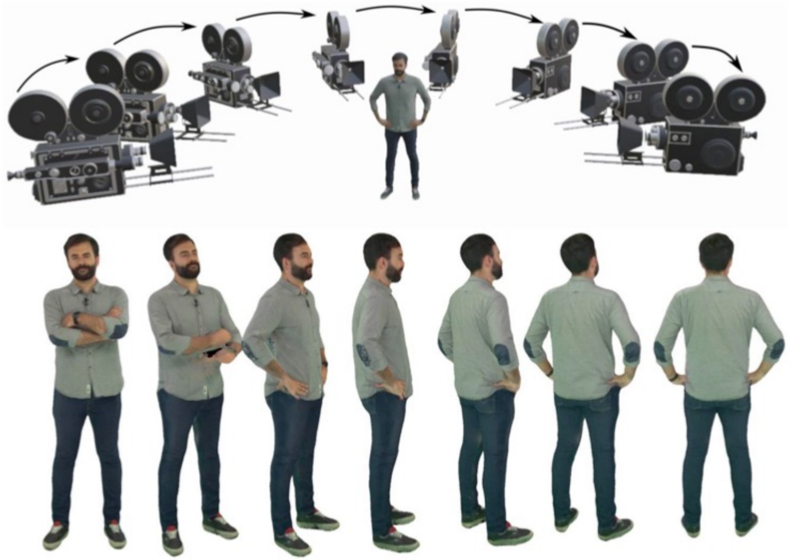
HMD AR

- **Differences depending on degradation type**

E. Alexiou, and T. Ebrahimi. "Impact of visualisation strategy for subjective quality assessment of point clouds." In *2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2018

Passive *mode of inspection*

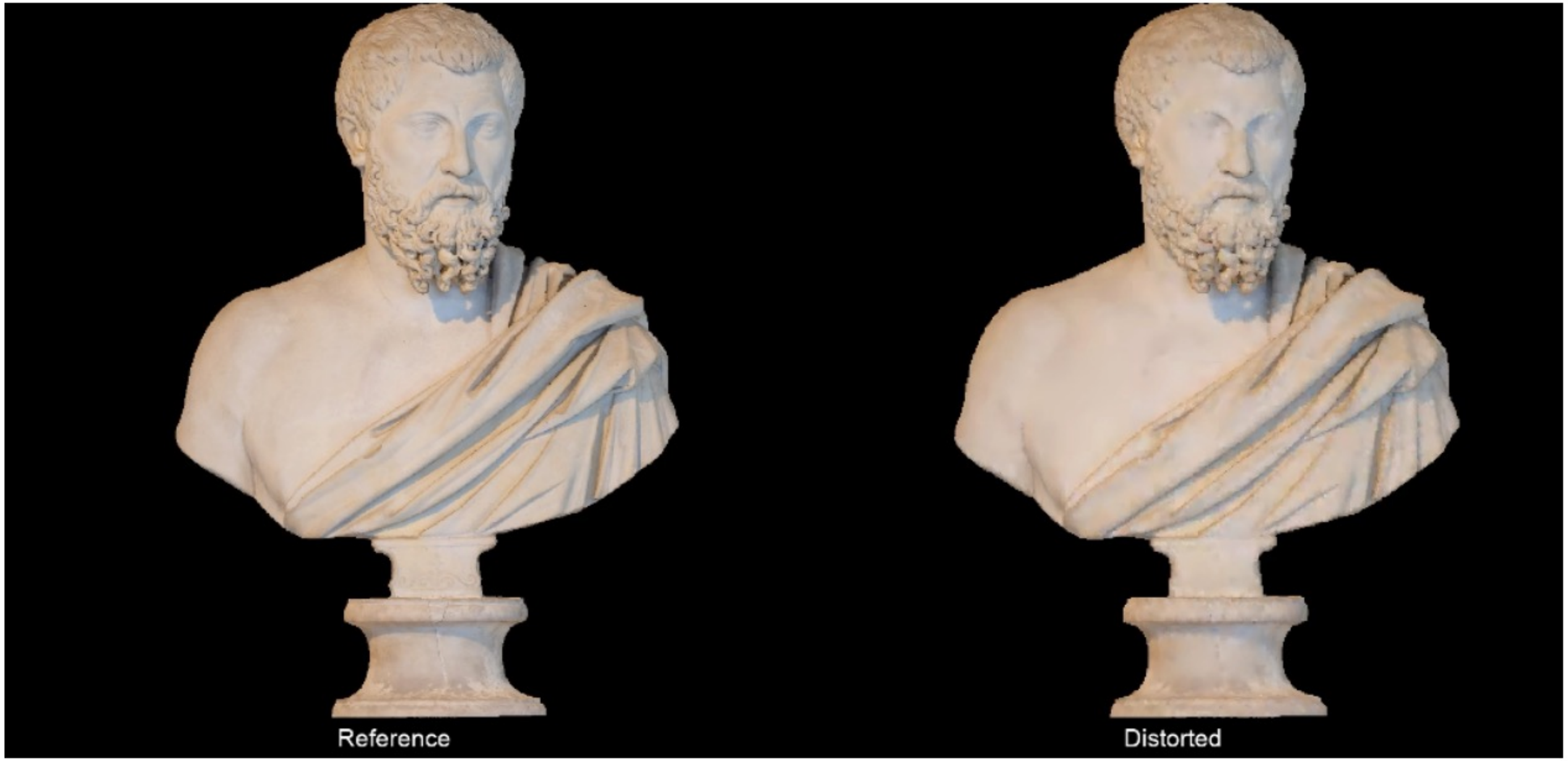
- Produce a **video** capturing the model from a pre-defined camera trajectory, which is shown to subjects
- **Advantages**:
  - Every stimulus is inspected in the exact same way by all participants – reproducibility and elimination of inter-viewer variations due to interactivity
  - Less complex setups – easier for crowdsourcing
  - Absence of external biases due to novelty effect



Disadvantages:
  - Less realistic type of consumption

E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

Passive desktop setup

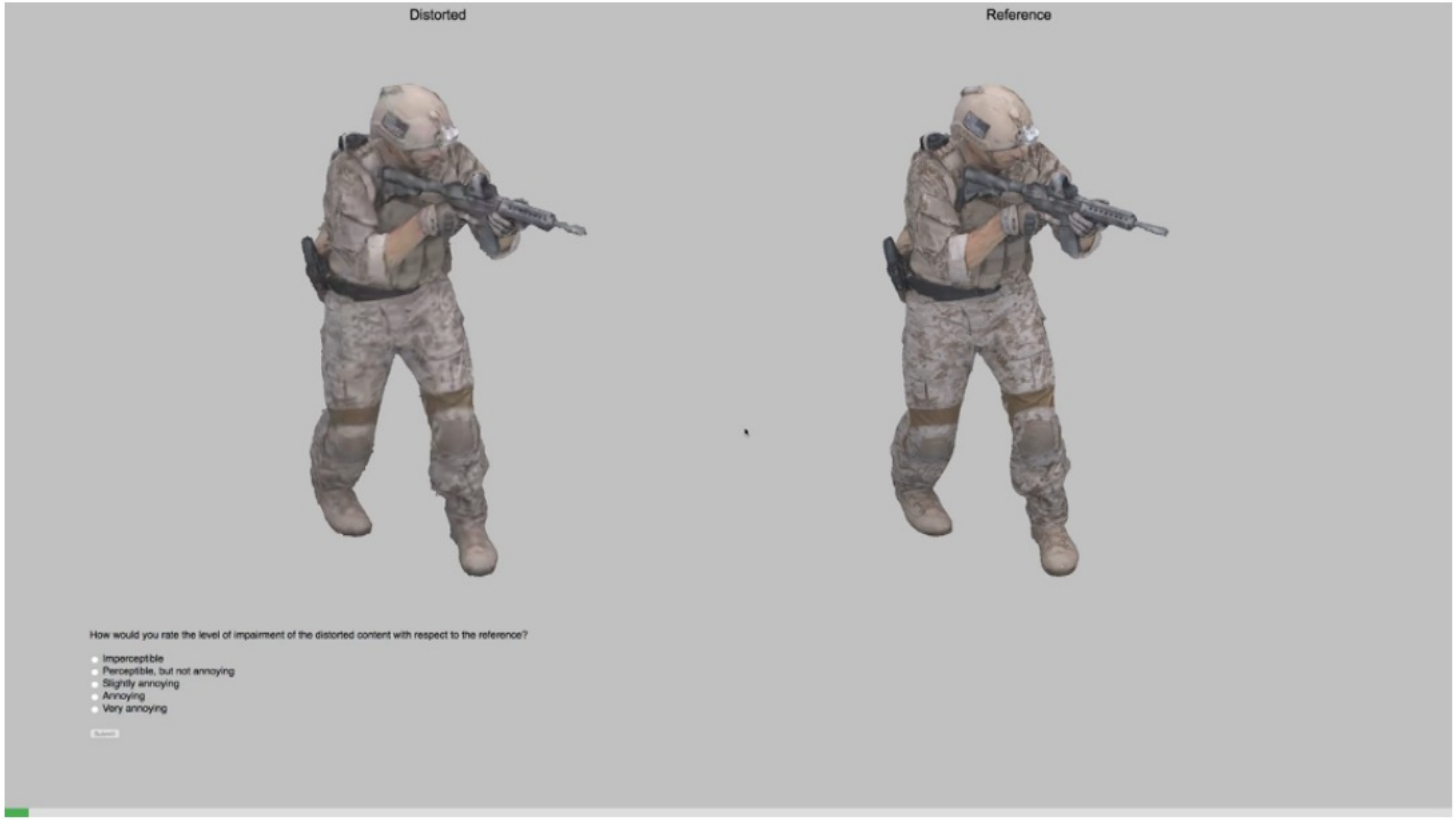


D. Lazzarotto, E. Alexiou, and T. Ebrahimi. "Benchmarking of objective quality metrics for point cloud compression." In *2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP)*, 2021

Interactive *mode of inspection*

- Use of 3D rendering engines to display the contents in a virtual scene and allow users to **naturally interact** with those
  - Mouse cursor or keyboard in desktop setups
  - Physical movements or controllers with head-mounted displays
- **Advantages**:
  - Scores inherently contain the effect of natural user interaction
  - Allow to examine interactions between interactivity patterns and perception of quality
- **Disadvantages**:
  - Individual experience per user
  - More complex setups

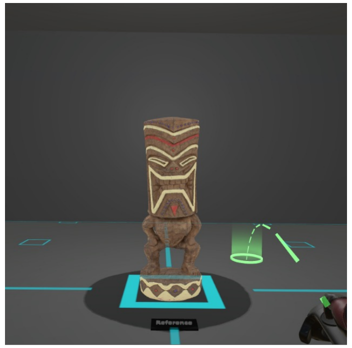
Interactive desktop setup [JS - webGL]

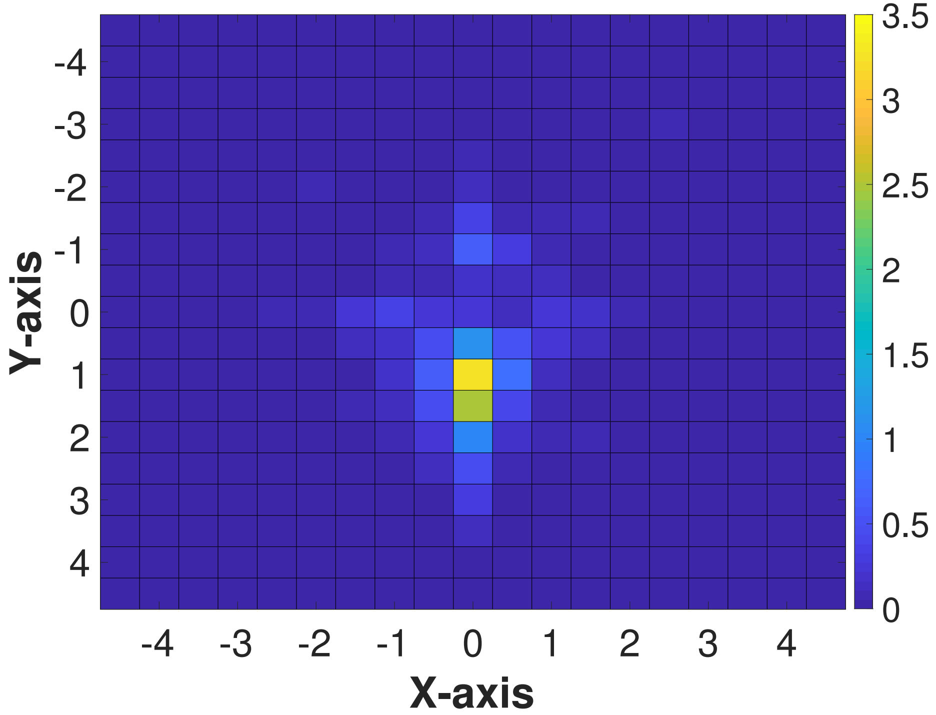


E. Alexiou, I. Viola, T. M. Borges, T. A. Fonseca, R. L. De Queiroz, and T. Ebrahimi. "A comprehensive study of the rate-distortion performance in MPEG point cloud compression." *APSIPA Transactions on Signal and Information Processing* 8 (2019): e27

From desktop to VR settings

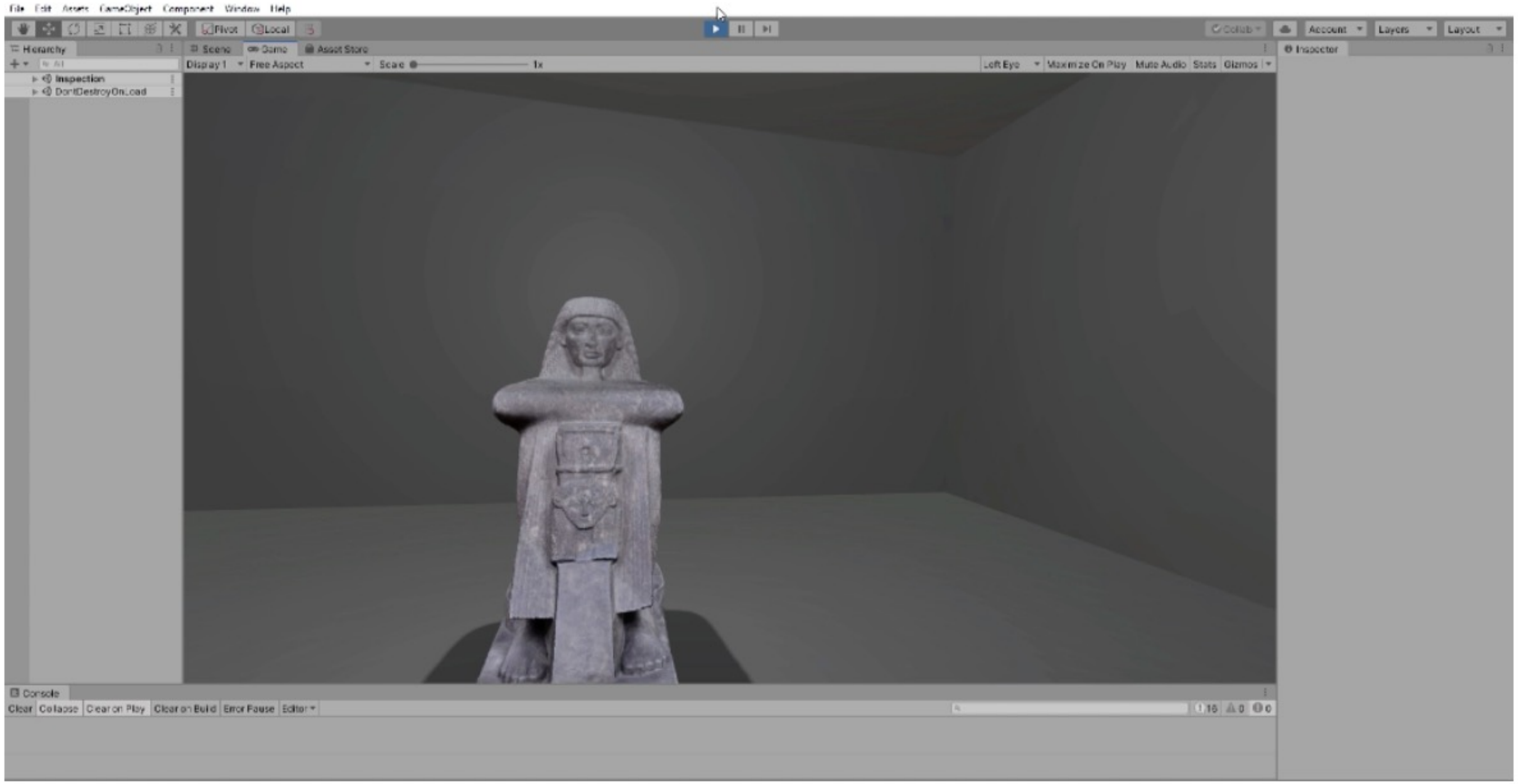
- VR settings bring immersiveness
- Entirely **controlled** and **reproducible** environment
- Allow to study navigation and behavioural patterns





E. Alexiou, N. Yang, and T. Ebrahimi. "PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality." In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020

Interactive VR setup with 6DoF [Unity]

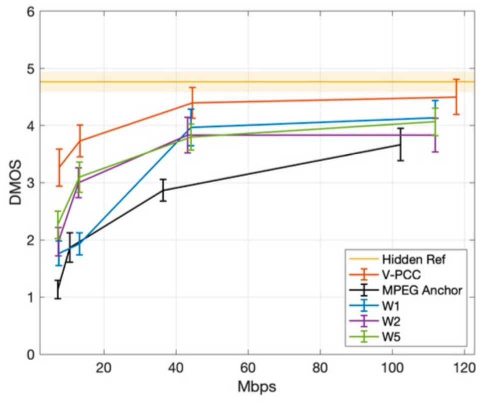


E. Alexiou, N. Yang, and T. Ebrahimi. "PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality." In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 2020

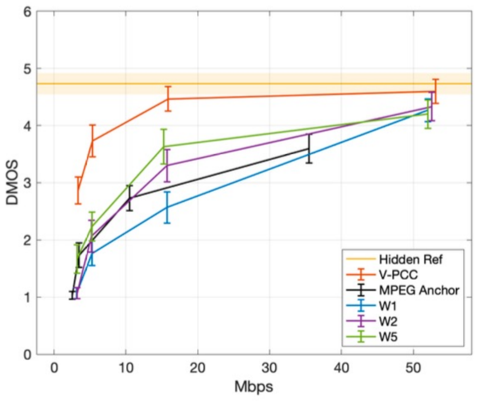
Adaptive tiling

(a) *Track viewport position* (b) *Compute quality levels* (c) *Request tile representations* (d) *Render synchronized tiles* (e) *Playout*

- Adaptive tiling to reduce bitrate
  - Users consume part of the content
- Utility functions for bit-rate allocation per tile
  - Consider user’s position
  - Different strategies (e.g., Greedy, Uniform, Weighted hybrid)
- Subjective evaluation in VR with 6DoF
- Results show substantial bit-rate savings vs nontiled



(a) *Longdress*

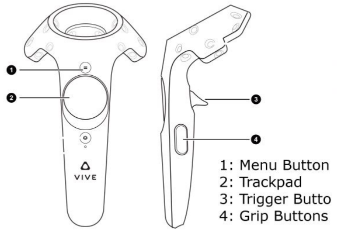


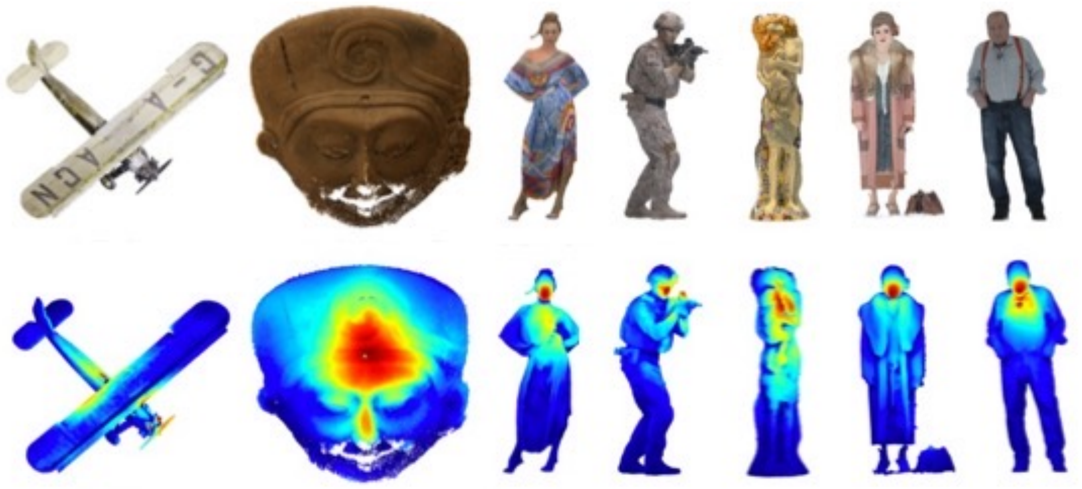
(b) *Loot*

S. Subramanyam, et al. "Subjective QoE Evaluation of User-Centered Adaptive Streaming of Dynamic Point Clouds." In *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*, 2022

From quality assessment to visual saliency

- Where people look in VR?
- 6DoF inspection
- Task dependent protocol
- Non distracting scene
- Real-life models
- Per session profiling
- Weighted fixations





E. Alexiou, P. Xu, and T. Ebrahimi. "Towards modelling of visual saliency in point clouds for immersive applications." In *2019 IEEE International Conference on Image Processing (ICIP)*, 2019

Outline

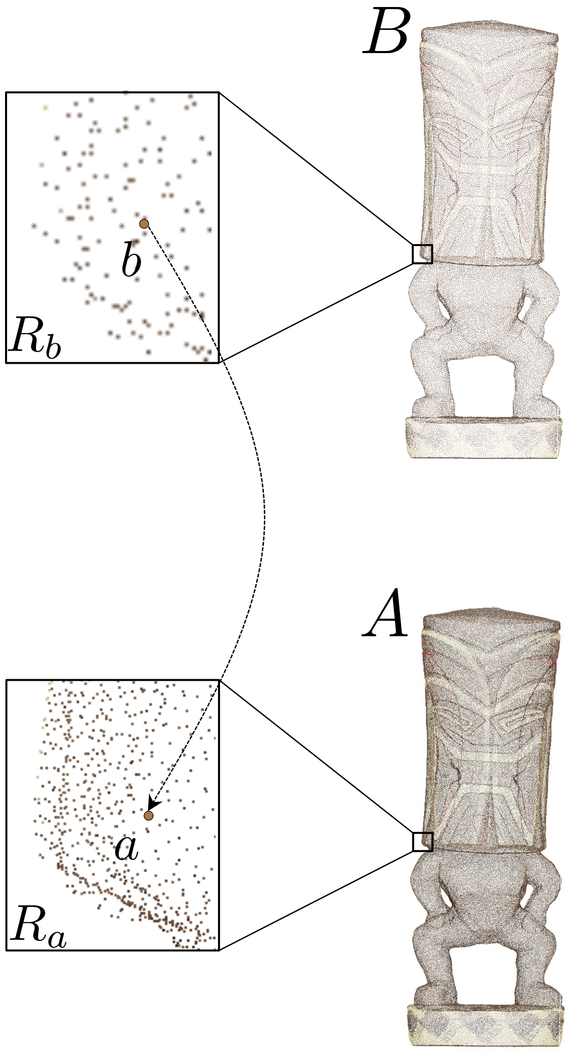
1. Introduction
2. Point cloud subjective quality assessment
3. Point cloud objective quality assessment

Objective quality assessment

- Methods that rely on **computational models** that predict the visual quality of a content
- **Advantages:**
  - Automatically executed for closed-loop optimizations in compression and transmission systems
- **Disadvantages:**
  - Reliability, as they do not always correlate well with subjective opinions
- **Challenges:**
  - Low complexity and high prediction accuracy across different types of content and distortions
- *Metrics may focus on the quantification of the signal error, or on quantification of perceptual degradations (perceptual metrics)*
- Metrics can be distinguished based on their requirement for reference data
  - Full-reference | Reduced-reference | No-reference
- Point cloud metrics can be clustered based on the domain their applied as:
  - Model-based | Image-based

Model-based quality metrics

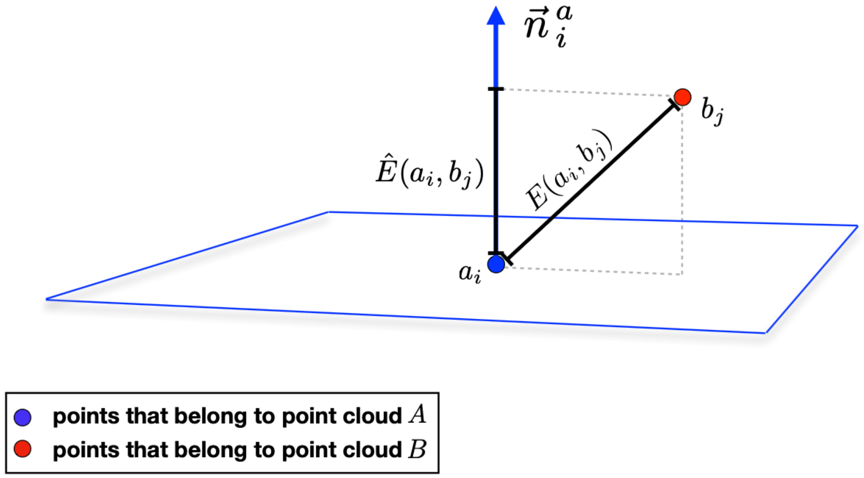
- Operate on 3D point cloud space
  - Measurements from point cloud geometry or color
- **Advantages:**
  - Applied on point cloud data
  - Rendering-agnostic
- **Disadvantages:**
  - Depend on configurations for features estimation
  - Multimodal nature of data with adhoc fusion

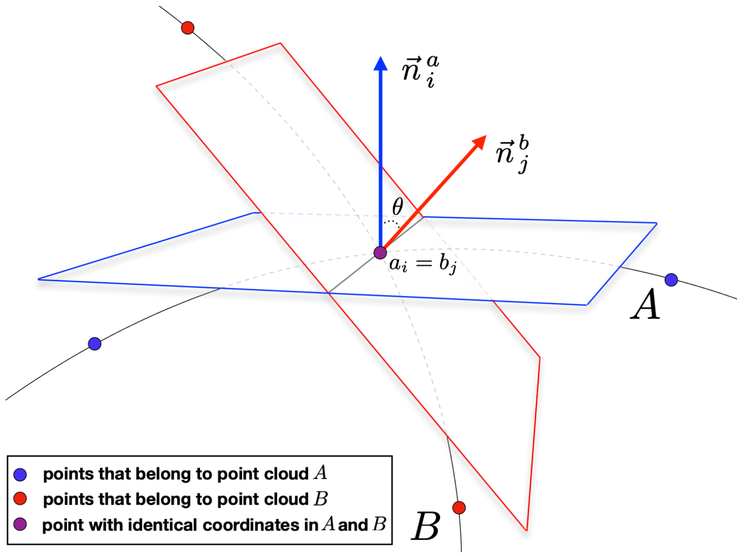


E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

Model-based quality metrics

- Early metrics are **Full-reference**, and based on **simple** geometric or color **distances**
  - Less perceptually relevant – similar to MSE/PSNR in 2D images
  - Full-reference metrics require a correspondence
- Point-to-point
  - Measures displacement of distorted samples from their reference positions
- Point-to-plane
  - Measures deviation of distorted sample from its local linear surface approximation
- Plane-to-plane
  - Measures angular similarity between local linear surface approximations

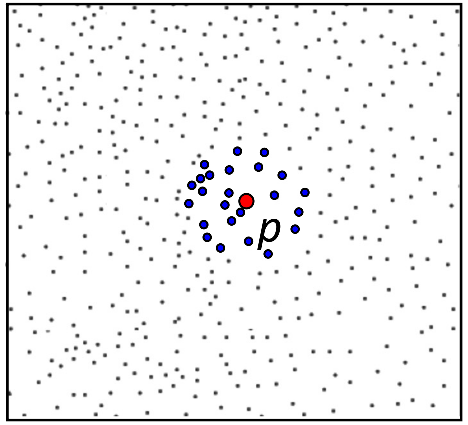




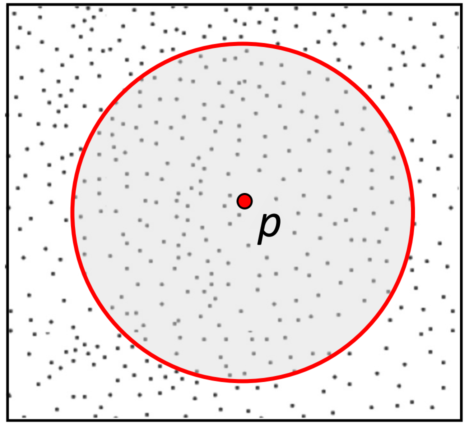
E. Alexiou, and T. Ebrahimi. "Point cloud quality assessment metric based on angular similarity." In *2018 IEEE International Conference on Multimedia and Expo (ICME)*, 2018

Point cloud features

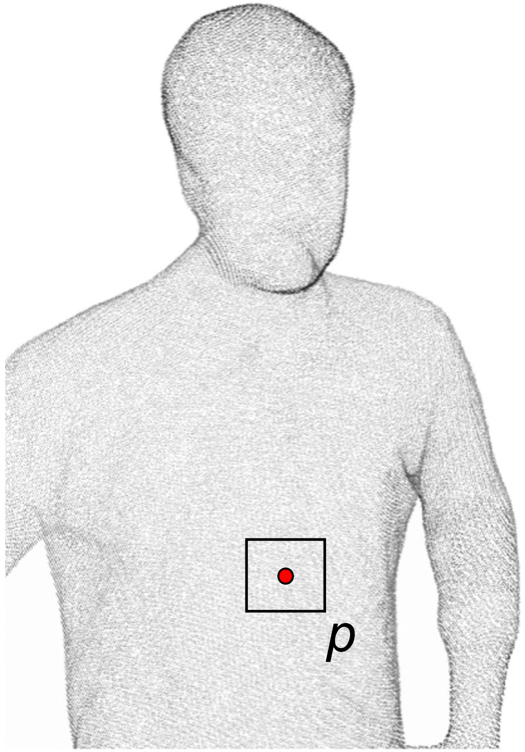
- Features exctracted from **neighborhoods** capturing **local properties** in geometry or color
  - Interpretability?
- How do you compute neighborhoods?
  - Properties of neighborhood formulation algorithms



k-nearest neighbors

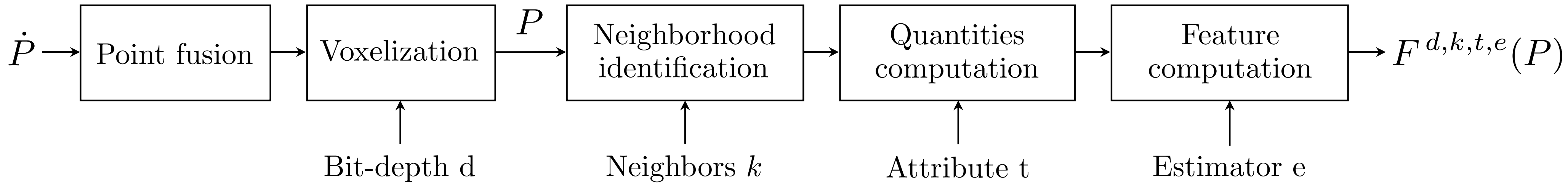


Spherical neighborhood of radius *r*



E. Alexiou, Y. Nehmé, E. Zerman, I. Viola, G. Lavoué, A. Ak, A. Smolic, P. Le Callet, and P. Cesar. "Subjective and objective quality assessment for volumetric video." In *Immersive Video Technologies*, Academic Press, 2023

PointSSIM



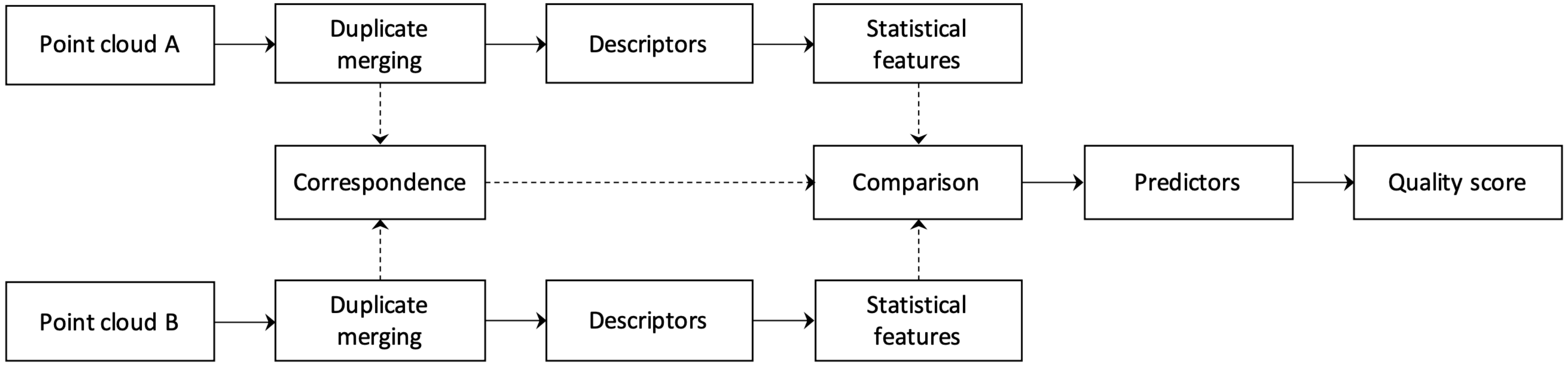
- Extend operation logic of **SSIM** to point cloud data
- Difference between features capturing local changes
- Applied on **location**, **normal**, **curvature** and **luminance** attributes
- Voxelization to eliminate cross-content density variations

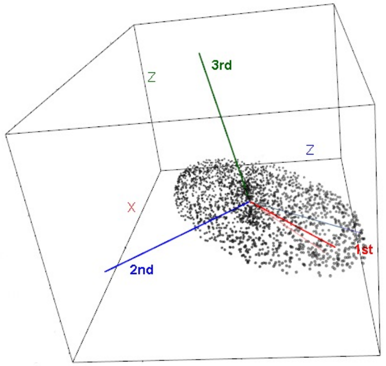
$$E_{B,A}(b) = \frac{\left|F_B(b) - F_A(a)\right|}{\max\left\{\left|F_B(b)\right|, \left|F_A(a)\right|\right\} + \varepsilon}$$

$$\texttt{PointSSIM}_{B,A} = \frac{1}{N_b}\sum_{b=1}^{N_b}\left(1 - E_{B,A}(b)\right)^k$$

E. Alexiou, and T. Ebrahimi. "Towards a point cloud structural similarity metric." In *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2020

PointPCA





- Perceptual capabilities of descriptors computed without requiring higher order polynomial fitting
- Geometric descriptors as location **dispersion**, **roughness** and **direction of in data dispersion** after PCA on spherical regions
- Textural descriptors as **luminance** and **dispersion** measurements
- Statistical features as 1st and 2nd order moments of descriptors on *k-nearest neighbors*
- Combination via Random Forests

	Descriptor	Definition
Geometric	Eigenvalues	$d_v^g = \lambda_v^g$, with $v \in \{1,2,3\}$
	Sum of eigenvalues	$d_4^g = \sum_v \lambda_v^g$
	Linearity	$d_5^g = (\lambda_1^g - \lambda_2^g)/\lambda_1^g$
	Planarity	$d_6^g = (\lambda_2^g - \lambda_3^g)/\lambda_1^g$
	Sphericity	$d_7^g = \lambda_3^g/\lambda_1^g$
	Anisotropy	$d_8^g = (\lambda_1^g - \lambda_3^g)/\lambda_1^g$
	Omnivariance	$d_9^g = \sqrt[3]{\lambda_1^g \cdot \lambda_2^g \cdot \lambda_3^g}$
	Eigenentropy	$d_{10}^g = -\sum_v \lambda_v^g \cdot \ln(\lambda_v^g)$
	Surface variation	$d_{11}^g = \lambda_3^g / \sum_v \lambda_v^g$
	Roughness	$d_{12}^g = \lvert(\mathbf{p}_i - \bar{\mathbf{p}}_i) \cdot \mathbf{e}_3^g\rvert$
	Parallelity$_x$	$d_{13}^g = 1 - \lvert\mathbf{u}_x \cdot \mathbf{e}_3^g\rvert$
	Parallelity$_y$	$d_{14}^g = 1 - \lvert\mathbf{u}_y \cdot \mathbf{e}_3^g\rvert$
	Parallelity$_z$	$d_{15}^g = 1 - \lvert\mathbf{u}_z \cdot \mathbf{e}_3^g\rvert$
Textural	Luminance	$d_1^t = \mathrm{Y}$
	Eigenvalues	$d_{\tilde{v}}^t = \lambda_{\tilde{v}}^t$, with $\tilde{v} \in \{1,2,3\}$
	Sum of eigenvalues	$d_5^t = \sum_{\tilde{v}} \lambda_{\tilde{v}}^t$
	Eigenentropy	$d_6^t = -\sum_{\tilde{v}} \lambda_{\tilde{v}}^t \cdot \ln(\lambda_{\tilde{v}}^t)$

E. Alexiou, X. Zhou, I. Viola, and P. Cesar. "PointPCA: Point Cloud Objective Quality Assessment Using PCA-Based Descriptors." *arXiv,* 2023

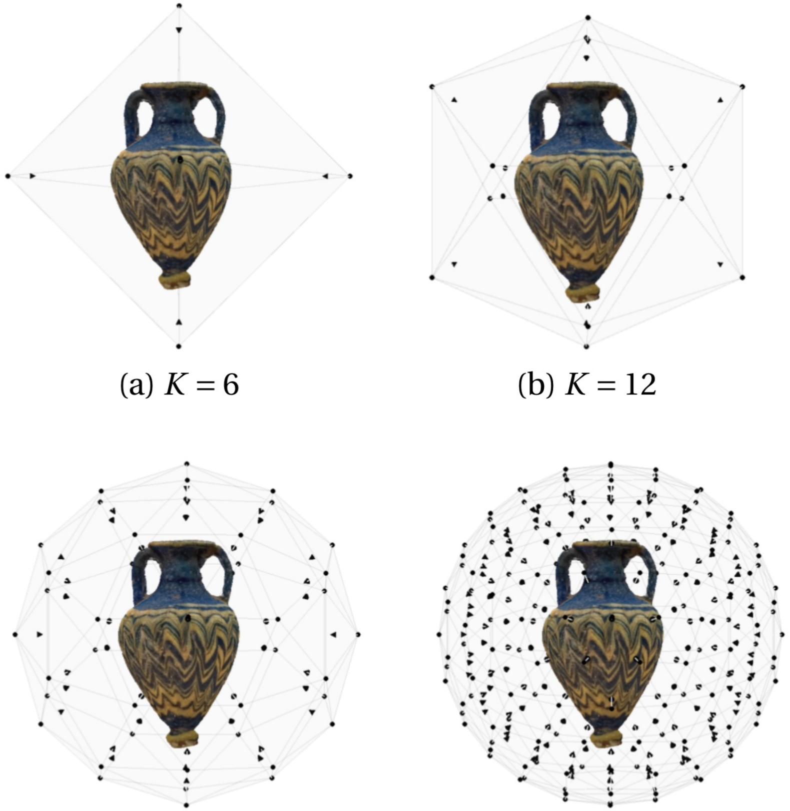
Image-based quality metrics

- Operate on 2D image space
  - Project onto planar surfaces
- **Advantages**:
  - Holistic capture for geometry, color, and rendering
  - Exploit existing sophisticated 2D metrics
- **Disadvantages:**
  - View-dependent
  - Rendering-dependent

E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

Effect of *number of views*

- Capture projections after rendering
- Apply 2D imaging algorithms on model views
  - PSNR, SSIM, **MS-SSIM**, VIFp
  - Average scores across model views
- Camera arrangements
  - K = {1, 6, 12, 42, 162}
- Including or excluding background?
  - **Union of foregrounds of original and distorted**



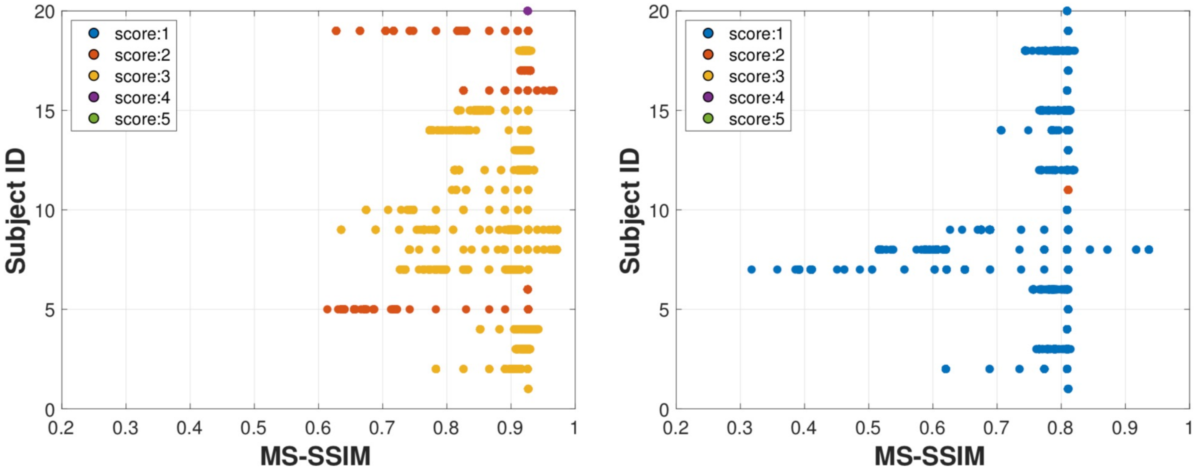
(a) $K = 6$ (b) $K = 12$ (c) $K = 42$ (d) $K = 162$

MS-SSIM	Inanimate objects				Human figures			
	PLCC	SROCC	RMSE	OR	PLCC	SROCC	RMSE	OR
$K = 1$	0.951	0.944	0.373	**0.519**	**0.952**	**0.935**	**0.279**	**0.519**
$K = 6$	**0.955**	0.944	**0.359**	0.556	0.933	0.927	0.328	**0.519**
$K = 12$	0.949	0.944	0.381	**0.519**	0.926	0.920	0.344	**0.519**
$K = 42$	0.949	**0.945**	0.383	**0.519**	0.926	0.915	0.345	0.556
$K = 162$	0.949	**0.945**	0.384	**0.519**	0.925	0.915	0.347	**0.519**

E. Alexiou and T. Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging", in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019

Exploiting *user views*

- Navigation tracks
  - Playback of all users interactions while evaluating stimuli
- **Advantages:**
  - Representative of the experience
- **Disadvantages:**
  - Time consuming
  - Substantial fluctuations between objective scores
- Results show that a fixed number of views leads to better results



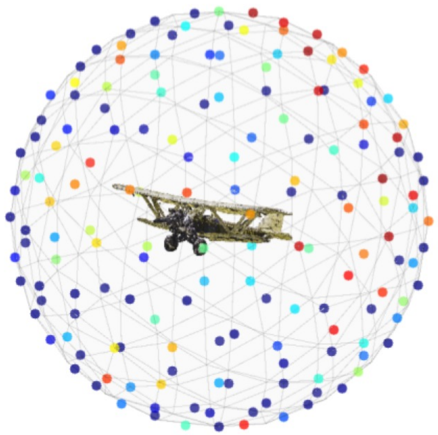
longdress, OD09_QP90 *longdress*, OD08_QP90

	Inanimate objects				Human figures			
	PLCC	SROCC	RMSE	OR	PLCC	SROCC	RMSE	OR
MS-SSIM	0.872	0.878	0.593	0.630	0.918	0.897	0.362	0.667
MS-SSIM ($K = 6$)	**0.955**	**0.944**	**0.359**	**0.556**	**0.933**	**0.927**	**0.328**	**0.519**

E. Alexiou, "Perceptual quality of point clouds with application to compression", PhD Thesis, 2021

Weighting *user views*

- Importance weights
  - Cluster viewpoints to fixed camera arrangements
  - Weights based on time of inspection
- **Advantages:**
  - Consider views that were inspected
  - Less computational costs: up to 70% reduction for K = 162
- **Disadvantages:**
  - Practicality
- Results show improvements over simple average on fixed camera arrangements



K = 162 (w = 0.006)

w = 0.044 **w = 0.017** **w = 0.003**

	Inanimate objects				Human figures			
	PLCC	SROCC	RMSE	OR	PLCC	SROCC	RMSE	OR
$K = 1$ (AVG)	0.951	0.944	0.373	0.519	0.952	0.935	0.279	0.519
$K = 1$ (WAVG)	0.951	0.944	0.373	0.519	0.952	0.935	0.279	0.519
$K = 6$ (AVG)	0.955	**0.944**	0.359	0.556	0.933	0.927	0.328	**0.519**
$K = 6$ (WAVG)	**0.956**	**0.944**	**0.356**	**0.519**	**0.949**	**0.935**	**0.289**	**0.519**
$K = 12$ (AVG)	0.949	0.944	0.381	**0.519**	0.926	0.920	0.344	**0.519**
$K = 12$ (WAVG)	**0.951**	**0.949**	**0.376**	0.556	**0.943**	**0.935**	**0.303**	**0.519**
$K = 42$ (AVG)	0.949	0.945	0.383	**0.519**	0.926	0.915	0.345	0.556
$K = 42$ (WAVG)	**0.951**	**0.947**	**0.374**	**0.519**	**0.949**	**0.933**	**0.289**	**0.519**
$K = 162$ (AVG)	**0.949**	**0.945**	0.384	**0.519**	0.925	0.915	0.347	**0.519**
$K = 162$ (WAVG)	**0.949**	0.942	**0.382**	0.556	**0.948**	**0.936**	**0.290**	**0.519**

E. Alexiou and T. Ebrahimi, "Exploiting user interactivity in quality assessment of point cloud imaging", in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019

Future challenges

- Objective metrics for dynamic point clouds
- Currently assessed using temporal pooling
  - Both for image-based and model-based
  - Similar to 2D content
- However, this approach doesn't consider temporal inconsistencies
  - Open research question

Thank you for your attention!