Perceptually Guided 3DGS Streaming and Rendering for Mixed Reality

Yunxiang Zhang[1*], Sai Harsha Mupparaju[1*], Kenneth Chen[1], Jenna Kang[1], Xinyu Zhang[2], Maito Omori[2], Kazuyuki Arimatsu[2], Qi Sun[1]

[1] New York University, USA   [2] Sony Interactive Entertainment, Japan

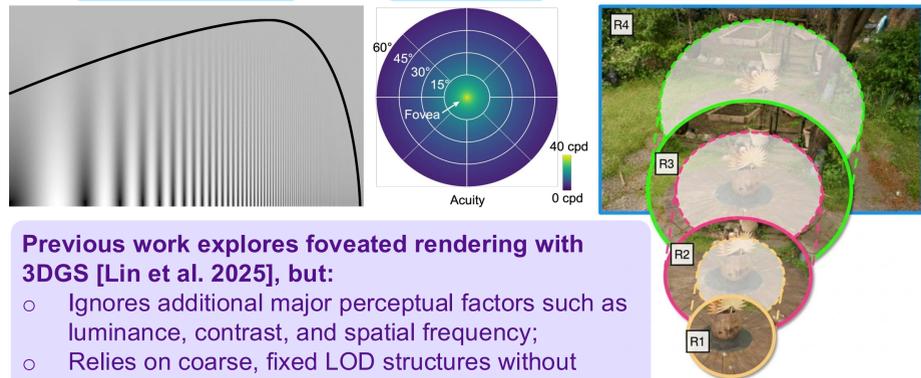**WACV 2026 TUCSON, AZ 3/6 – 3/10**

Introduction

Why is 3DGS not yet MR-ready? Because MR requires:
- High refresh rates (≥ 60 FPS for real-time immersive interaction);
- High-resolution stereo rendering (≥ 4K × 2K for retinal-level clarity);
- Tightly limited power and compute budgets in untethered setups.

Contrast Sensitivity

Visual Acuity

Foveated 3DGS
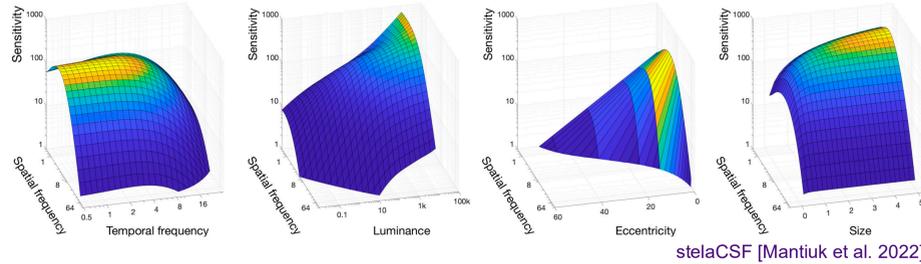
MetaSapiens [Lin et al. 2025]

Previous work explores foveated rendering with 3DGS [Lin et al. 2025], but:
- Ignores additional major perceptual factors such as luminance, contrast, and spatial frequency;
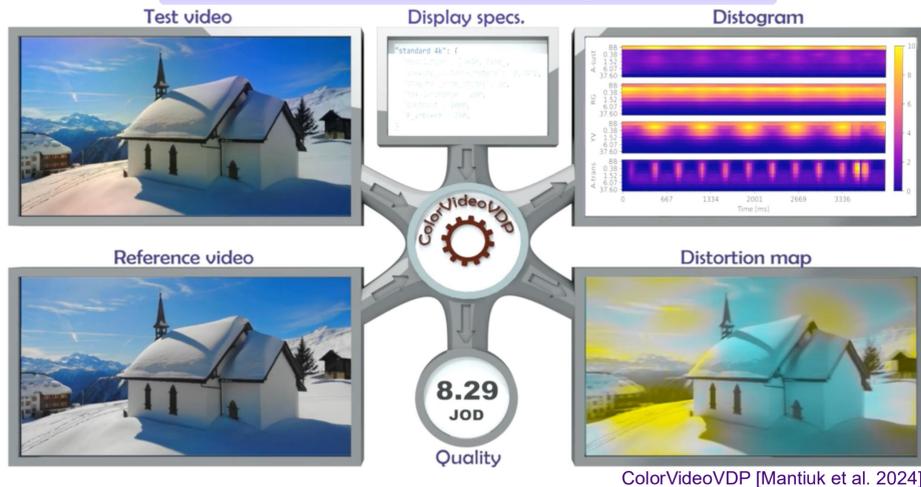- Relies on coarse, fixed LOD structures without adapting to local scene content or available budgets.

Perceptual Factors and Their Impact on Contrast Sensitivity

stelaCSF [Mantiuk et al. 2022]

Visual Difference Predictor for Images and Videos

Test video

Display specs.

Distogram

Reference video

ColorVideoVDP

8.29 JOD

Quality

Distortion map

ColorVideoVDP [Mantiuk et al. 2024]

Method

### Step 1: Build an LOD Structure for 3D Gaussians

We construct a continuous LOD structure for 3D Gaussians by adapting the tree representation from Hierarchical 3DGS [Kerbl et al. 2024]. Here, LOD is defined as a screen-space projected size threshold. Gaussians corresponding to a particular LOD can be efficiently retrieved via a cut over the tree structure.

LOD Fetching
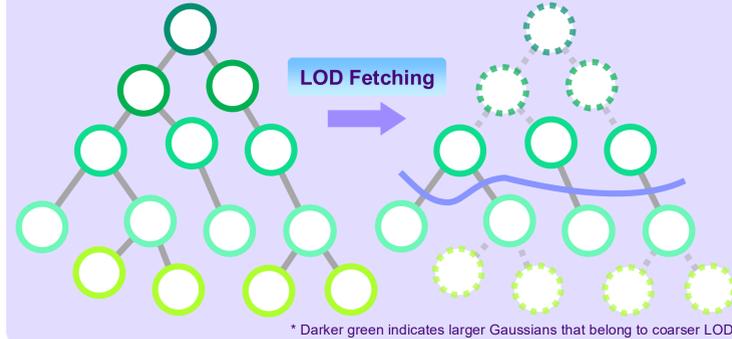
* Darker green indicates larger Gaussians that belong to coarser LOD

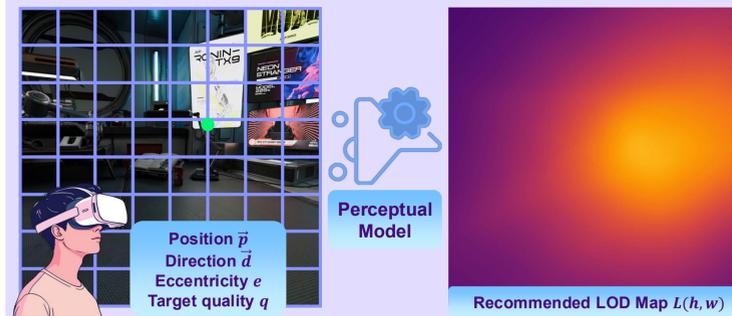### Step 2: Learn a Scene-Aware Perceptual Model for LOD Prediction

We adopt FovVideoVDP [Mantiuk et al. 2021], a perceptual quality metric that models the spatial, temporal, and eccentricity-dependent characteristics of human vision, and use it to generate data tuples of viewpoint, perceptual quality measured in JODs, and corresponding LOD values.
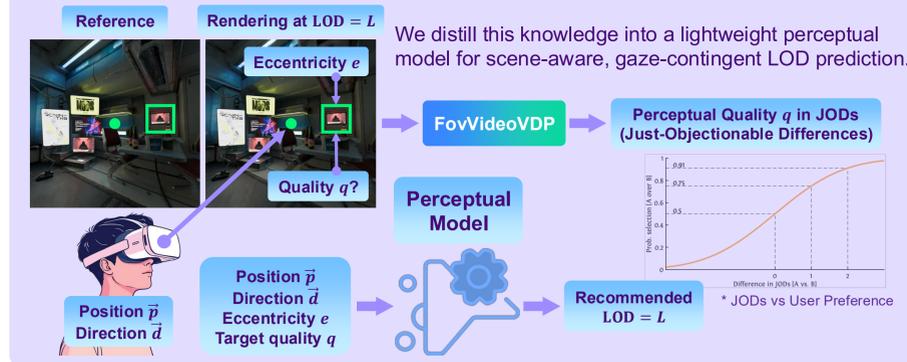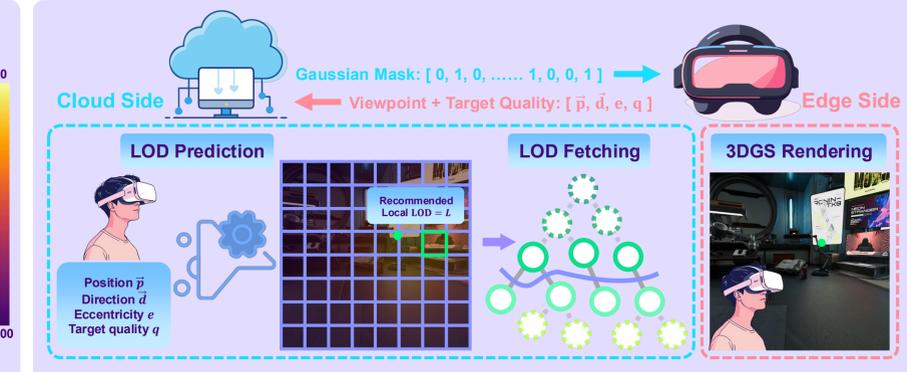
We distill this knowledge into a lightweight perceptual model for scene-aware, gaze-contingent LOD prediction.

Reference

Rendering at LOD = L

Eccentricity e

FovVideoVDP

Quality q?

Perceptual Model

Perceptual Quality q in JODs (Just-Objectionable Differences)

Position $\vec{p}$
Direction $\vec{d}$
Eccentricity e
Target quality q

Recommended LOD = L

Position $\vec{p}$
Direction $\vec{d}$

* JODs vs User Preference

### Step 3: Modulate LOD across the Visual Field

We partition the visual field into a dense grid of viewing frustums and apply the spatially varying LOD map to adaptively allocate budget to each frustum.

Position $\vec{p}$
Direction $\vec{d}$
Eccentricity e
Target quality q

Perceptual Model

Recommended LOD Map $L(h, w)$

Step 4: An Edge-Cloud Collaborative Rendering Framework

Gaussian Mask: [0, 1, 0, ... 1, 0, 0, 1]

Viewpoint + Target Quality: [$\vec{p}, \vec{d}, e, q$]

Cloud Side   **Edge Side**

LOD Prediction   LOD Fetching   3DGS Rendering

Recommended Local LOD = L

Position $\vec{p}$
Direction $\vec{d}$
Eccentricity e
Target quality q

Perceptually Guided and Spatially Adaptive LOD Modulation across Eye Gaze Positions and Compute Budget Levels

700K Gaussians / **300K Gaussians**

Gaze Position 1 / Gaze Position 2 (repeated)
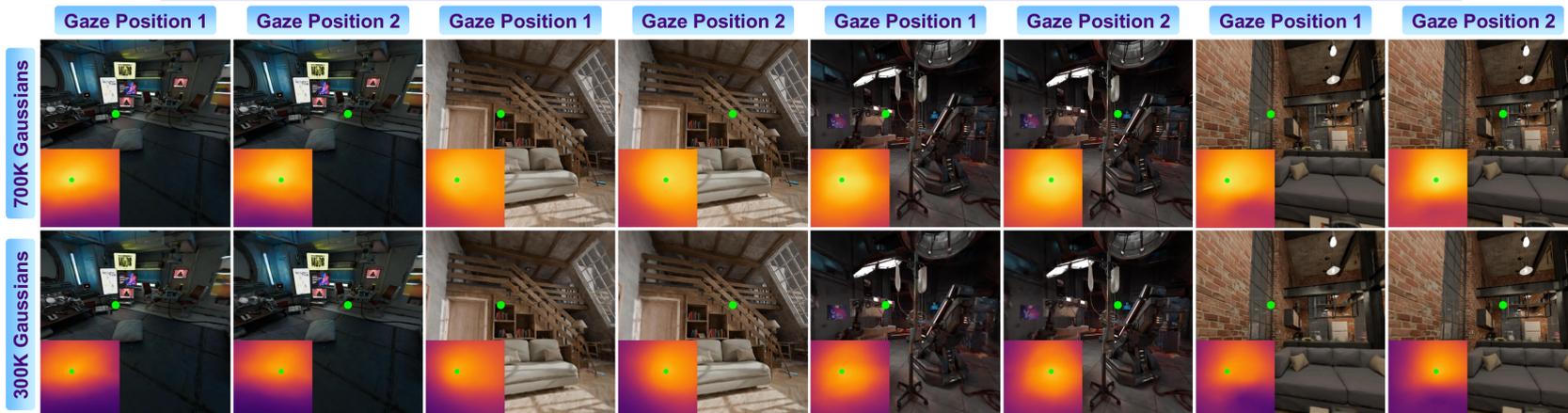
Evaluation

Comparison with Uniform and Foveated LOD Baselines

Reference | Uniform | Foveated | Perceptual

88% reduction in compute

MR User Study: Pairwise Video Quality Comparison

Which test video (**A** or **B**) more closely matches the **Reference** in visual fidelity?

Test A | Reference | Test B
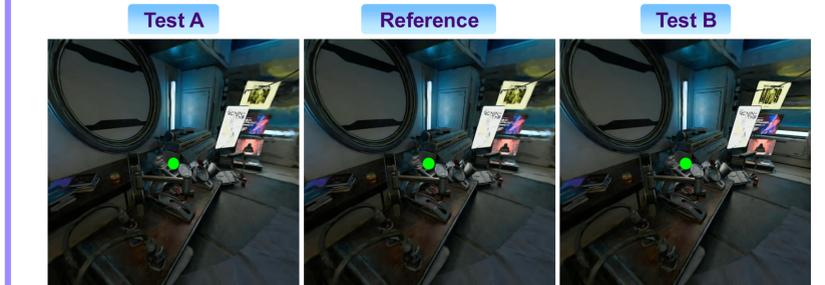
Perceptual Quality in JODs (Just-Objectionable Differences)

JODs Evaluated across 100 Test Views

q_{ref} Reference 2560K Gaussians

88% reduction in compute
81% reduction in compute
73% reduction in compute

— Uniform
— Foveated
— Perceptual

Number of Gaussians [×10⁵]

JODs Converted from User Study Data

q_{ref} Reference 2560K Gaussians

88% reduction in compute
81% reduction in compute
73% reduction in compute

— Uniform
— Foveated
— Perceptual

Number of Gaussians [×10⁵]