ORAL (SPOTLIGHT)

Neural Harmonic Textures

for High-Quality Primitive-Based Neural Reconstruction

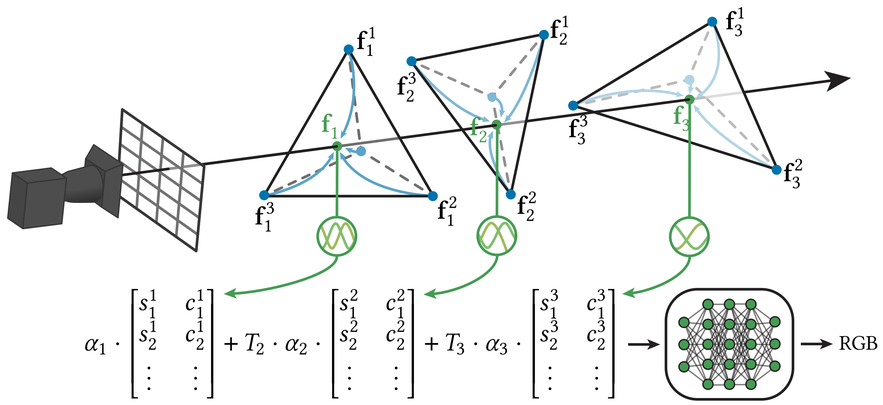
Jorge Condor[1,2] · Nicolas Moënne-Loccoz[1] · Merlin Nimier-David[1] · Piotr Didyk[2] · Zan Gojcic[1] · Qi Wu[1]

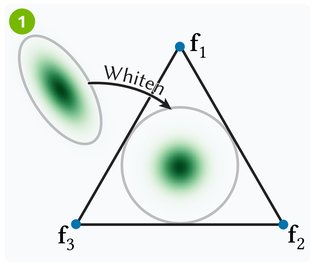
[1]NVIDIA [2]Università della Svizzera italiana, Lugano, Switzerland

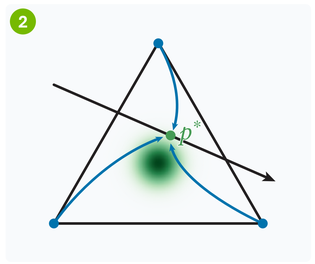
ECCV 2026 | MALMÖ | SEPT 8–12

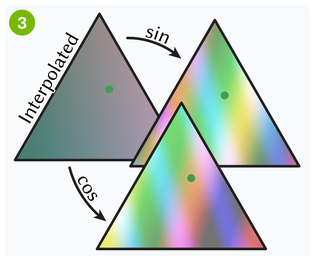
Project page · Paper · Code

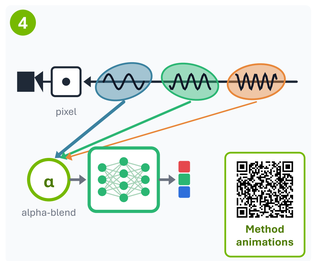
OUR METHOD











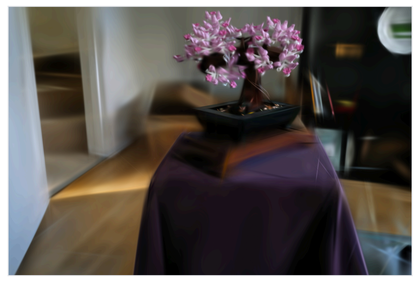
1 Canonical transform
Each Gaussian is whitened to a unit isotropic one; latents sit on its fixed bounding tetrahedron.

2 Interpolation
Barycentric interpolation at the max-response point **p*** of each ray-primitive hit.

3 Harmonic texturing
sin and **cos** of the interpolated feature: interpolation sets the frequency, opacity the amplitude.

4 Deferred shading
Harmonics are **alpha-blended** along the ray; a tiny MLP decodes colour **once per pixel** (gsplat, fused CUDA kernels).

MOTIVATION: APPEARANCE AND GEOMETRY ARE COUPLED



3DGUT-MCMC · 10K Gaussians · **23.97 dB**

+ NHT (ours) · 10K Gaussians · **27.63 dB**

Same primitive budget, **+3.7 dB**: appearance no longer limited by the primitive count

vs. explicit per-primitive textures: BBSplat [Svitov et al. 2024], MipNeRF360

Method	PSNR ↑	LPIPS ↓	MB ↓	FPS ↑ (A100)
BBSplat · 300K	27.09	0.255	215.5	17.3
NHT · 300K (equal primitives)	27.83	0.273	**49.3**	**137.8**
NHT · 1M (equal memory)	**28.36**	**0.232**	161.6	84.0

10K Gaussians and 2D fitting videos

RESULTS

Method	PSNR ↑	SSIM ↑	LPIPS ↓	FPS ↑
3DGUT + SH	27.38	0.867	0.252	**243**
3DGUT + NHT (ours)	**28.04**	**0.874**	**0.237**	202

Controlled comparison in gsplat: 3DGUT [Wu et al. 2025] + MCMC [Kheradmand et al. 2024], 1M primitives, 30k iterations, average of MipNeRF360, Tanks & Temples and Deep Blending, RTX A6000 Ada.

+0.66 dB PSNR
over 3DGUT + SH at 1M primitives

163 FPS
MipNeRF360 average at 1M primitives, fused kernels, RTX 4090 (up to 261 FPS on garden)

30 it/s
training step (forward + backward), MipNeRF360 average, 1M primitives, RTX 4090

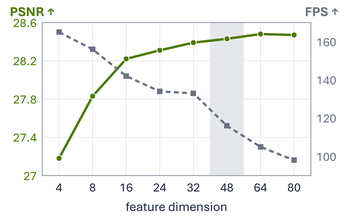
11 min
garden trained to 30k iterations at 1M primitives, fused kernels, RTX 4090

Method	MipNeRF360 PSNR ↑	MipNeRF360 SSIM ↑	MipNeRF360 LPIPS ↓	Tanks & Temples PSNR ↑	Tanks & Temples SSIM ↑	Tanks & Temples LPIPS ↓	Deep Blending PSNR ↑	Deep Blending SSIM ↑	Deep Blending LPIPS ↓
Mip-NeRF 360	27.60	0.788	0.275	22.22	0.754	0.290	29.40	0.899	0.306
ZipNeRF	28.55	0.829	0.218	23.64	0.836	0.179	–	–	–
2DGS	27.22	0.804	0.275	22.85	0.827	0.244	29.56	0.904	0.325
3DGS-MCMC	27.99	0.830	0.229	24.46	0.866	0.174	29.49	0.912	0.306
3DGUT-MCMC	27.82	0.826	0.233	24.20	0.861	0.180	29.87	0.913	0.309
Beta Splatting-MCMC	28.12	0.831	0.238	24.54	0.866	0.196	29.56	0.907	0.316
Spherical Voronoi	28.56	**0.835**	0.228	24.80	0.871	0.172	30.34	0.914	**0.299**
Triangle Splatting	27.00	0.808	0.231	23.05	0.843	0.191	28.92	0.891	0.308
Textured Gaussians	27.35	0.827	–	24.26	0.854	–	28.33	0.891	–
Radiance Meshes	27.15	0.810	0.274	23.13	0.851	0.200	29.39	0.901	0.362
Neural Harmonic Textures (ours)	**28.74**	0.834	**0.216**	**25.68**	**0.882**	**0.141**	**30.94**	**0.919**	0.302

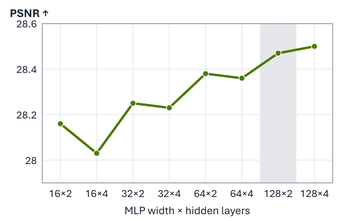
Result videos

ABLATIONS

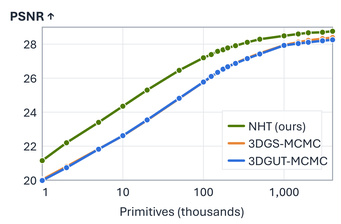
Features per primitive



MLP width × hidden layers



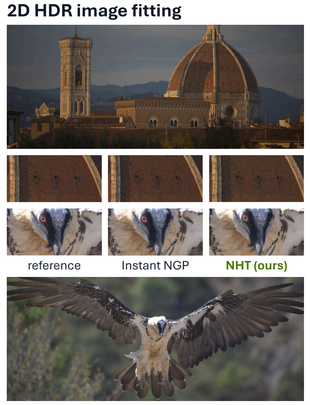
Quality vs. primitive count



APPLICATIONS

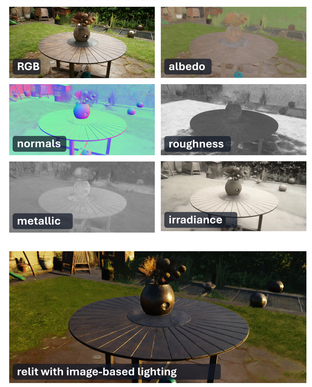
One representation, any signal: only the decoder head changes

2D HDR image fitting



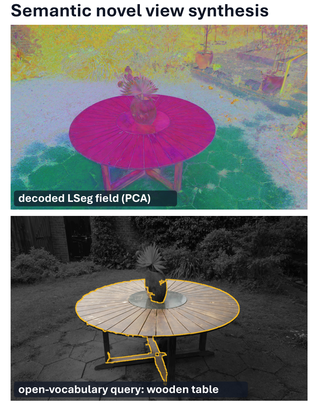
45.7 MP RAW at 100× compression: **38 % lower LPIPS** than Instant NGP [Müller et al. 2022] at the same size.

PBR material stacks



Maps from RGB2X intrinsics [Zeng et al. 2024], fit jointly with RGB and rendered at **60+ FPS**; relit from the decoded stack.

Semantic novel view synthesis



Joint RGB + **512-D LSeg** [Li et al. 2022] field in real time, about 3× less memory than Feature 3DGS [Zhou et al. 2024] and 10x faster; open-vocabulary masking from novel views.

TAKEAWAYS

Expressive: textures detach appearance from geometry: fewer primitives, sharper detail. **Fast**: one deferred MLP evaluation per pixel; real time, minutes of training. **General**: drops into 3DGS, 3DGUT, 2DGS and Triangle Splatting; decodes any signal. **Scalable**: feature and MLP sizes trade quality for memory and speed.

PBR, semantic and 2D demos