NeurIPS 2026

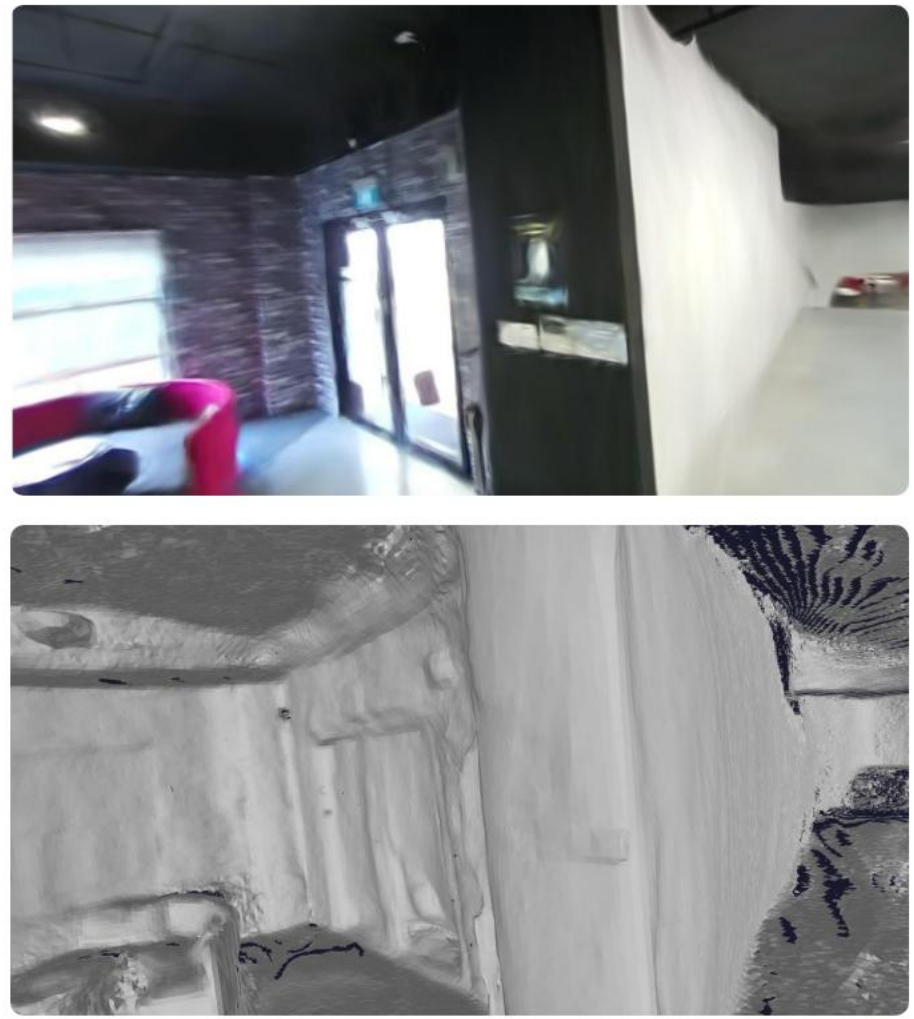
AffineGS

Unlocking Accurate Geometry in 3DGS via Spatially Varying Affine Rectification

Tze Ho Elden Tse[1,2*] · Jizong Peng[1*] · Richard Chen[1] · Angela Yao[2]

[1]dConstruct Robotics [2]National University of Singapore [*]Equal contribution

eldentse.github.io/AffineGS



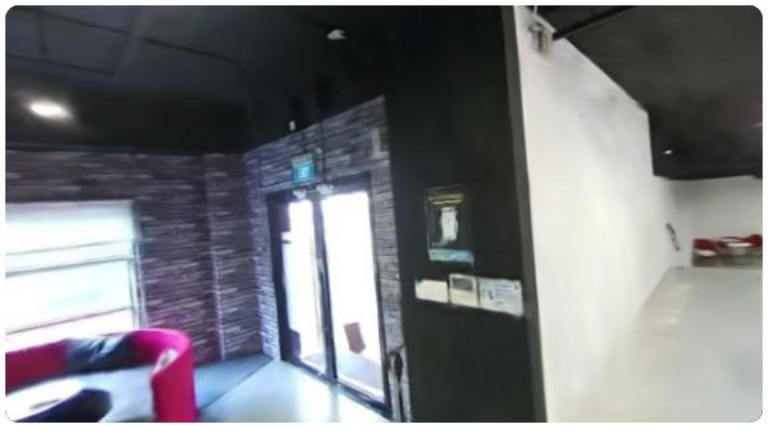
Rendered view (top) and extracted mesh (bottom), with AffineGS

MOTIVATION

3DGS renders look right. The geometry underneath doesn't.



Reference image



3DGS rendering photorealistic



3DGS extracted mesh noisy

Gaussians overfit the training views → floaters, multi-layered surfaces, noisy meshes.
A problem for robotics, AR/VR and anything that needs a reliable 3D model.

BACKGROUND

The usual fix: monocular depth priors, globally aligned

Sensor depth

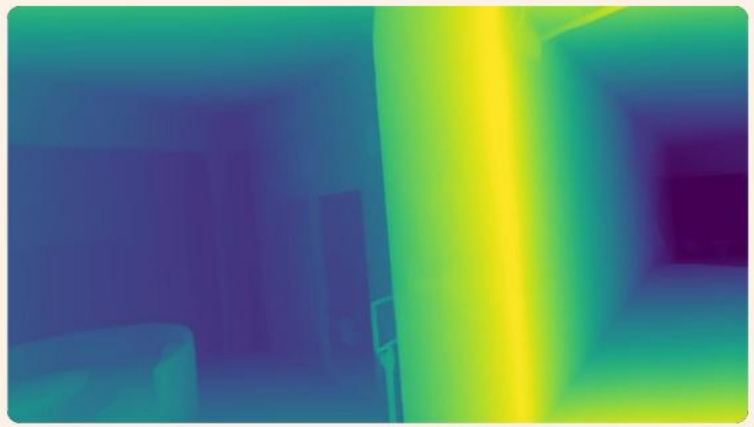
e.g. DN-Splatter

✓ accurate geometry
× needs RGB-D hardware
× not available for most captures

Monocular depth prior

e.g. Depth-Anything-V2

✓ from any RGB image
✓ strong structural cues
✓ scalable



monocular disparity *d*

EXISTING PRACTICE: ONE GLOBAL ALIGNMENT

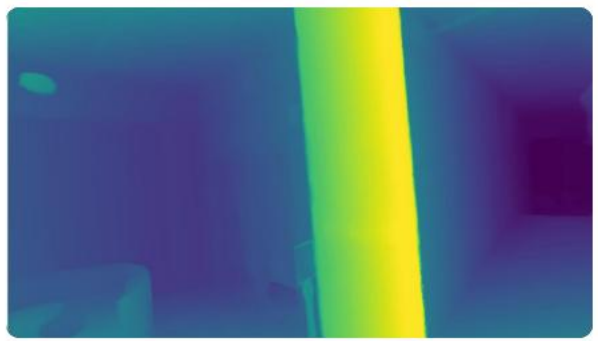
$$\frac{1}{d} \approx a\mathcal{D} + b$$

one scale *a* and one shift *b* for the **whole image**
d: monocular disparity · $\mathcal{D}$: depth rendered by 3DGS

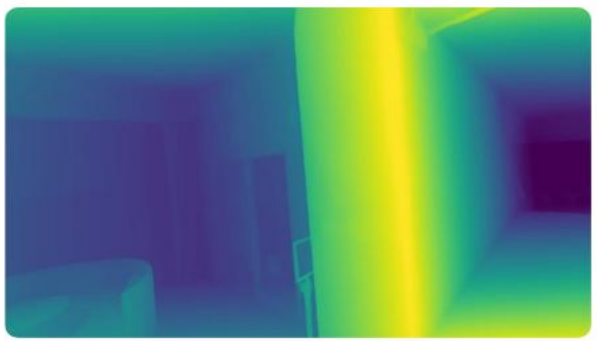
the assumption we question

KEY OBSERVATION

Monocular depth is distorted piecewise-linearly, not globally



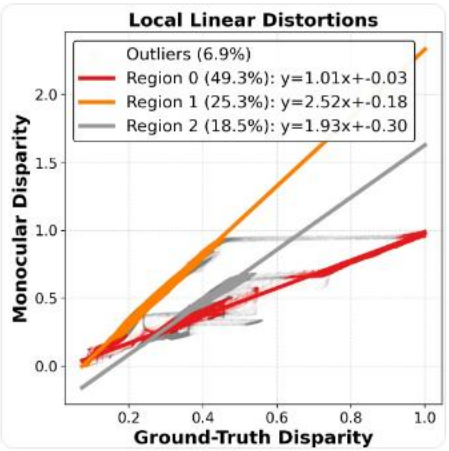
Ground-truth disparity



Monocular disparity

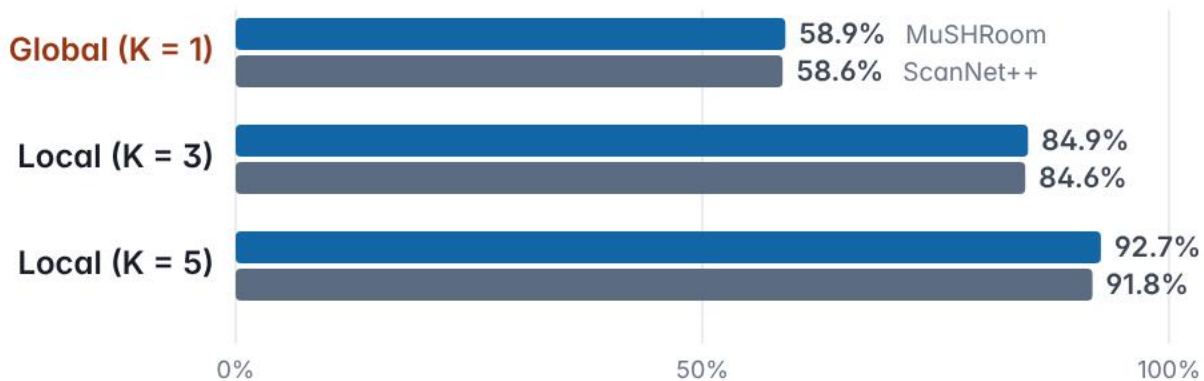


Spatially coherent affine regions



mono vs. GT: **3 lines**, not 1

PIXELS ALIGNED BY K AFFINE FITS (DEPTH-ANYTHING-V2, RESIDUAL < 0.02)



> One global transform either **averages the error** everywhere, or **fits one region** and warps the rest.

METHOD · 1 / 2

Spatially varying affine rectification

MODEL: SCALE AND SHIFT BECOME FIELDS

$$\frac{1}{d(u,v)} = \alpha(u,v)\,\mathcal{D}(u,v) + \beta(u,v)$$

d: monocular disparity · $\mathcal{D}$: rendered 3DGS depth · global methods fix α, β to constants

SOLVE FROM GRADIENTS, NO SENSOR DEPTH

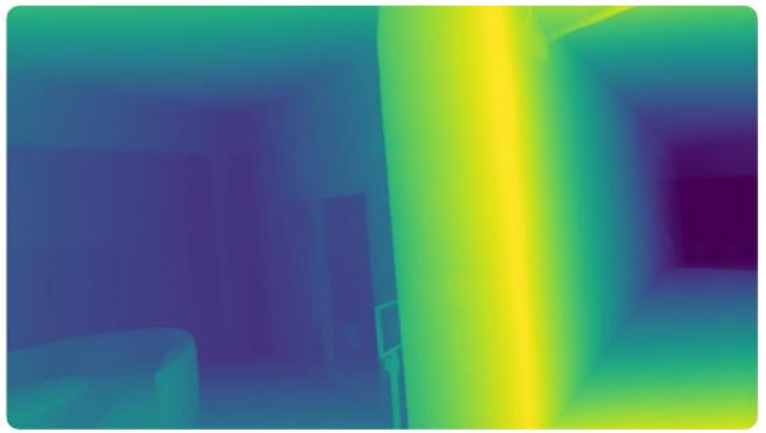
$$-\tfrac{1}{d^2}\nabla d \;\approx\; \alpha\,\nabla\mathcal{D} \qquad (\nabla\alpha \approx 0,\ \nabla\beta \approx 0)$$

$$\alpha = \frac{\|\nabla d\|}{d^2\,\|\nabla\mathcal{D}\| + \epsilon}, \qquad \beta = \frac{1}{d} - \alpha\,\mathcal{D}$$

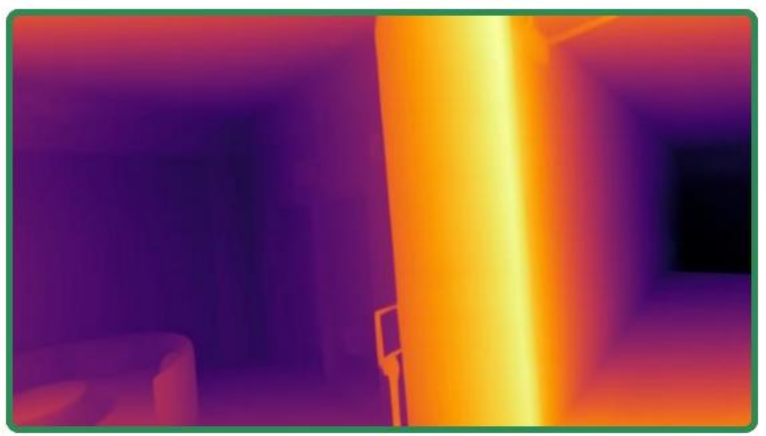
EDGE-AWARE GUIDED FILTER (GUIDE = d)

→ α_{smooth}, β_{smooth}: smooth within surfaces, sharp across boundaries

re-estimated every iteration · no sensor depth

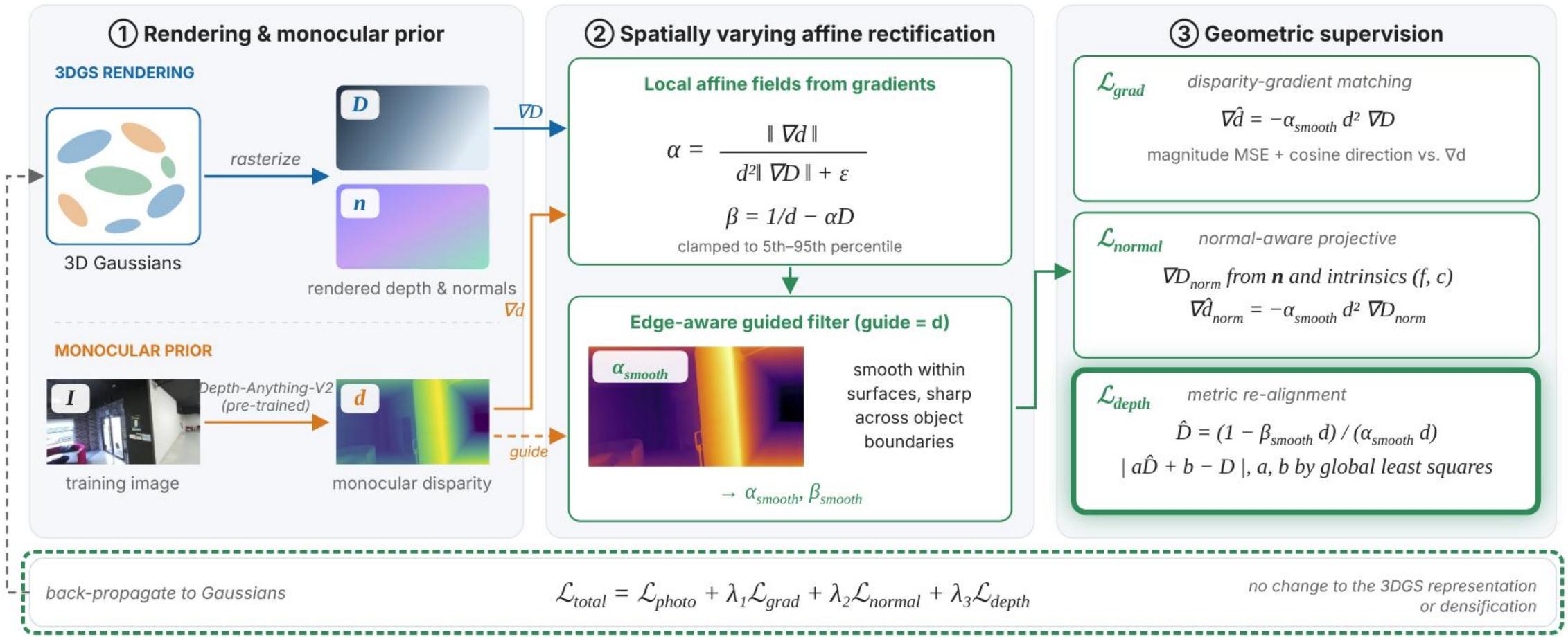


monocular disparity d



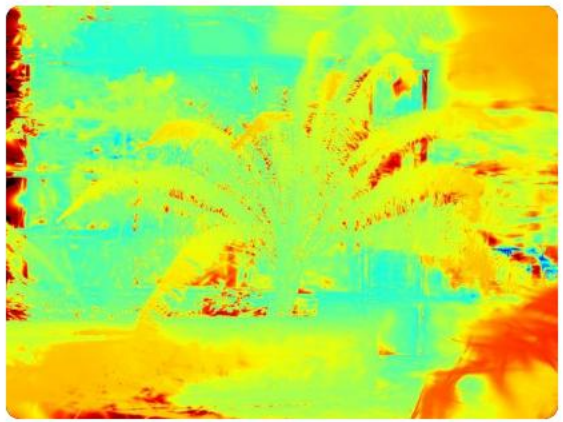
estimated scale field α_{smooth}

Rectified fields supervise geometry through three losses

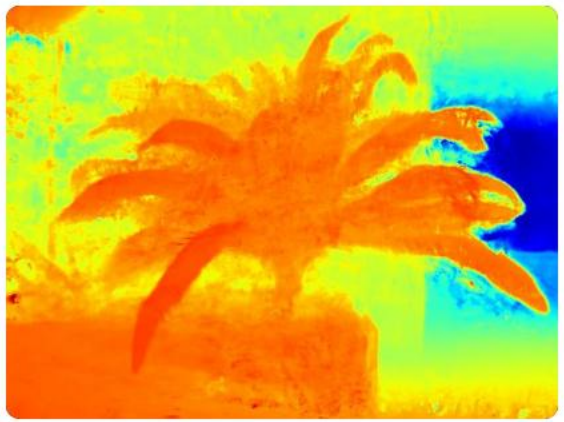


RESULTS · SPARSE VIEWS

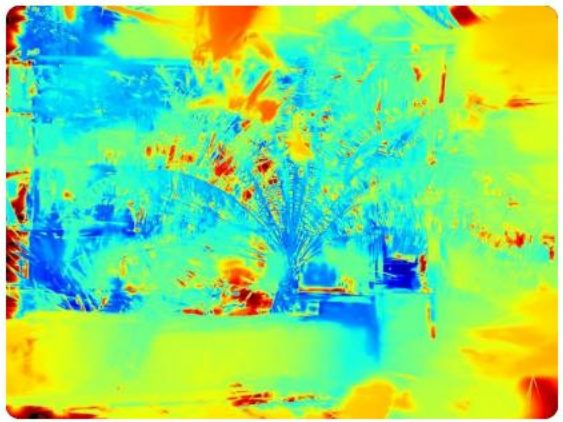
Sparse views (NeRF-LLFF): no structural collapse



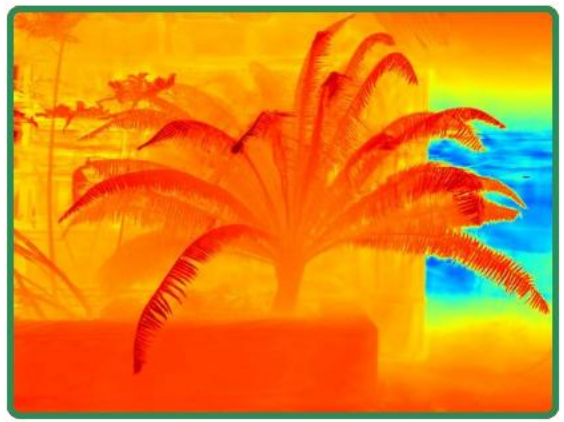
3DGS



DRGS

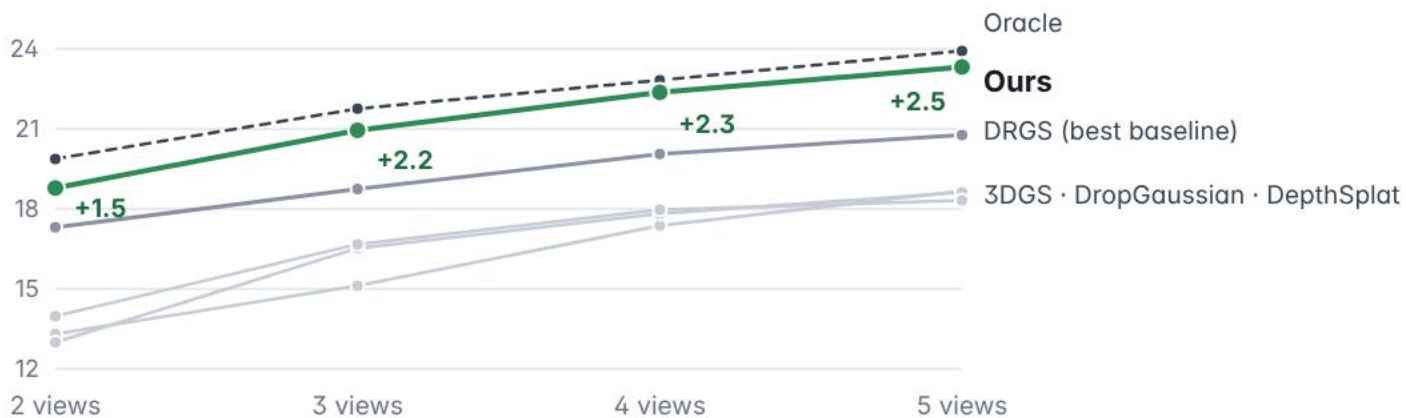


DropGaussian



Ours

MEAN PSNR (DB) VS. NUMBER OF TRAINING VIEWS



+1.5–2.5 dB

over the best baseline (DRGS) at every view count

within 0.5–1.1 dB of the Oracle, which uses dense pseudo-GT depth

RESULTS · DEPTH & NOVEL VIEWS

Same monocular prior, 6× lower depth error (MuSHRoom)



DN-Splatter
same mono prior, global



DN-Splatter
uses sensor depth



Ours
same mono prior, rectified

Rendered normals

DEPTH ERROR, ABS REL ↓



RENDERING QUALITY, PSNR ↑

24.49
Ours, no sensor depth

24.58
DN-Splatter, sensor depth

RESULTS · MESHES

Cleaner meshes, without sensor depth



DN-Splatter



Ours (ScanNet++)

MESH F-SCORE ↑ ON MUSHROOM (5 CM)



ABLATION · ALL THREE LOSSES

F-score **0.728** → **0.855**

PSNR **−0.35 dB** only

SUMMARY

Takeaways

1. **Monocular depth priors are distorted piecewise-linearly.**
   A single global scale-and-shift is ill-posed.

2. **Spatially varying, edge-aware affine fields** rectify the prior, with normal-aware gradient losses.

3. **Much better 3DGS geometry, same photorealism.**
   No sensor depth, no architectural changes.

Project page & code
eldentse.github.io/AffineGS
github.com/eldentse/AffineGS

Thank you!