ICCV HONOLULU HAWAII
OCT 19-23, 2025

SurfaceSplat: Connecting Surface Reconstruction and Gaussian Splatting

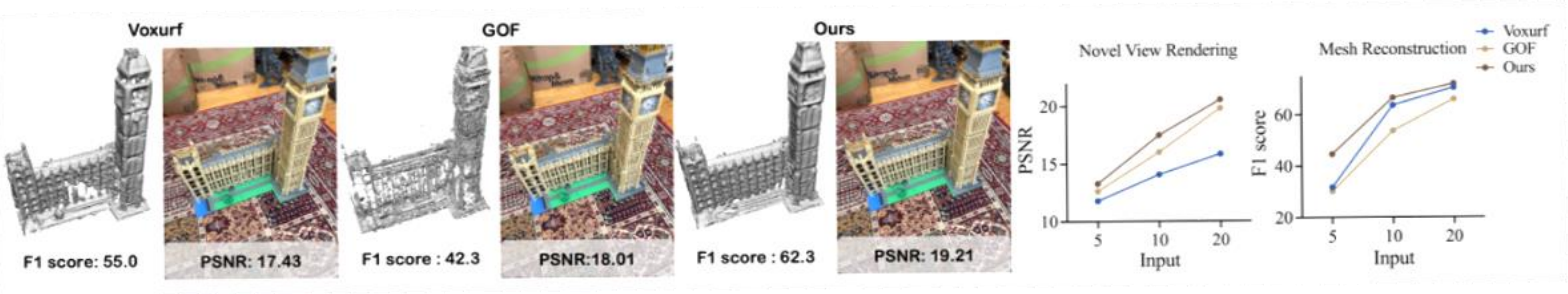
Zihui Gao[1,3*], Jia-Wang Bian[2*], Guosheng Lin[3], Hao Chen[1†], Chunhua Shen[1]

[1]Zhejiang University, China [2]ByteDance Seed [3]Nanyang Technological University, Singapore

[*]Equal contribution [†]Corresponding author

https://github.com/aim-uofa/SurfaceSplat

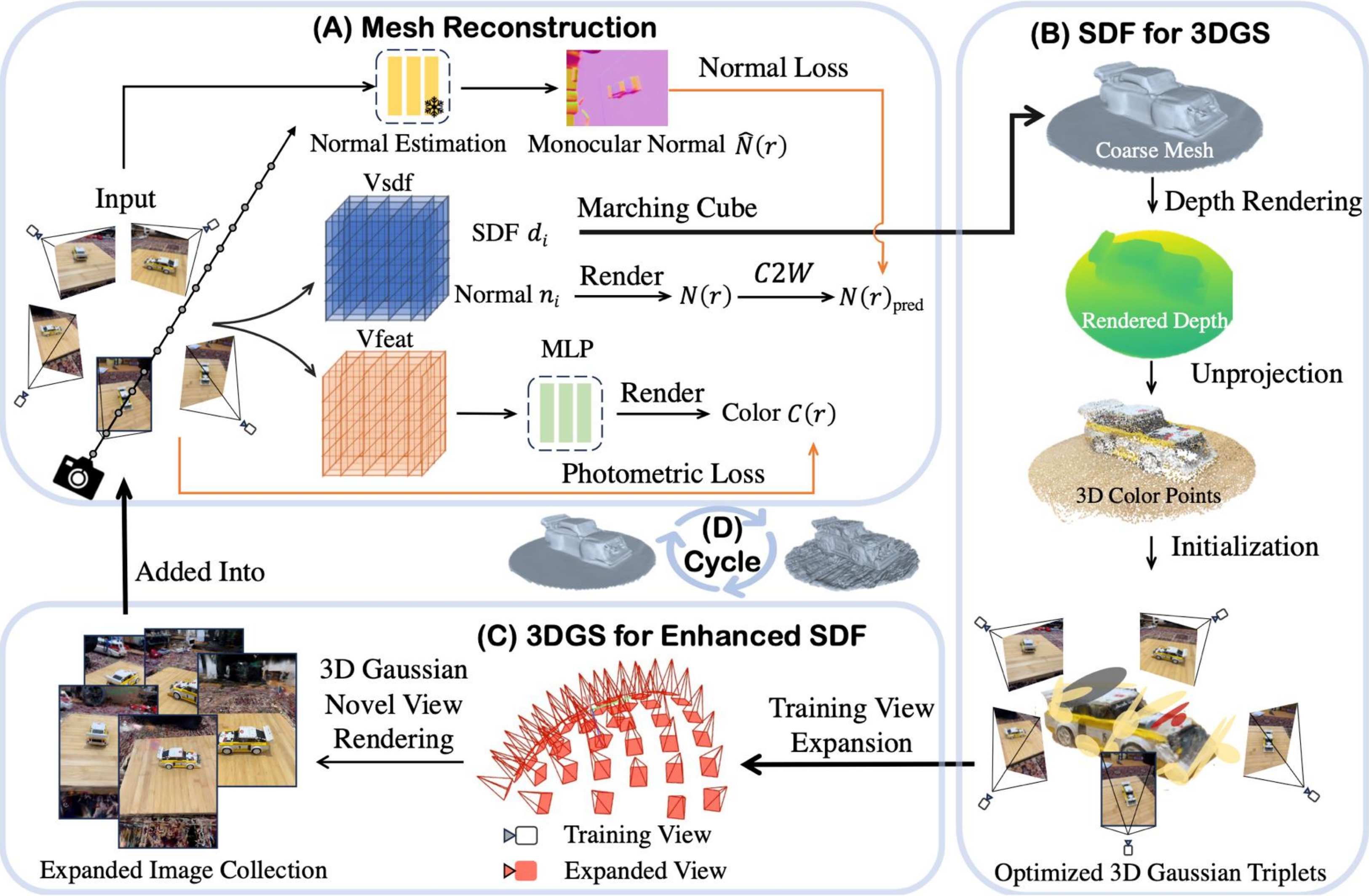
Motivation



What We Found:

- Signed Distance Function (SDF)-based methods like Voxurf capture global structures well but lack fine details.
- 3D Gaussian Splatting (3DGS)-based approaches lack global geometry coherence. 3DGS-based methods like GOF and 2DGS leverage a pre-computed sparse point cloud for image rendering.
- SDF-based methods outperform 3DGS in surface reconstruction, while 3DGS excels in image rendering, as illustrated in figure above.

Method



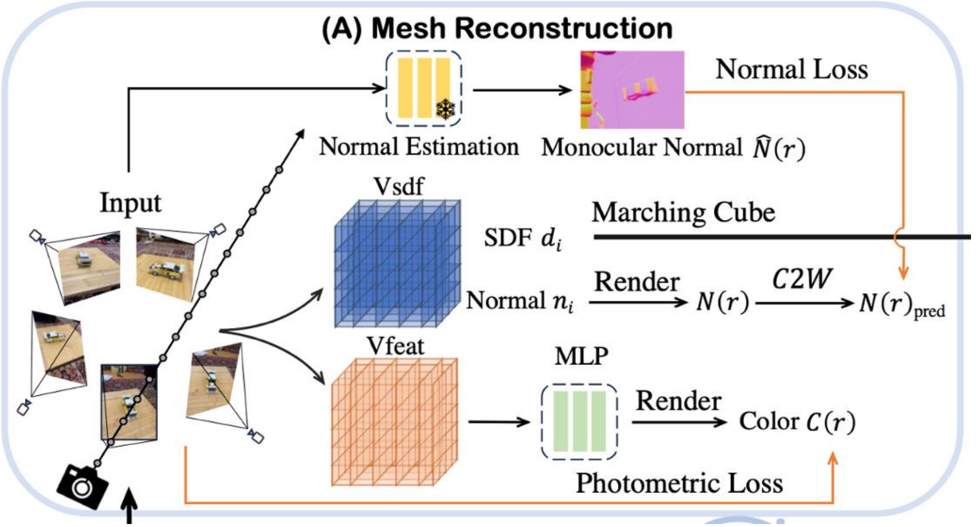
Two key ideas:

(i) SDF for Improved 3DGS:To address the limitation of 3DGS in learning global geometry, we first fit the global structure using an SDF-based representation. We then initialize 3DGS by sampling point clouds from the mesh surface

(ii) 3DGS for Enhanced SDF: To compensate for the inability of SDF-based methods, we leverage the improved 3DGS from the first step to render additional novel view- point images, expanding the dataset.

Method

(A) Coarse mesh reconstruction:



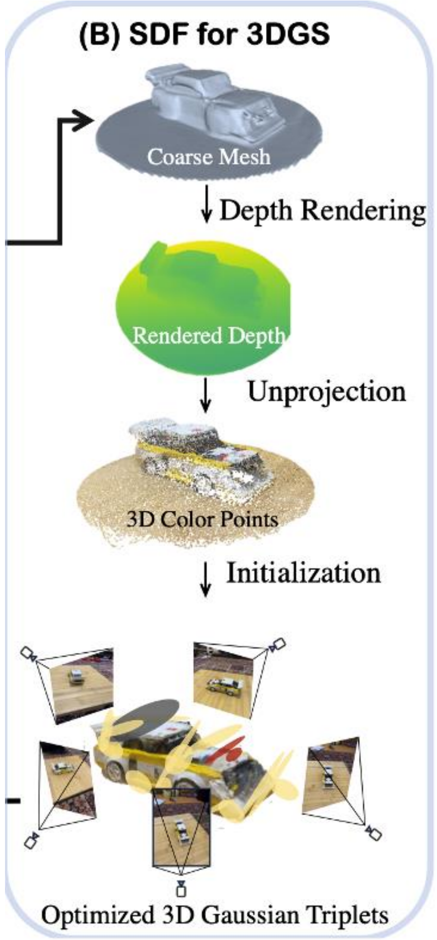
We adopt the coarse-stage surface reconstruction from Voxurf. The loss function is

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + \mathcal{L}_{TV}\left(V^{(\text{sdf})}\right) + \mathcal{L}_{\text{smooth}}\left(\nabla V^{(\text{sdf})}\right)$$

We also introduce a normal consistency loss to improve training stability

$$\mathcal{L}_{\text{normal}} = \sum \left(\|\hat{N}(\mathbf{r}) - \bar{N}(\mathbf{r})\|_1\right)$$

(B) SDF for 3DGS:

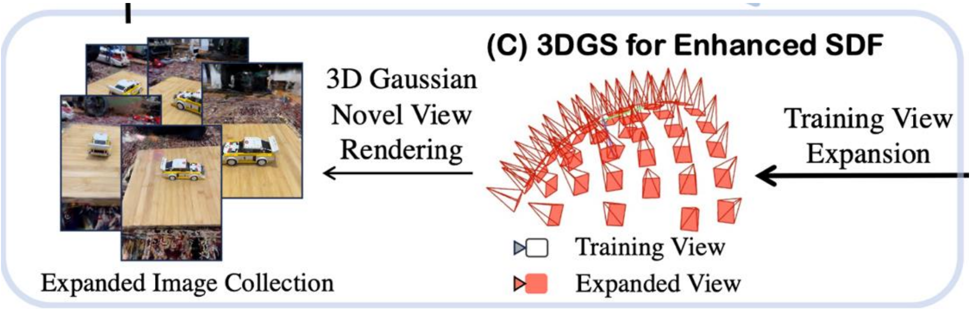


Point clouds are sampled from the mesh surface to initialize 3DGS.

- Mesh cleaning: cluster the connected mesh triangles and remove the floaters.
- Sample surface points: project the reconstructed mesh onto the training views using their known camera poses to generate depth maps and sample points from valid depth regions.

Method

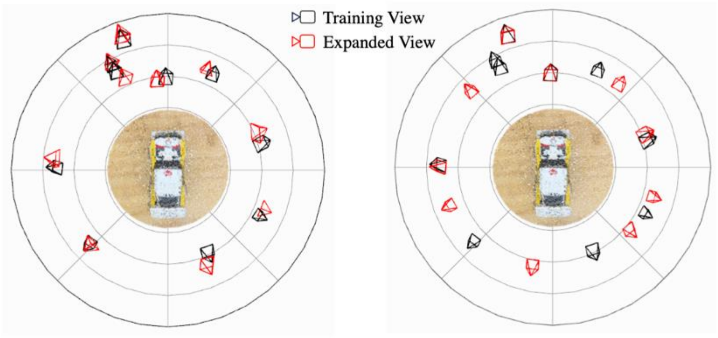
(C) 3DGS for enhanced SDF:



3DGS renders new viewpoint images to expand the training set, refining the mesh.

$$\{\mathcal{I}_{\text{new}}\} = \text{Splat}(\mathcal{G}, \{\boldsymbol{\pi}_{\text{new}}\})$$

We explore two methods for generating new camera poses.



(a) Camera position perturbation (b) Camera poses interpolation

(D) Cyclic Optimization:

Steps B and C can be repeated for iterative optimization, progressively improving performance.

Rendering Step: optimize a 3DGS model for rendering novel view images.

$$\mathcal{I}^{(n)} = \mathcal{R}\left(\mathcal{M}_c^{(n)}\right)$$

Meshing Step: refine the current mesh by finetuning it using both the newly rendered images and the original input images.

$$\mathcal{M}_f^{(n)} = \mathcal{O}\left(\mathcal{M}_c^{(n)}, \mathcal{I}^{(n)}\right)$$

We update the refined mesh.

$$\mathcal{M}_c^{(n+1)} = \mathcal{M}_f^{(n)}$$

Experiments

Table 1. Surface reconstruction and novel view synthesis results on MobileBrick. The results are averaged over all 18 test scenes with an initial input of 10 images per scene. PSNR-F is computed only on foreground regions. The best results are **bolded**.

	Mesh Reconstruction							Rendering		Time
	$\sigma = 2.5mm$			$\sigma = 5mm$			CD (mm)↓	PSNR↑	PSNR-F↑	
	Accu.(%)↑	Recall(%)↑	F1↑	Accu.(%)↑	Recall(%)↑	F1↑				
Voxurf [43]	62.89	62.54	62.42	80.93	80.61	80.38	13.3	14.34	18.34	55 mins
MonoSDF [55]	41.56	32.47	36.22	57.88	48.19	52.21	37.7	14.71	15.42	6 hrs
2DGS [15]	49.83	45.32	47.10	72.65	64.88	67.96	14.8	17.12	18.52	10 mins
GOF [57]	50.24	61.11	54.96	74.99	82.68	78.16	11.0	16.52	18.36	50 mins
3DGS [17]	\	\	\	\	\	\	\	17.19	19.12	10 mins
SparseGS [44]	\	\	\	\	\	\	\	16.93	18.74	30 mins
Ours	68.36	**69.79**	68.97	86.79	**86.82**	86.65	**9.7**	17.48	20.45	1 hr
Ours (Two cycles)	**69.61**	68.89	**69.14**	**87.79**	85.93	**86.74**	9.9	**17.58**	**20.55**	1.6 hr

Experiments

Table 2. Surface reconstruction results on DTU with 5 input views. Values indicate Chamfer Distance in millimeters (mm). ”-” denotes failure cases where COLMAP could not generate point clouds for 3DGS initialization. GSDF-10 is reported with 10 input images, as it fails in sparser settings. The best results are **bolded**, while the second-best are underlined.

Scan	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean	Time
Voxurf [43]	2.74	4.50	3.39	1.52	2.24	2.00	2.94	**1.29**	2.49	1.28	2.45	4.69	0.93	2.74	1.29	2.43	50 mins
MonoSDF [55]	**1.30**	3.45	**1.45**	**0.61**	**1.43**	**1.17**	**1.07**	1.42	**1.49**	**0.79**	3.06	2.60	0.60	2.21	2.87	1.70	6 hrs
SparseNeuS [22]	3.57	3.73	3.11	1.50	2.36	2.89	1.91	2.10	2.89	2.01	2.08	3.44	1.21	2.19	2.11	2.43	Pretrain + 2 hrs ft
2DGS [15]	4.26	4.80	5.53	1.50	3.01	1.99	2.66	3.65	3.06	2.54	2.15	-	0.96	2.17	1.31	2.84	6 mins
GOF (TSDF) [57]	7.30	5.80	6.03	2.79	4.23	3.41	3.44	4.37	3.75	2.99	3.19	-	2.64	3.67	2.25	4.03	50 mins
GOF [57]	4.37	3.68	3.84	2.29	4.40	3.28	2.84	4.64	3.40	3.76	3.56	-	3.06	2.95	2.91	3.55	50 mins
GSDF-10 [54]	6.89	6.82	7.97	6.54	5.22	1.91	5.56	4.38	7.01	3.69	6.33	6.33	3.95	6.30	2.09	5.40	3 hrs
Ours	1.55	**2.64**	1.52	1.40	1.51	1.46	1.23	1.43	1.82	1.19	**1.49**	**1.80**	**0.54**	**1.19**	**1.04**	**1.45**	1 hr

Ablations

(1) Efficacy of 3DGS for Improving SDF

Table 3. Surface reconstruction results with varying numbers of input views on MobileBrick (porsche) and DTU (scan69). The Baseline represents a pure SDF-based reconstruction without the assistance from 3DGS. δ indicates the improvement.

	MobileBrick / F1 score			DTU / CD		
Input	Baseline	Ours	δ	Baseline	Ours	δ
5	33.50	**43.11**	+9.61	2.940	**1.230**	-1.710
10	59.66	**62.37**	+2.71	1.362	**1.165**	-0.197
20	63.18	**63.88**	+0.7	1.043	**0.965**	-0.078

(2) Efficacy of SDF for enhancing 3DGS

Table 4. 3DGS rendering results with different initializations, averaged across all 18 MobileBrick test scenes.

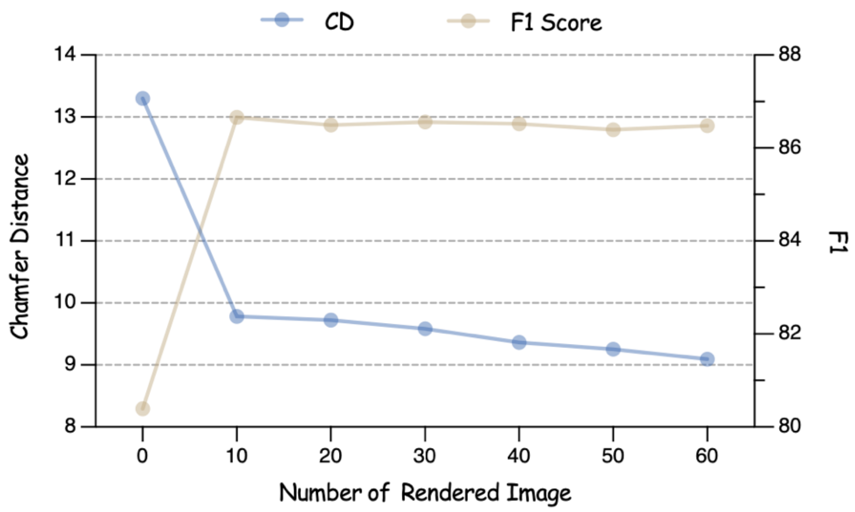
Method	Foreground PSNR
3DGS (COLMAP)	19.13
3DGS w/ mesh clean	19.88
3DGS w/ normal and mesh clean	**20.45**

(3) Different pose expansion strategies

Table 5. Ablation study on pose expansion strategies for in MobileBrick (aston) with 10 input images.

	F1↑	Recall(%)↑	CD (mm)↓
Baseline	55.8	49.9	8.7
Camera position perturbation	59.9	57.4	6.6
Camera poses interpolation	**60.8**	**59.1**	**6.4**

(4) Number of newly rendered views



Visualization

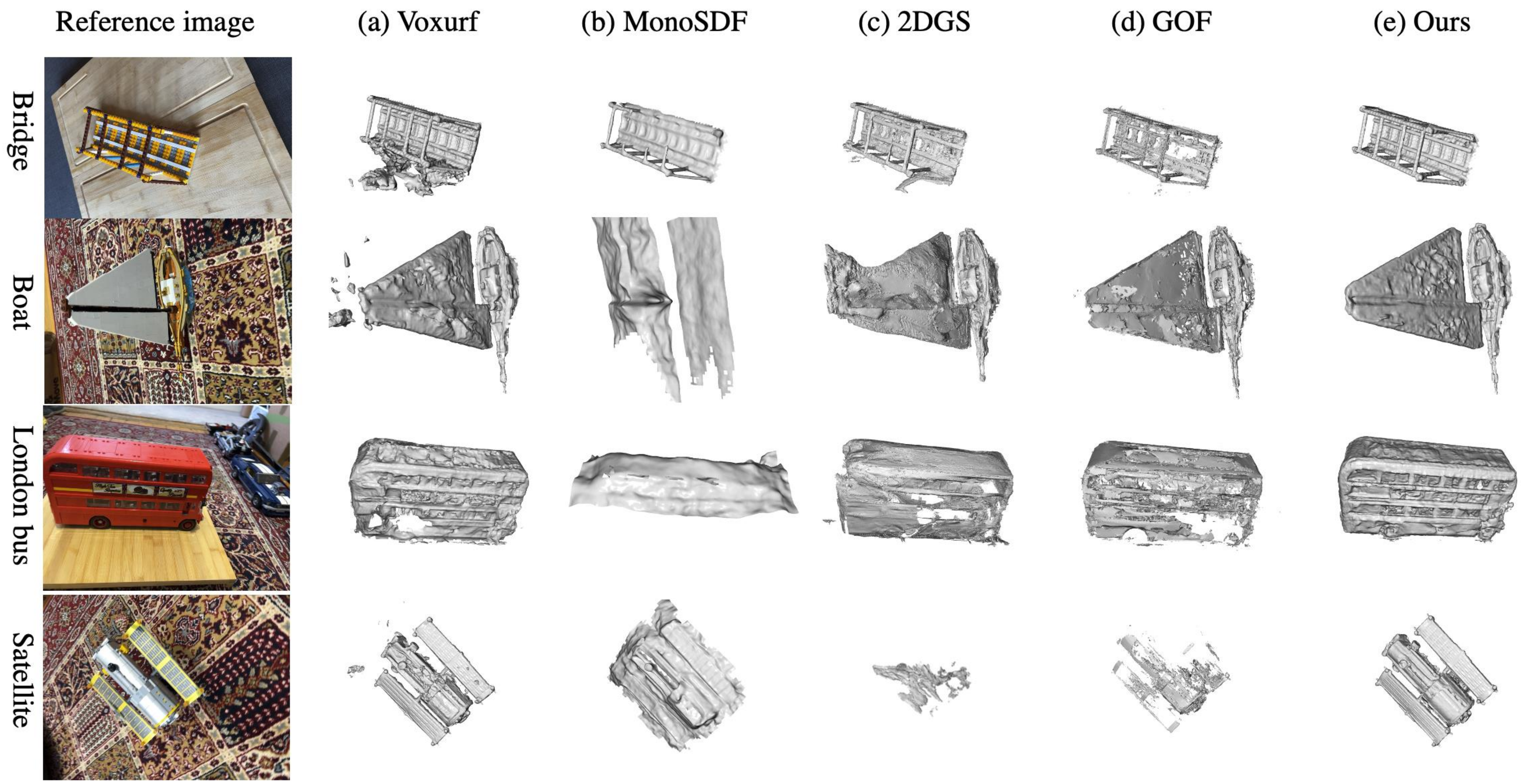


Figure 5. Qualitative mesh reconstruction comparisons on MobileBrick. See more visual results in supplementary material.

Visualization

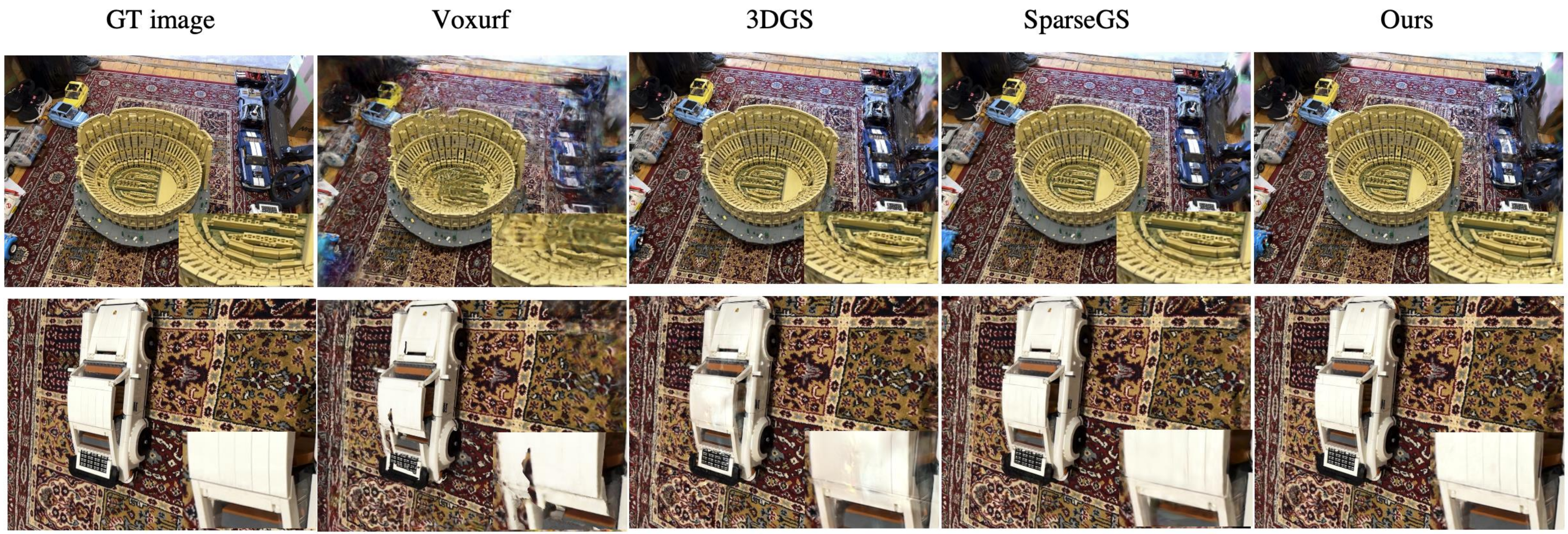


Figure 6. Qualitative novel view synthesis comparisons on MobileBrick.

Thanks for watching!