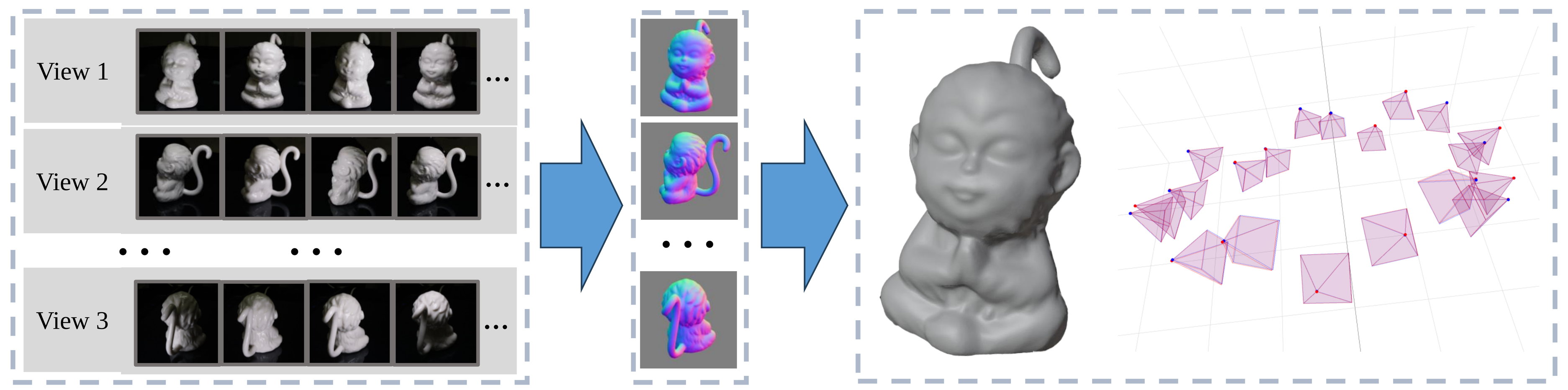
PMNI: Pose-free Multi-view Normal Integration for Reflective and Textureless Surface Reconstruction

Mingzhi Pei[1] Xu Cao[2] Xiangyi Wang[1] Heng Guo[1†] Zhanyu Ma[1]
[1]Beijing University of Posts and Telecommunications [2]Independent Researcher

CVPR Nashville JUNE 11-15, 2025

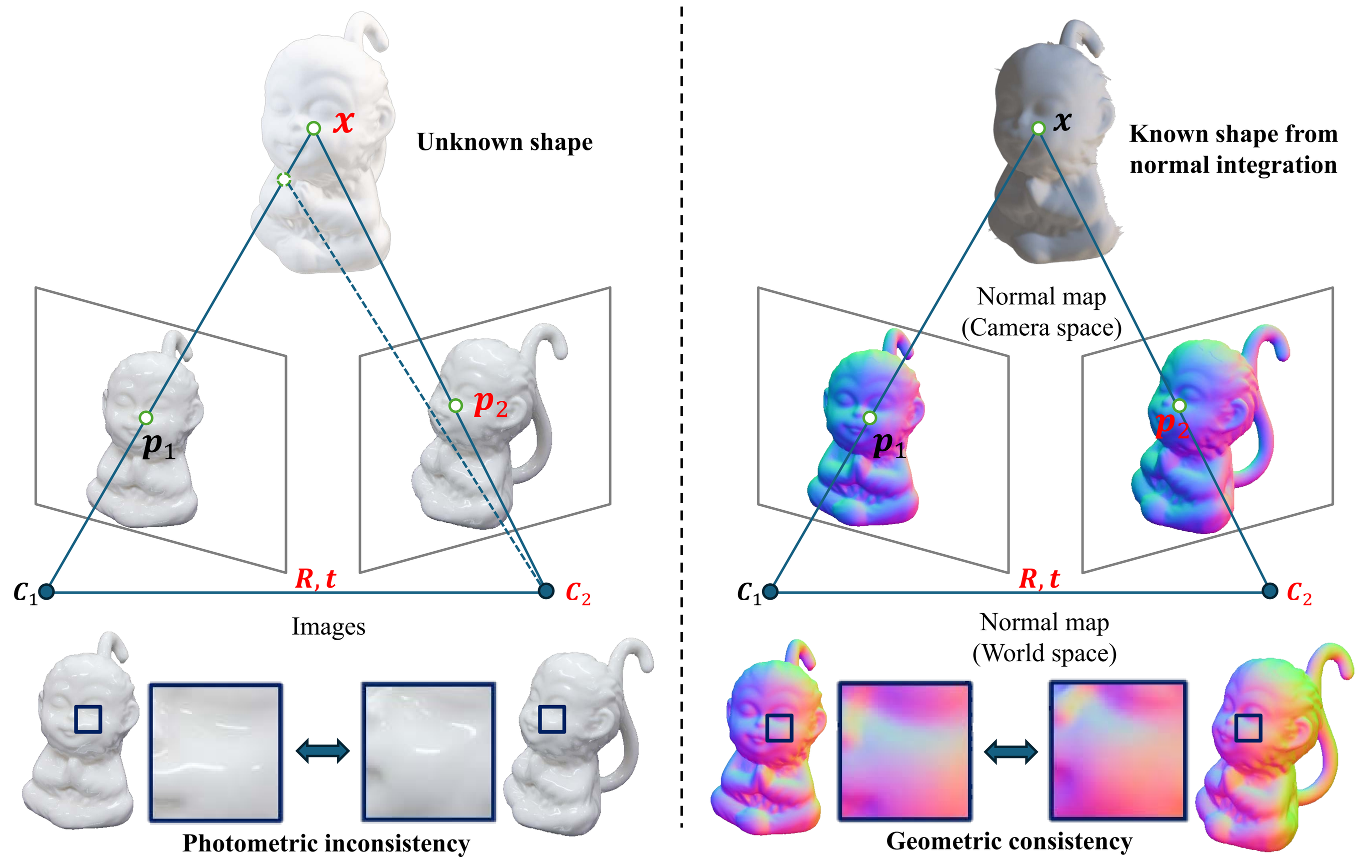
For more information, please visit:
https://pmz-enterprise.github.io/PMNI/

Problem definition



Input: Multi-light Multi-view Images   Normal   Output: Shape & Pose

Motivation & Contribution



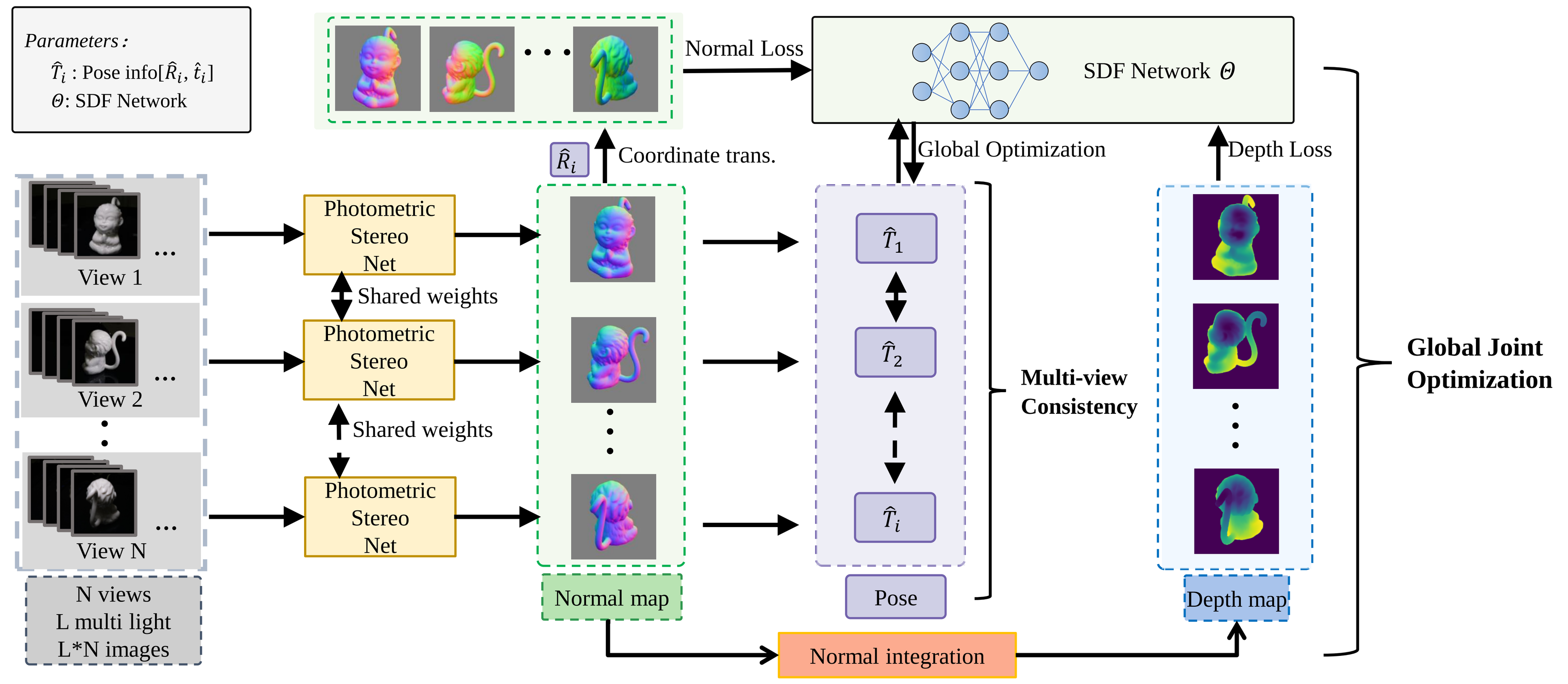
Motivation:

- Reflective and textureless surfaces remain challenging for 3D reconstruction without precise camera pose calibration, as RGB feature extraction and matching fails due to photometric inconsistency.
- Existing methods often require calibration boards, limiting their use in casual-capture setups.
- Surface normal maps maintain geometric consistency, offering a robust alternative to RGB images for shape and pose estimation.

Contribution:

- PMNI is the first method achieving high-quality reflective surface reconstruction without camera pose calibration.
- By leveraging multi-view surface normal maps from photometric stereo, PMNI can jointly optimize both the surface shape and camera poses.

Method



World-to-camera surface normal loss: $\mathcal{L}_{normal} = \sum_{i=1}^{N} \sum_{\mathbf{p}} \left|\mathbf{R}_i \mathbf{n}_i^w(\mathbf{p}) - \mathbf{n}_i^c(\mathbf{p})\right|_2^2$

Normal integration loss: $\mathcal{L}_{ni} = \sum_{i=1}^{N} \left|\mathbf{z}_i^r - \alpha_i \mathbf{z}_i^{ni}\right|$

Multi-view normal consistency loss: $\mathcal{L}_c = \sum_{i}^{N-1} \gamma_i(\mathbf{x}) \left\|\bar{\mathbf{R}}\bar{\mathbf{n}}^c(\mathbf{p}) - \mathbf{R}_i \mathbf{n}_i^c(\pi_i(\mathbf{x}))\right\|_2^2$

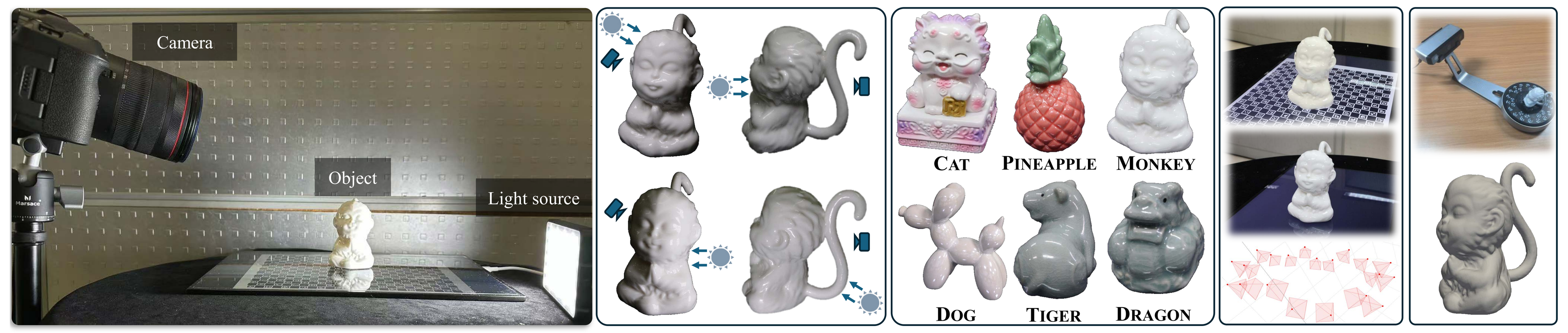
Mask loss: $\mathcal{L}_{mask} = \sum_{i}^{N} \sum_{\mathbf{p}} \mathrm{BCE}\left(\hat{o}_i(\mathbf{p}), o_i(\mathbf{p})\right)$

Eikonal loss: $\mathcal{L}_{eikonal} = \sum_{\mathbf{x}} \left(\|\nabla f(\mathbf{x})\|_2 - 1\right)^2$

Joint optimization of pose and shape:

$$\mathcal{L} = \lambda_0 \mathcal{L}_{normal} + \lambda_1 \mathcal{L}_{ni} + \lambda_2 \mathcal{L}_c + \lambda_3 \mathcal{L}_{eikonal} + \lambda_4 \mathcal{L}_{mask}$$

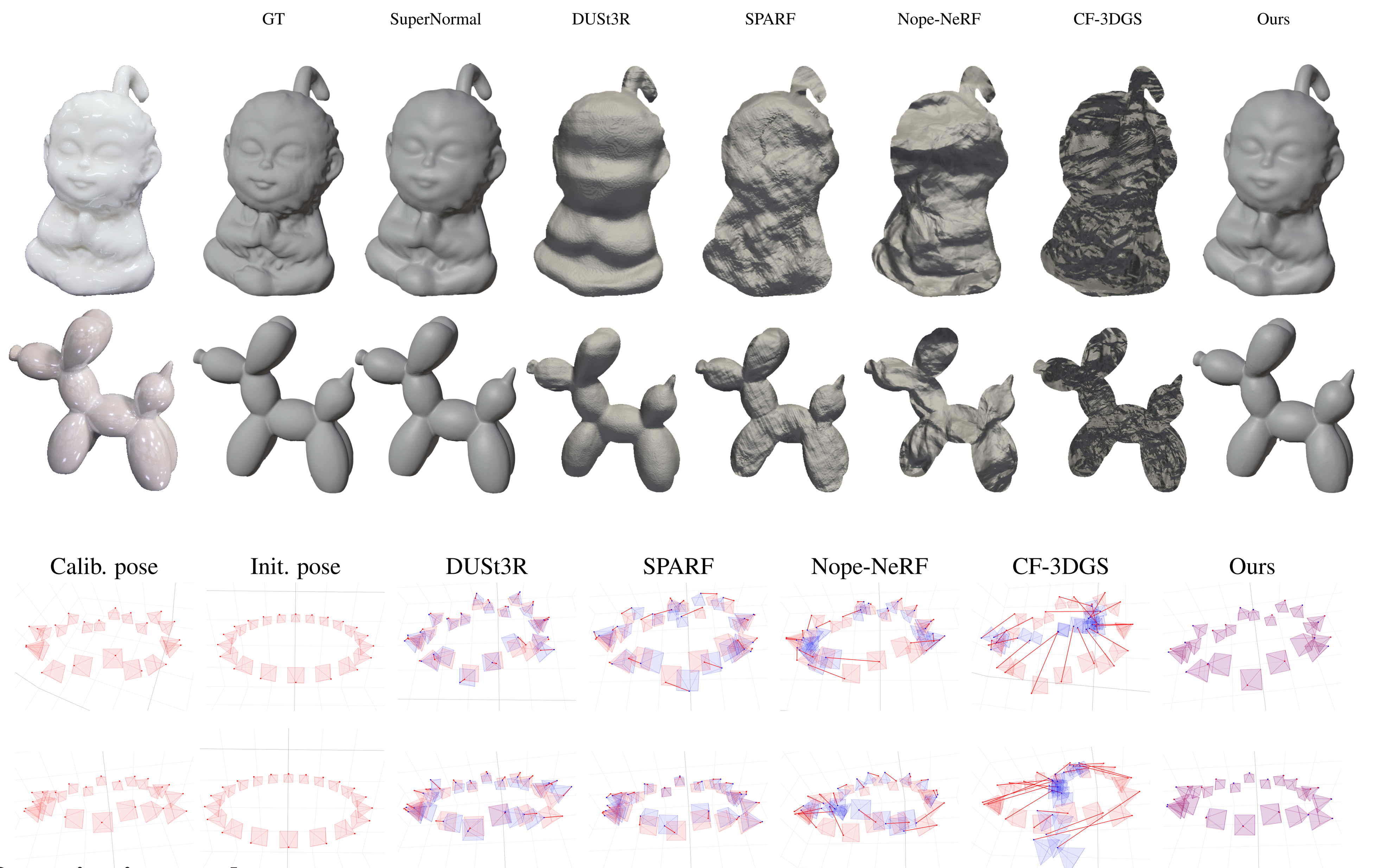
Capture setup & Dataset



Capture setup   Multi-view/Multi-light images   Target objects   Pose calibration   Mesh scan

Experimental results

Qualitative results:



Quantitative results:

Method	RPEr (°) ↓							RPEt ↓							Relative Depth Error ↓						
	Monkey	Cat	Pineapple	Dog	Dragon	Tiger	Avg	Monkey	Cat	Pineapple	Dog	Dragon	Tiger	Avg	Monkey	Cat	Pineapple	Dog	Dragon	Tiger	Avg
DUSt3R	<u>3.175</u>	<u>2.049</u>	<u>2.640</u>	<u>2.216</u>	2.602	4.839	<u>2.920</u>	<u>0.329</u>	<u>0.199</u>	0.247	0.490	0.224	0.335	0.304	<u>0.062</u>	0.056	0.046	0.147	0.046	0.075	0.072
Nope-NeRF	9.371	8.472	7.513	8.674	8.467	8.282	8.463	0.695	0.596	0.610	0.774	0.654	0.637	0.661	0.276	0.191	0.305	0.489	0.231	0.176	0.278
SPARF	7.233	6.395	3.485	3.620	<u>0.731</u>	<u>0.695</u>	3.693	0.375	0.203	<u>0.146</u>	<u>0.261</u>	<u>0.041</u>	<u>0.058</u>	<u>0.181</u>	0.099	<u>0.055</u>	<u>0.038</u>	<u>0.131</u>	<u>0.029</u>	<u>0.050</u>	<u>0.067</u>
CF-3DGS	16.867	16.664	17.276	14.789	15.625	16.659	16.313	0.947	0.796	1.092	0.878	0.998	1.124	0.972	0.363	0.360	0.475	0.488	0.477	0.502	0.444
Ours	**0.230**	**0.356**	**0.258**	**0.258**	**0.439**	**0.582**	**0.354**	**0.015**	**0.020**	**0.016**	**0.019**	**0.027**	**0.035**	**0.022**	**0.011**	**0.017**	**0.008**	**0.010**	**0.011**	**0.026**	**0.014**

Reference:

[1] SuperNormal [Xu Cao *et al.*, CVPR24]
[2] SDM-UniPS [Ikehata *et al.*, CVPR23]
[3] SPARF [Prune Truong *et al.*, CVPR23]
[4] CF-3DGS [Yang Fu *et al.*, CVPR24]
[5] Nope-NeRF [Wenjing Bian *et al.*, CVPR23]
[6] DUSt3R [Shuzhe Wang *et al.*, CVPR24]

Bring-home Message

- We propose PMNI, the first pose-free method for high-quality 3D reconstruction of reflective surfaces using multi-view surface normal maps.
- By jointly optimizing shape and camera poses, our method achieves state-of-art performance without precise calibration.