Supplementary Material of MF-SDF: Neural Implicit Surface Reconstruction using Mixed Incident Illumination and Fourier Feature Optimization

Xueyang Zhou,[1] Xukun Shen[1,2] and Yong Hu[1,2]

[1]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, P.R. China
{by2106166, xkshen, huyong}@buaa.edu.cn
[2]School of New Media Art and Design, Beihang University, Beijing, P.R. China

1. Additional Quantitative Results

Due to space constraints, the text does not give the quantitative indicators corresponding to all visualization results. Here we add the chamfer distance and hausdorff distance between the reconstruction results on the two datasets and the grount-truth mesh. We first calculate the Chamfer distance and the max Hausdorff distance on our dataset. These numerical results can be understood in conjunction with the visualization results of the main text. It can be observed that our method achieves good numerical results in most experiments, indicating that our hybrid lighting representation and shadow-aware photometric loss, combined with the Fourier feature optimization method, can reconstruct consistent, smooth, and detail-rich surfaces on real-world datasets. In addressing highly reflective surfaces, our approach can rival or surpass existing neural inverse rendering reconstruction methods, with even more outstanding performance in detail representation. Compared to current leading high-precision surface detail reconstruction methods, our results are on par with theirs and can handle the reconstruction of highly reflective surface objects, which they cannot process.

	NeuS	NeRO	TensoSDF	Ours
pottery	0.01630	0.01779	0.01750	**0.01135**
blue vase	**0.02801**	0.02977	0.03346	0.03142
blue teapot	0.02643	0.03154	0.04330	**0.02114**
silver teapot	0.02533	0.02583	0.02457	**0.02435**
bronze tripod	0.05450	0.04195	0.06483	**0.04096**
bronze rangette	0.03622	0.02281	0.03039	**0.01866**
black ceramic	0.04036	0.02318	0.02150	**0.02042**

Table 1: *The Chamfer distance↓ results on our dataset,* ***Bold*** *means the best score and underline means the second best*

	NeuS	NeRO	TensoSDF	Ours
pottery	**0.009281**	0.01133	0.01854	0.01034
blue vase	0.02334	**0.02237**	0.04121	0.02297
blue teapot	0.03358	0.02797	0.0559	**0.02642**
silver teapot	0.04728	0.02309	0.02694	**0.02197**
bronze tripod	0.1591	0.1594	0.1849	**0.1551**
bronze rangette	0.1409	0.0200	0.1035	**0.0122**
black ceramic	0.12179	0.04353	0.0407	**0.0267**

Table 2: *The max Hausdorff distance↓ results on our dataset,* ***Bold*** *means the best score and underline means the second best*

	NeRO	TensoSDF	Neuralangelo	Ours
Baking	0.03986	0.05534	0.02995	**0.02631**
Ball	0.01908	0.01930	**0.01074**	0.01941
Blocks	0.03573	0.03466	0.03583	**0.02521**
Car	0.02861	0.03307	**0.01951**	0.02300
Teapot	0.02166	0.02818	0.02197	**0.01684**

Table 3: *Chamfer distance ↓ result on standford dataset,* ***Bold*** *means the best score and underline means the second best*

We subsequently added additional Chamfer Distance numerical results corresponding to those in the original text on the StanfordORB dataset. We still maintain a leading position in most results, but we found that Neuralangelo outperformed us in two results. but it is worth noting that a lower CD does not necessarily mean a smoother surface. The lower the CD value, the more abundant the surface details seem to be. It can be seenthat Neuralangelo achieved the best results in two experiments, but did not produce the most consistent surface(back to the main text for visualization) corresponding to their score.

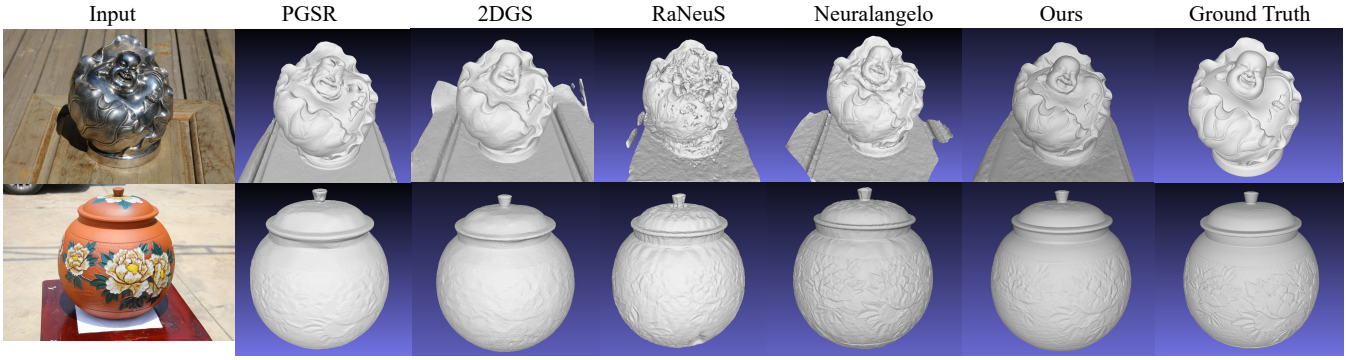

Figure 1: *The reconstruction results of the existing leading methods on two representative data in our dataset.*

2. Comparison with Some Existing Advanced Methods

As we can see from the Figure1, we conducted validation experiments using state-of-the-art multi-view reconstruction methods on our representative dataset, primarily focusing on methods based on Gaussian Splatting (GS) and Multi-resolution Hash Grid (NGP). We found that, on real-world datasets, GS-based methods either struggle with reconstructing highly reflective objects, such as PGSR [CLY*24], or suffer from significant loss of detail, as seen in 2DGS [HYC*24]. And both will produce a stripe like artifact on the surface. While NGP-based methods [WTNT24, LME*23] can reconstruct more prominent details, they also encounter reconstruction errors with highly reflective objects. Therefore, in our main text, we selected Neuralangelo [LME*23] and NeuS [WLL*21] as our baselines—they are not the most recently published but are more reliable.

3. Some Descriptions of Our Dataset

Our data set is collected with the SLR camera (Canon 5DS). We place the object being captured on a flat surface and evenly spaced in a 360-degree circle around the object on the upper sphere, and at three different elevation levels, similar to the StandfordORB dataset [KZY*23]. These objects included those with highly reflective surfaces, and the captured images had a resolution of 8688*5792 pixels. We also use the COLMAP [SF16] to obtain the camera's intrinsic and extrinsic parameters, as in Figure2. For more importantly, we employed the KSCAN-Magic and Artec Space Spider scanners to acquire the ground truth 3D mesh of the object surface, as in Figure3.

4. Some Explanations on the calculation of Spherical Gaussian Illumination

For strong illumination approximated by the Spherical Gaussians, We use the same technique as PhySG [ZLW*21], and also refer to the hemispherical approximation technique for the phase cosine term proposed by Meder et al. [MB18],

$$L_{SG}(\omega_i) = \sum_{k=1}^{M} G(\omega_i; \xi_k, \lambda_k, \mu_k) \tag{1}$$

Figure 2: *We collect data along the loop of the upper hemisphere, and use the COLMAP [SF16] to obtain the camera's intrinsic and extrinsic parameters.*

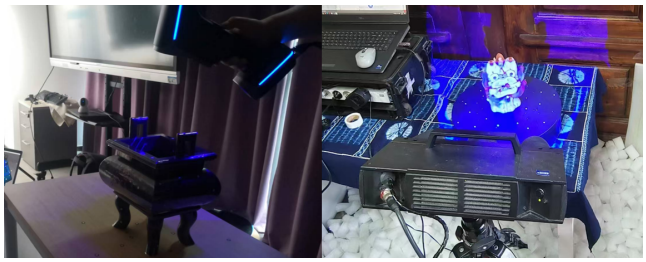

Figure 3: *we use the 3D scanners to obtain the ground truth mesh.*

Since we also need to use split-sum approximation to calculate the interaction process of lighting and surface materials represented by SG (as 2),

$$\mathbf{c}_{\text{specularSG}} \approx \underbrace{\int_{\Omega} L(\omega_i) D(r, \mathbf{t}) d\omega_i}_{L_{\text{SG}}} \cdot \underbrace{\int_{\Omega} \frac{DFG}{4(\omega_o \cdot \mathbf{n})} d\omega_i}_{M_{\text{specular}}} \tag{2}$$

we conduct spherical Gaussian approximation for the distribution term D, and can carry out closed form integration.

$$D(\mathbf{h}) = G(\mathbf{h}; \mathbf{n}, \frac{\lambda}{4\mathbf{h}\cdot\omega_o}, \mu) \tag{3}$$

where $\mathbf{h} = ||(\mathbf{t}+\omega_o)/2||$ known as the half vector.

5. Results using Photometric Consistency Loss without Shadow Aware Treatment

Our shadow aware multi-view consistency loss is based on Geo-NeuS [FXOT22], but actually our inspiration comes from the analysis of the experiments of Geo-NeuS. For Geo-NeuS, the photometric consistency loss is calculated for all points on the surface of the object to be reconstructed. This strategy works well for objects with rich surface texture and a large number of patterns, but for texture-less surfaces and black or reflective surfaces, this algorithm will lead to extremely unstable training. In combination with the SDF loss of sparse point clouds, the gradient disappears many times in the experiment. For experiments that are difficult to converge, the results often have incorrect geometry, and the foreground objects even overlap with the background. As shown in the 4 below. However, our shadow awareness photometric consistency loss

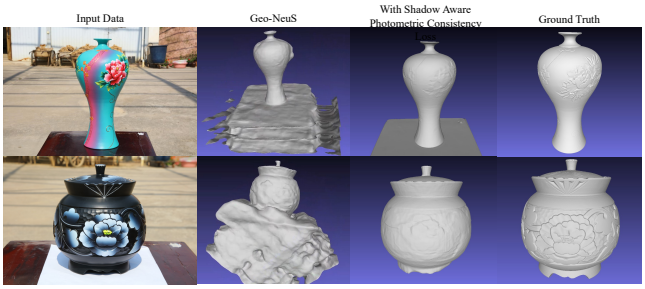

Figure 4: *we use the MeshLab [CMRC] for showing the effects.*

can well avoid this problem. By injecting shadow hints into the 3D neural field, the algorithm can distinguish where the shadow surface needs to be smoothed, rather than the weak texture surface whose material color is black.

6. Other Details

For other details of our method, such as network implementation and other intermediate parameter settings, they will be found in the open-source code after our article is published.

References

[CLY*24] CHEN D., LI H., YE W., WANG Y., XIE W., ZHAI S., WANG N., LIU H., BAO H., ZHANG G.: Pgsr: Planar-based gaussian splatting for efficient and high-fidelity surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics* (2024). 2

[CMRC] CIGNONI P., MUNTONI A., RANZUGLIA G., CALLIERI M.: Meshlab. doi:10.5281/zenodo.5114037. 3

[FXOT22] FU Q., XU Q., ONG Y. S., TAO W.: Geo-neus: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. *Advances in Neural Information Processing Systems 35* (2022), 3403–3416. 3

[HYC*24] HUANG B., YU Z., CHEN A., GEIGER A., GAO S.: 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers* (2024), pp. 1–11. 2

[KZY*23] KUANG Z., ZHANG Y., YU H.-X., AGARWALA S., WU E., WU J., ET AL.: Stanford-orb: a real-world 3d object inverse rendering benchmark. *Advances in Neural Information Processing Systems 36* (2023), 46938–46957. 2

[LME*23] LI Z., MÜLLER T., EVANS A., TAYLOR R. H., UNBERATH M., LIU M.-Y., LIN C.-H.: Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 8456–8465. 2

[MB18] MEDER J., BRÜDERLIN B.: Hemispherical gaussians for accurate light integration. In *Computer Vision and Graphics: International Conference, ICCVG 2018, Warsaw, Poland, September 17-19, 2018, Proceedings* (2018), Springer, pp. 3–15. 2

[SF16] SCHONBERGER J. L., FRAHM J.-M.: Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 4104–4113. 2

[WLL*21] WANG P., LIU L., LIU Y., THEOBALT C., KOMURA T., WANG W.: Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *arXiv preprint arXiv:2106.10689* (2021). 2

[WTNT24] WANG Y., TAN D. J., NAVAB N., TOMBARI F.: Raneus: Ray-adaptive neural surface reconstruction. In *2024 International Conference on 3D Vision (3DV)* (2024), IEEE, pp. 53–63. 2

[ZLW*21] ZHANG K., LUAN F., WANG Q., BALA K., SNAVELY N.: Physg: Inverse rendering with spherical gaussians for physics-based material editing and relighting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 5453–5462. 2