PolGS: Polarimetric Gaussian Splatting for Fast Reflective Surface Reconstruction
Supplementary Material

Yufei Han[1], Bowen Tie[1], Heng Guo[1,2*], Youwei Lyu[1], Si Li[1*], Boxin Shi[3,4], Yunpeng Jia[1], Zhanyu Ma[1]
[1]Beijing University of Posts and Telecommunications [2]Xiong'an Aerospace Information Research Institute
[3]State Key Laboratory of Multimedia Information Processing, School of Computer Science, Peking University
[4]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University
{hanyufei, tiebowen, guoheng, youweilv, lisi, mazhanyu}@bupt.edu.cn
shiboxin@pku.edu.cn xibei156@163.com

A. Polarimetric image formation model

In this section, we provide a more detailed explanation of the polarimetric theory utilized in our method.

According to the polarized Bidirectional Reflectance Distribution Function (BRDF) model [1], the output Stokes vector $\mathbf{s}_o(\mathbf{v})$ can be separated into diffuse and specular components as:

$$\mathbf{s}_o(\mathbf{v}) = \int_\Omega \mathbf{H}_d \mathbf{s}_i(\boldsymbol{\omega})\ d\boldsymbol{\omega} + \int_\Omega \mathbf{H}_s \mathbf{s}_i(\boldsymbol{\omega})\ d\boldsymbol{\omega}. \tag{1}$$

The diffuse and specular Stokes components under a single light source can be formulated respectively as:

$$\mathbf{H}_d \mathbf{s}_i(\boldsymbol{\omega}) = \rho_d L(\boldsymbol{\omega}) \boldsymbol{\omega}^\top \mathbf{n} T_i^+ T_i^- \begin{bmatrix} T_o^+ \\ T_o^- \cos(2\phi_n) \\ -T_o^- \sin(2\phi_n) \\ 0 \end{bmatrix}, \tag{2}$$

$$\mathbf{H}_s \mathbf{s}_i(\boldsymbol{\omega}) = \rho_s L(\boldsymbol{\omega}) \frac{DG}{4\mathbf{n}^\top \mathbf{v}} \begin{bmatrix} R^+ \\ R^- \cos(2\phi_h) \\ -R^- \sin(2\phi_h) \\ 0 \end{bmatrix}. \tag{3}$$

Unlike PANDORA [5] and NeRSP [7], we use the Gaussian Splatting structure to represent the diffuse component instead of a complex implicit Multi-Layer Perceptron (MLP) network. We also use a cubemap encoder followed by 3DGS-DR [9] to predict the environment map and the specular component from different views. Therefore, we define the diffuse color $C = \int_\Omega \rho L(\boldsymbol{\omega}) \boldsymbol{\omega}^\top \mathbf{n} T_i^+ T_i^-\ d\boldsymbol{\omega}$ and specular color $L_r = \rho_s \int_\Omega L(\boldsymbol{\omega}) \frac{DG}{4\mathbf{n}^\top \mathbf{v}}\ d\boldsymbol{\omega}$, and the final $\mathbf{s}_o(\mathbf{v})$ can be represented as:

$$\mathbf{s}_o(\mathbf{v}) = C \begin{bmatrix} T_o^+ \\ T_o^- \cos(2\phi_n) \\ -T_o^- \sin(2\phi_n) \\ 0 \end{bmatrix} + L_r \begin{bmatrix} R^+ \\ R^- \cos(2\phi_h) \\ -R^- \sin(2\phi_h) \\ 0 \end{bmatrix}. \tag{4}$$

To make the diffuse color view-independent, we adjust the order of the spherical harmonics (SH) coefficients to zero, facilitating the radiance decomposition. In Section D, we conduct an experiment to show the differences resulting from the SH coefficient adjustment under different numbers of input views.

B. Additional results on real-world datasets

This section presents more results on real-world datasets tested with different methods.

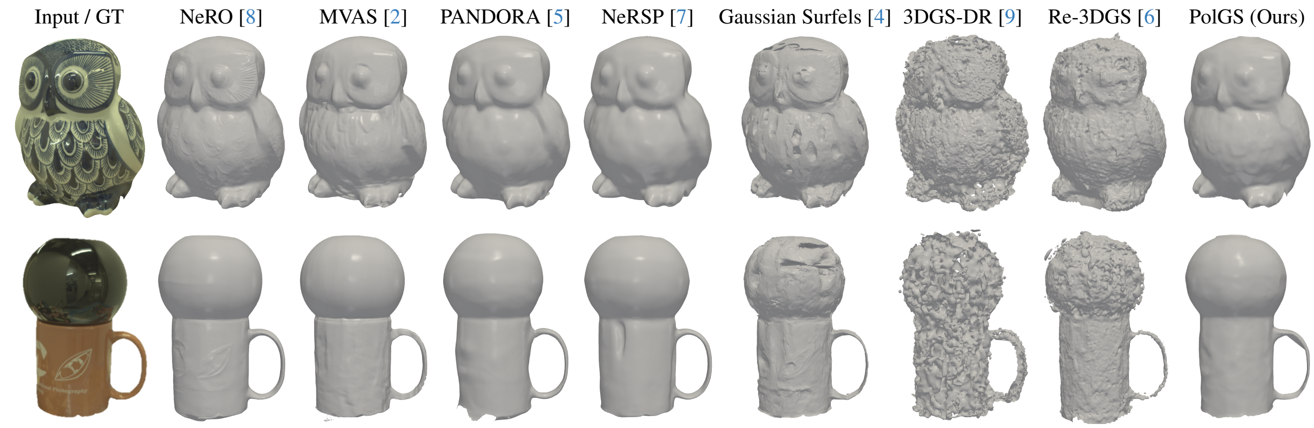

Figure 1. Qualitative shape recoveries on PANDORA [5].

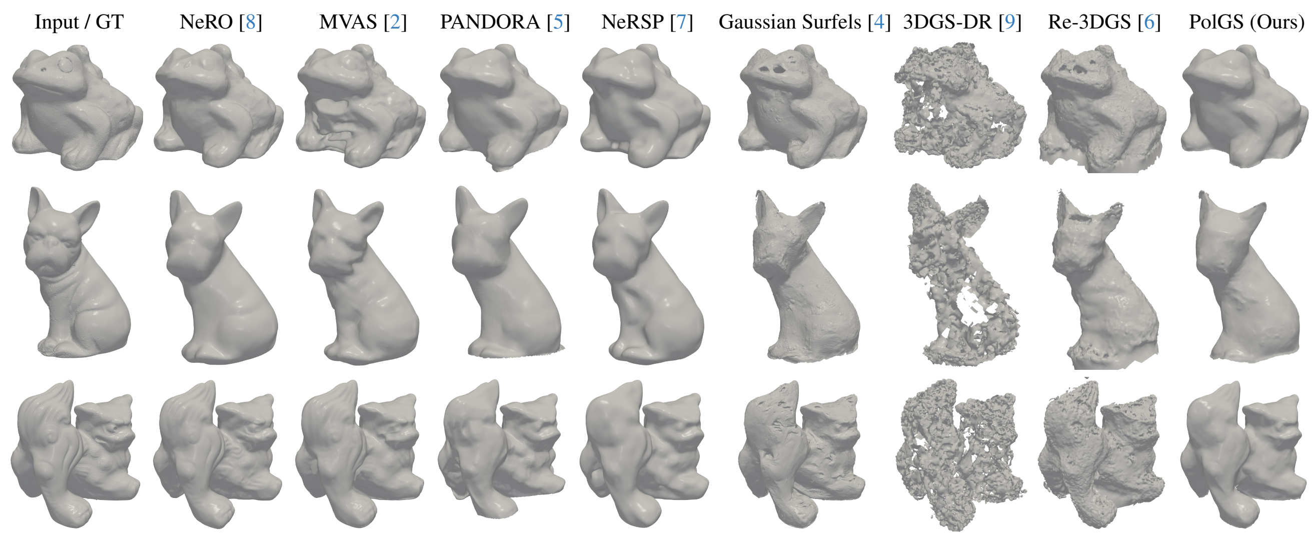

Figure 2. Qualitative shape recoveries on RMVP3D [7].

B.1. Comparison on PANDORA [5]

Figure 1 illustrates mesh reconstruction results using different methods on PANDORA [5]. The qualitative comparison demonstrates that our PolGS achieves superior reconstruction quality compared to other 3DGS-based approaches.

B.2. Comparison on RMVP3D [7]

Figure 2 presents the mesh reconstruction results using different methods on RMVP3D [7]. The qualitative results indicate that our PolGS not only surpasses other 3DGS-based approaches but also achieves performance closer to SDF-based methods in a shorter time.

C. Additional results on synthetic dataset

In this section, we display more results on the synthetic dataset tested with different methods.

C.1. Comparison on SMVP3D [7]

Figure 3 shows the surface normal estimation results using different methods on SMVP3D [7]. The qualitative and quantitative results both demonstrate that our PolGS achieves better reconstruction quality compared to other 3DGS-based approaches.

D. Analysis of SH coefficient adjustment

In the main paper, we display the surface normal prediction before and after the SH (Spherical Harmonics) coefficient adjustment. In Fig. 4, we provide additional experimental results with different input numbers of views, along with SH coefficient adjustment. We show the diffuse and specular separation with SH coefficients of order 0 and 3 with input number views 9 and 18. It can be observed that the ability to perform radiance decomposition decreases as the SH coef-

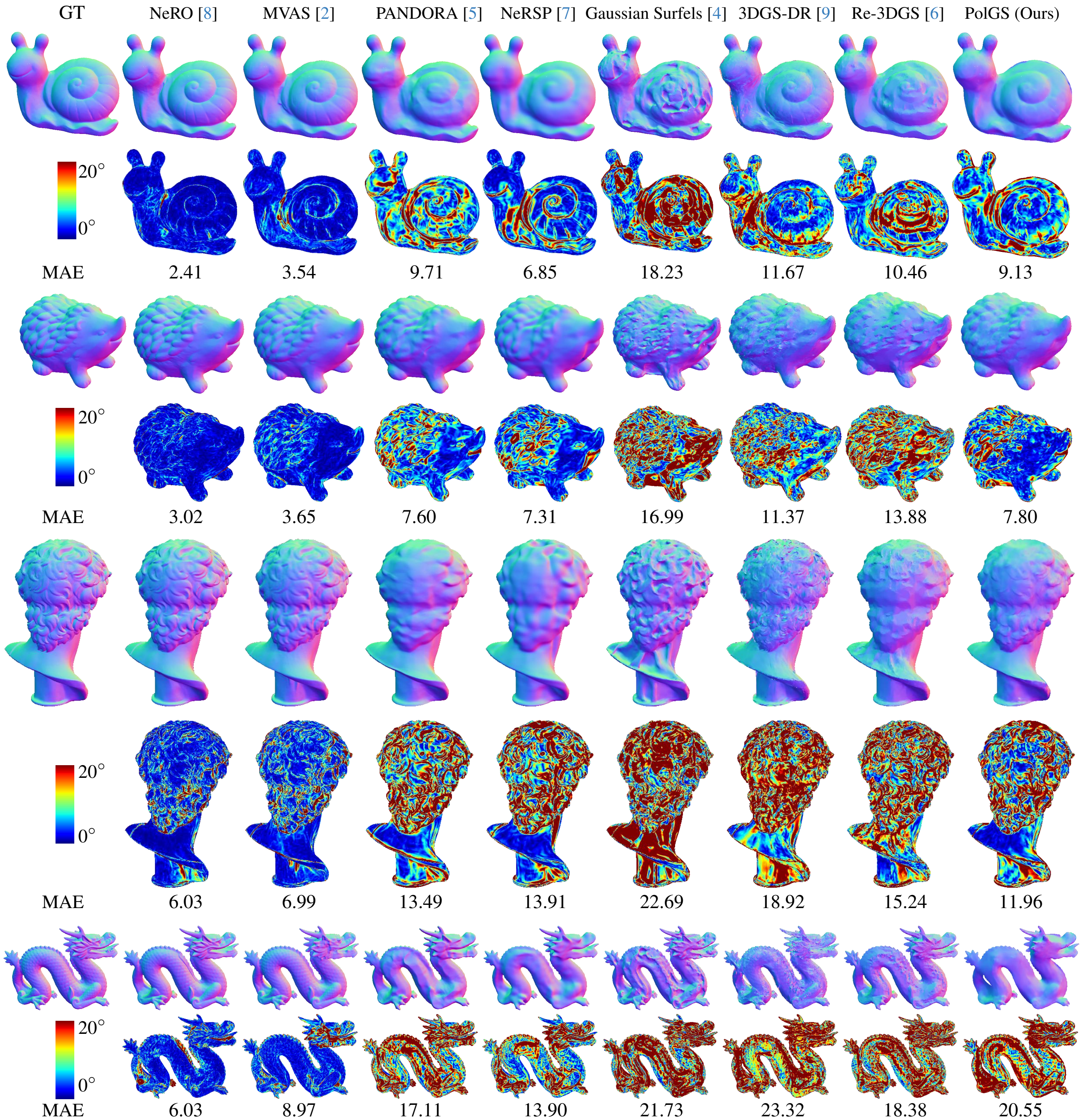

Figure 3. Qualitative comparisons on surface normal estimation of SQUIRREL in SMVP3D [7], where our 3DGS-based method can outperform existing methods based on the same representation and achieves comparable results with SDF-based methods such as NeRSP [7] and PANDORA [5] while with higher efficiency.

ficient order increases, especially when the number of input views is not significantly high. However, with more input views, the use of higher-order spherical harmonic functions tends to better fit the diffuse color, so the difference with the 0th-order results is not as significant. The above phenomenon further demonstrates the reasonableness of our adjustment of the SH coefficients.

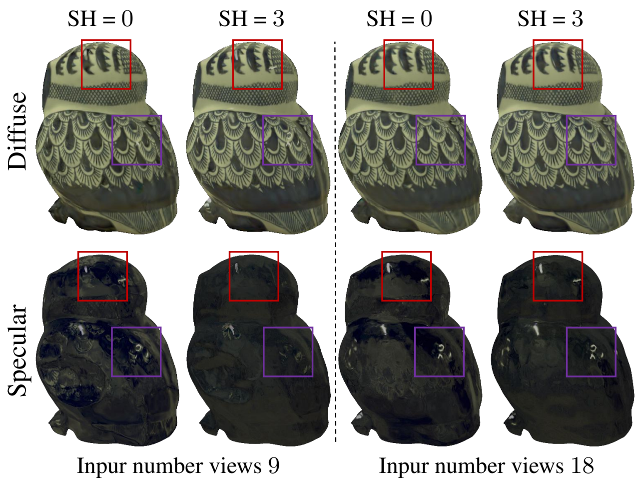

Figure 4. Radiance decomposition results obtained using different spherical harmonic (SH) coefficient orders, specifically 0, 1, and 3.

E. Implementation details

In this section, we introduce more implementation details about experiment settings and our training strategy.

E.1. Experiment settings

We test our model and compare it with other methods on three datasets: one synthetic dataset (SMVP3D [7]) and two real-world datasets (PANDORA [5] and RMVP3D [7]).

SMVP3D [7] This dataset contains five objects, and the input Stokes parameters and images are all at a resolution of 512×512 pixels. We select 36 different views around each object to ensure adequate coverage. The specific input views are displayed in Fig. 5 (a).

PANDORA [5] The PANDORA [5] dataset includes three objects, and we select two of them for testing. The original input resolution is 2048×2448, which we resize to 512×612 during training and evaluation. We use 35 views to train on each object. The specific input views are displayed in Figure 2(b).

RMVP3D [7] Containing four objects, we select three from the RMVP3D [7] dataset for testing. The original input resolution is 1024×1224, resized to 512×612 for consistency during training and evaluation. We employ 35 views to train on each object. The specific input views are displayed in Figure 2(c).

E.2. Training and Evaluation

During the training period, we first perform 1,000 epochs for warm-up. The total number of training epochs is 15,000, meaning we introduce the polarimetric loss $\mathcal{L}_{pol}$ and perform radiance decomposition after the initial 1,000 epochs.

After training, we use the mesh extraction method from Gaussian Surfels [4]. Specifically, we set the depth coefficient to 8 to reduce artifacts from the generated point cloud.

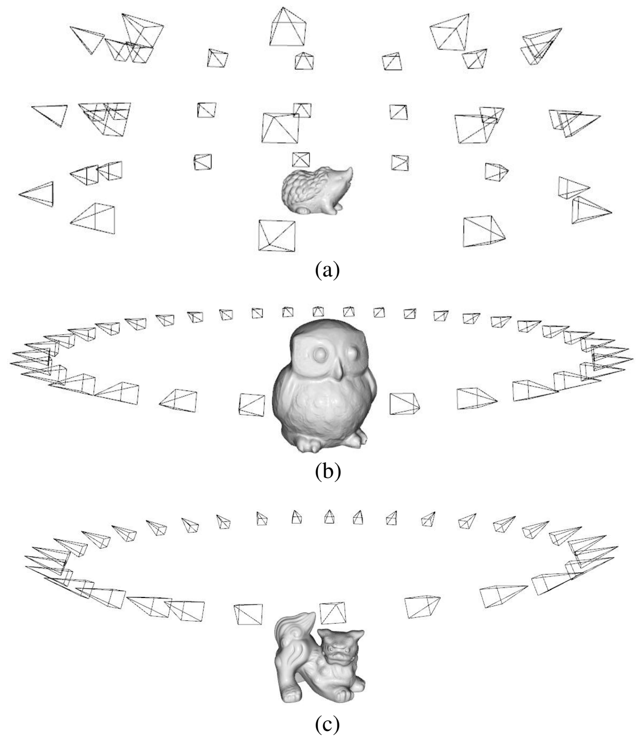

Figure 5. Visualization of view distributions of different datasets. **(a)** Input views display on SMVP3D [7]. **(b)** Input views display on PANDORA [5]. **(c)** Input views display on RMVP3D [7].

F. Failure case

As shown in Fig. 3, the reconstruction results are not satisfactory, especially for the DRAGON model. This is due to the fact that the reconstructed objects are elongated, and the quality of the reconstructed mesh decreases considerably when the input view is not in the form of an equatorial surround as in PANDORA [5] or RMVP3D [7]. Additionally, the number of input views greatly affects the quality of the reconstruction, as can be observed in the 3DGS-based approach,such as 3DGS-DR [9] and Re-3DGS [6]. In their paper, with 100 input images as training data, they are able to recover the normal vectors very well. However, in our experiments, only $30 - 40$ views were used as inputs, so both our results and other 3DGS-based methods are much less effective in shape recovery. There is still a gap between this and SDF-based methods, and this is an area for future improvement.

G. Limitations

Although we utilize the additional polarimetric prior to constrain the surface normal representation, our method relies on assumption of an unpolarized environment light and ignores the inter-reflections among the surfaces. These constraints may limit its practicality in real-world scenarios. Furthermore, the polarimetric information can be compromised by noise in real-world data, potentially affecting the accuracy of our results.

References

[1] Seung-Hwan Baek, Daniel S Jeon, Xin Tong, and Min H Kim. Simultaneous acquisition of polarimetric svbrdf and normals. *ACM Transactions on Graphics*, 2018. 1

[2] Xu Cao, Hiroaki Santo, Fumio Okura, and Yasuyuki Matsushita. Multi-View Azimuth Stereo via Tangent Space Consistency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 2, 3

[3] Guangcheng Chen, Yicheng He, Li He, and Hong Zhang. PISR: Polarimetric neural implicit surface reconstruction for textureless and specular objects. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2024.

[4] Pinxuan Dai, Jiamin Xu, Wenxiang Xie, Xinguo Liu, Huamin Wang, and Weiwei Xu. High-quality surface reconstruction using gaussian surfels. In *ACM SIGGRAPH*, 2024. 2, 3, 4

[5] Akshat Dave, Yongyi Zhao, and Ashok Veeraraghavan. Pandora: Polarization-aided neural decomposition of radiance. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 1, 2, 3, 4

[6] Jian Gao, Chun Gu, Youtian Lin, Zhihao Li, Hao Zhu, Xun Cao, Li Zhang, and Yao Yao. Relightable 3D gaussians: Realistic point cloud relighting with brdf decomposition and ray tracing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2024. 2, 3, 4

[7] Yufei Han, Heng Guo, Koki Fukai, Hiroaki Santo, Boxin Shi, Fumio Okura, Zhanyu Ma, and Yunpeng Jia. NeRSP: Neural 3D Reconstruction for Reflective Objects with Sparse Polarized Images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. 1, 2, 3, 4

[8] Yuan Liu, Peng Wang, Cheng Lin, Xiaoxiao Long, Jiepeng Wang, Lingjie Liu, Taku Komura, and Wenping Wang. NeRO: Neural geometry and brdf reconstruction of reflective objects from multiview images. *ACM Transactions on Graphics*, 2023. 2, 3

[9] Keyang Ye, Qiming Hou, and Kun Zhou. 3D Gaussian Splatting with Deferred Reflection. In *ACM SIGGRAPH*, 2024. 1, 2, 3, 4