

PolGS: Polarimetric Gaussian Splatting for Fast Reflective Surface Reconstruction

Yufei Han¹, Bowen Tie¹, Heng Guo^{1,2*}, Youwei Lyu¹, Si Li^{1*}, Boxin Shi^{3,4}, Yunpeng Jia¹, Zhanyu Ma¹

¹Beijing University of Posts and Telecommunications ²Xiong'an Aerospace Information Research Institute

³State Key Laboratory of Multimedia Information Processing, School of Computer Science, Peking University

⁴National 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

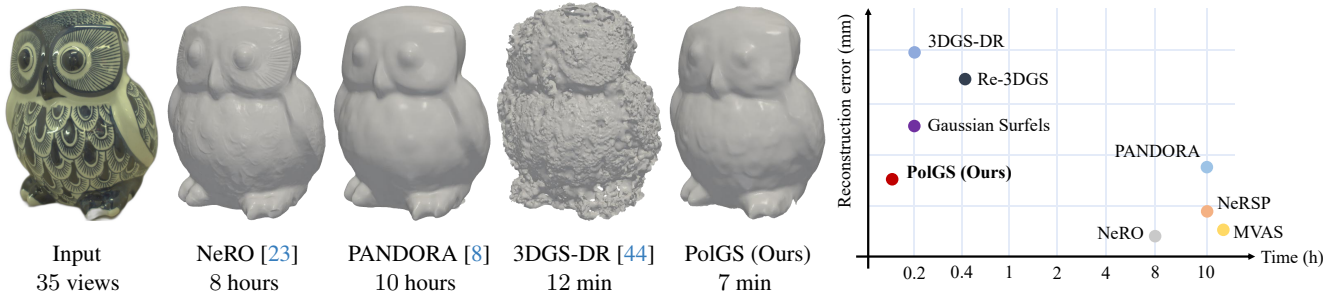


Figure 1. Comparison of efficiency and accuracy on reflective surface reconstruction. Our method takes the shortest time while comparable shape reconstruction accuracy (measured by Chamfer Distance in millimeters) with the existing method based on neural implicit surface representation [23].

Abstract

Efficient shape reconstruction for surfaces with complex reflectance properties is crucial for real-time virtual reality. While 3D Gaussian Splatting (3DGS)-based methods offer fast novel view rendering by leveraging their explicit surface representation, their reconstruction quality lags behind that of implicit neural representations, particularly in the case of recovering surfaces with complex reflective reflectance. To address these problems, we propose PolGS, a Polarimetric Gaussian Splatting model allowing fast reflective surface reconstruction in 10 minutes. By integrating polarimetric constraints into the 3DGS framework, PolGS effectively separates specular and diffuse components, enhancing reconstruction quality for challenging reflective materials. Experimental results on the synthetic and real-world dataset validate the effectiveness of our method. Project page: <https://yu-fei-han.github.io/polgs>.

1. Introduction

Fast and accurate reconstruction of reflective surfaces is essential for applications like real-time virtual reality and in-

verse rendering. Reflective surfaces meet unique challenges in 3D reconstruction as their specular properties require precise handling of light interactions to capture realistic surface details. It is desired to develop a fast and accurate reconstruction method for reflective surfaces.

For reflective surface reconstruction, most existing methods rely on implicit neural representations [28], such as Ref-NeRF [34] and NeRO [23], which model complex surface details effectively but suffer from slow processing times due to their implicit neural network structure. While the signed distance function (SDF) offers a better geometry representation, the MLP network incurs significant time costs. In contrast, the recent 3D Gaussian Splatting (3DGS) [14] technique offers a promising approach for fast novel view rendering through explicit surface representation. However, this representation lacks the same level of detail as implicit methods due to limitations in representing fine geometry and surface normals. Additionally, 3DGS-based reconstruction methods primarily focus on diffuse surfaces, leaving reflective surfaces under-explored and resulting in lower quality reconstructions for these challenging materials.

To address the challenge of reflective surface reconstruction, polarized images are often utilized to enhance shape representation [3, 8, 12, 17]. By decomposing the dif-

* Corresponding authors.

fuse and specular components according to the polarimetric bidirectional reflectance distribution function (pBRDF) model [2], the surface normal can be effectively constrained. However, existing polarization-based methods rely on implicit neural representations, which significantly slow down the reconstruction process.

To achieve high efficiency in reflective surface reconstruction, we propose PolGS, a novel method that integrates polarimetric information into the 3DGS architecture for the first time. Unlike previous approaches, the explicit surface representation and splatting rendering in 3DGS present unique challenges for directly applying polarimetric constraints. To overcome this, we adopt an enhanced 3DGS-based method, Gaussian Surfels [7], as our baseline, which offers better surface representation capabilities. To ensure that each Gaussian kernel retains a view-independent diffuse color, we modify the spherical harmonics (SH) coefficients to zero-order. Next, we introduce a Cubemap Encoder module, inspired by 3DGS-DR [44], to extract the specular component. Finally, we use the polarimetric constraint during the separation of diffuse and specular components [2]. This approach resolves shape ambiguities that arise when relying solely on RGB inputs, enabling more accurate reconstruction of *reflective* surfaces, especially for *texture-less* object.

As shown in the teaser Fig. 1 and Table 1, PolGS achieves reconstruction quality comparable with SDF-based methods but with over 80-times speed improvement (compared to the fastest and best neural-based reconstruction work in our experiment using RGB input [23]), making it highly suitable for virtual reality applications and real-time inverse rendering.

In summary, we advance the reflective surface reconstruction by proposing:

- PolGS, the first method to incorporate polarimetric information into 3DGS, accelerating the progress of surface reconstruction.
- A pBRDF module integrated into 3DGS, which effectively constrains the diffuse and specular components of reflective surfaces, enhancing the accuracy of surface reconstruction.
- Our method achieves reconstruction quality comparable with SDF-based approaches while significantly improving reconstruction time within 10 min. Compared to existing 3DGS methods, PolGS delivers enhancements in reconstruction quality.

2. Related works

Our PolGS aims to reconstruct reflective surface with 3DGS by using polarized images, so we summarize recent progresses in 3D reconstruction techniques, focusing on reflective surfaces through SDF-based methods, 3DGS methods, and polarized image-based reconstruction, respectively.

Table 1. Comparison of different methods in reflective 3D reconstruction: the top four methods are SDF-based, while the bottom four methods are based on 3DGS.

Input	Method	Reflective	Accuracy	time (h)
RGB Images	NeRO [23]	✓	high	8
Azimuths	MVAS [3]	✓	high	11
Pol. Images	PANDORA [8]	✓	medium	10
Pol. Images	NeRSP [12]	✓	high	10
RGB Images	Gaussian Surfels [7]	✗	low	0.2
RGB Images	3DGS-DR [44]	✓	low	0.2
RGB Images	Re-3DGS [10]	✓	low	0.4
Pol. Images	PolGS (Ours)	✓	medium	0.1

3D reconstruction based on neural SDF representation

Novel view synthesis has achieved great success using Neural Radiance Fields (NeRF [28]). Motivated by the structure of the multi-layer perceptron (MLP) network within NeRFs, numerous 3D reconstruction methods have emerged that leverage implicit neural representations to predict object surfaces. Some approaches [31, 42] proposed the signed distance field (SDF) into the neural radiance field, effectively representing the surface as an implicit function. Other works [22, 32, 37, 39, 40, 43] extend it by proposing efficient framework in detailed surface reconstruction.

These methods based on SDF represent complex scenes implicitly, but they typically suffer from high computational demands and are not suitable for real-time applications. Although some methods [22, 40] utilize hash grids and instant-NGP [30] structure, it is still a challenge for them to reconstruct the mesh efficiently facing the reflective surface.

Ref-NeRF [34] uses the Integrated Directional Encoding (IDE) structure to estimate the specular reflection components of the object surface by using predicted roughness, view direction, and surface normals. NeRO [23] improves it by generating the physically-based rendering (PBR) parameters, and NeP [36] can better deal with the glossy surface. TensoSDF [19] combines a novel tensorial representation [4] with the radiance and reflectance field for robust geometry reconstruction. However, these approaches can not avoid high optimization time.

3D reconstruction based on 3DGS 3DGS [14] aims to address limitations of neural radiance fields by representing complex spatial points using 3D Gaussian ellipsoids. However, 3D ellipsoids cannot conform effectively to actual object surfaces, resulting in inaccuracies in shape representation when producing point clouds.

To overcome these drawbacks, several extensions and modifications have been proposed. SuGaR [11] approximates 2D Gaussians with 3D Gaussians, NeuSG [6], GSDF [45] and 3DGSR [24] integrate an extra SDF network for representing surface normals to supervise the Gaussian Splatting geometry. 2D Gaussian Splatting [13]

and Gaussian Surfels [7] have taken a different approach by transforming the 3D ellipsoids into 2D ellipses for modeling. This transformation allows for more refined constraints on depth and normal consistency, addressing the surface approximation issues more effectively. GOF [46] achieves more realistic mesh generation through its innovative opacity rendering strategy. However, these methods do not focus on reflective surface reconstruction.

Re-3DGS [10] associates extra properties, including normal vectors, BRDF parameters, and incident lighting from various directions to make photo-realistic relighting. 3DGS-DR [44] presents a deferred shading method to effectively render specular reflection with Gaussian splatting. Despite these advancements, these methods still face challenges in many scenarios and cannot provide accurate geometric expression.

3D reconstruction using polarized images Polarized images are widely used in Shape from Polarization (SfP) [1, 2, 9, 15, 16, 26, 27, 29, 33, 41], reflection removal [20, 25, 38], and some downstream tasks [21, 48] due to the strong physics-preliminary information in the Stokes field. The SfP task aims to predict the surface normal captured by the polarization camera under the single distant light [26, 33] or unknown ambient light [1, 15]. Multi-view 3D reconstruction works using polarized images [47] try to settle down the π and $\pi/2$ ambiguities with the Angle of Polarization (AoP). PANDORA [8] first uses polarized images in neural 3D reconstruction work, following the relevant constraints of pBRDF [2]. MVAS [3] leverages multi-view AoP maps to generate tangent spaces for surface points during the optimization process, which can reconstruct mesh without rendering supervision. NeRSP [12] combines the photometric and geometric cues from polarized images and generates better results under sparse views for reflective surfaces. PISR [5] focuses on texture-less specular surface and integrates the multi-resolution hash grid for efficiency. NeISFs [17, 18] consider the inter-reflection and models multi-bounce polarized light paths during rendering. Despite these advancements, computational cost remains a significant limitation for many polarized-based 3D reconstruction methods.

3. Preliminaries

3.1. Gaussian surfels model

Our PolGS selects Gaussian Surfels [7] as our base framework due to its stronger geometry expression ability. According to Gaussian Surfels [7], we use a set of unstructured Gaussian kernels $\{G_i = \{\mathbf{x}_i, \mathbf{s}_i, \mathbf{r}_i, o_i, C_i\} | i \in \mathcal{N}\}$ to represent the structure of 3DGS, where $\mathbf{x}_i \in \mathbb{R}^3$ denotes the center position of each Gaussian kernel, $\mathbf{s}_i = [s_i^x, s_i^y, 0]^\top \in \mathbb{R}^3$ is the scaling factors of x and y axes after flattening the 3D Gaussians [14], $\mathbf{r}_i \in \mathbb{R}^4$ is the rotation quaternion, $o_i \in \mathbb{R}$

is the opacity, and $C_i \in \mathbb{R}^k$ represents the spherical harmonic coefficients of each Gaussian. And Gaussian distribution can be defined by the covariance matrix Σ of a 3D Gaussian as:

$$G(\mathbf{x}; \mathbf{x}_i, \Sigma_i) = \exp(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_i)^\top \Sigma_i^{-1}(\mathbf{x} - \mathbf{x}_i)), \quad (1)$$

where Σ_i can be represented as:

$$\begin{aligned} \Sigma_i &= \mathbf{R}(\mathbf{r}_i) \mathbf{s}_i \mathbf{s}_i^\top \mathbf{R}(\mathbf{r}_i)^\top \\ &= \mathbf{R}(\mathbf{r}_i) \text{Diag}[(s_i^x)^2, (s_i^y)^2, 0] \mathbf{R}(\mathbf{r}_i)^\top, \end{aligned} \quad (2)$$

where $\text{Diag}[\cdot]$ indicates a diagonal matrix and $\mathbf{R}(\mathbf{r}_i)$ is a 3×3 rotation matrix represented by $\mathbf{r}_i$.

Gaussian splatting According to 3D GS [14], novel view rendering process can be represented as:

$$C = \sum_{i=0}^n T_i \alpha_i c_i, \quad (3)$$

where $T_i = \prod_{j=0}^{i-1} (1 - \alpha_j)$ is the transmittance, $\alpha_i = G'(\mathbf{u}; \mathbf{u}_i, \Sigma'_i) o_i$ is alpha-blending weight, which is the product of opacity and the Gaussian weight based on pixel $\mathbf{u}$. In order to speed up the rendering process, the 3D Gaussian in Eq. (2) is re-parameterized in 2D ray space [50] as G' :

$$G'(\mathbf{u}; \mathbf{u}_i, \Sigma'_i) = \exp(-\frac{1}{2}(\mathbf{u} - \mathbf{u}_i)^\top \Sigma'^{-1}_i(\mathbf{u} - \mathbf{u}_i)), \quad (4)$$

where $\Sigma'_i = (\mathbf{J}_k \mathbf{W}_k \Sigma_i \mathbf{W}_k^\top \mathbf{J}_k^\top)[2, : 2]$. The $\mathbf{W}_k$ is a viewing transformation matrix for input image k and $\mathbf{J}_k$ is the affine approximation of the projective transformation. Σ'_i represents the covariance matrix in the 2D ray space.

The depth $\tilde{D}$ and normal $\tilde{N}$ for each pixel can also be calculated via Gaussian splatting and alpha-blending:

$$\tilde{D} = \frac{1}{1 - T_{n+1}} \sum_{i=0}^n T_i \alpha_i d_i(\mathbf{u}), \quad (5)$$

$$\tilde{N} = \frac{1}{1 - T_{n+1}} \sum_{i=0}^n T_i \alpha_i \mathbf{R}_i[:, 2]. \quad (6)$$

Specifically, according to Gaussian Surfels [7], the depth of pixel $\mathbf{u}$ for each Gaussian kernel i is computed by calculating the intersection of the ray cast through pixel $\mathbf{u}$ with the Gaussian ellipse during splatting. So $d_i(\mathbf{u})$ can be represented by Taylor expansion as:

$$d_i(\mathbf{u}) = d_i(\mathbf{u}_i) + (\mathbf{W}_k \mathbf{R}_i)[2, :] \mathbf{J}_{pr}^{-1}(\mathbf{u} - \mathbf{u}_i), \quad (7)$$

where $\mathbf{J}_{pr}^{-1}$ is the Jacobian inverse mapping one pixel from image space to tangent plane of the Gaussian surfel as in [49], and $(\mathbf{W}_k \mathbf{R}_i)$ transforms the rotation matrix of a Gaussian surfel to the camera space.

3.2. Polarimetric image formation model

Polarization cameras can capture polarized images in a single shot, which can be represented as 4 different polarized angles of images $I = [I_0, I_{45}, I_{90}, I_{135}]$. The Stokes vector $S = \{s_0, s_1, s_2, s_3\}$ can be calculated by:

$$S = \begin{bmatrix} \frac{1}{2}(I_0 + I_{45} + I_{90} + I_{135}) \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ 0 \end{bmatrix}, \quad (8)$$

where we assume the light source is no circularly polarized here, thus the s_3 is 0.

According to PANDORA [8] and NeRSP [12], we assume the incident environmental illumination is unpolarized, the Stokes vector for the incident light direction ω can be expressed as:

$$\mathbf{s}_i(\omega) = L(\omega)[1, 0, 0, 0]^\top, \quad (9)$$

where $L(\omega)$ represents the light intensity. The polarization camera captures outgoing light that becomes partially polarized due to reflection, which is modeled using a 4×4 Muller matrix $\mathbf{H}$. The outgoing Stokes vector $\mathbf{s}$ then is formulated as the integral of the incident Stokes vector multiplied by this Mueller matrix:

$$\mathbf{s}(\mathbf{v}) = \int_{\Omega} \mathbf{H} \mathbf{s}_i(\omega) d\omega, \quad (10)$$

where $\mathbf{v}$ indicates the view direction and Ω is the integral domain. Following the polarized BRDF (pBRDF) model [2], the output Stokes vector can be divided into diffuse and specular components, represented by $\mathbf{H}_d$ and $\mathbf{H}_s$ respectively. $\mathbf{s}$ can be represented as:

$$\mathbf{s}(\mathbf{v}) = \int_{\Omega} \mathbf{H}_d \mathbf{s}_i(\omega) d\omega + \int_{\Omega} \mathbf{H}_s \mathbf{s}_i(\omega) d\omega. \quad (11)$$

Based on the derivations from PANDORA [8] and NeRSP [12], the output Stokes vector can be further specified as:

$$\mathbf{s}(\mathbf{v}) = L_d \begin{bmatrix} T_o^+ \\ T_o^- \cos(2\phi_n) \\ -T_o^- \sin(2\phi_n) \\ 0 \end{bmatrix} + L_s \begin{bmatrix} R^+ \\ R^- \cos(2\phi_h) \\ -R^- \sin(2\phi_h) \\ 0 \end{bmatrix}, \quad (12)$$

where $L_d = \int_{\Omega} \rho L(\omega) \omega^\top \mathbf{n} T_i^+ T_i^- d\omega$ denotes the diffuse radiance associated with the surface normal $\mathbf{n}$, Fresnel transmission coefficients [2] $T_{i,o}^+$ and $T_{i,o}^-$. The diffuse albedo is represented by ρ , and ϕ_n is the azimuth angle of the incident light.

Similarly, $L_s = \int_{\Omega} L(\omega) \frac{DG}{4\mathbf{n}^\top \mathbf{v}} d\omega$ denotes the specular radiance, which involves Fresnel reflection coefficients [2] R^+ and R^- , and the incident azimuth angle ϕ_h concerning

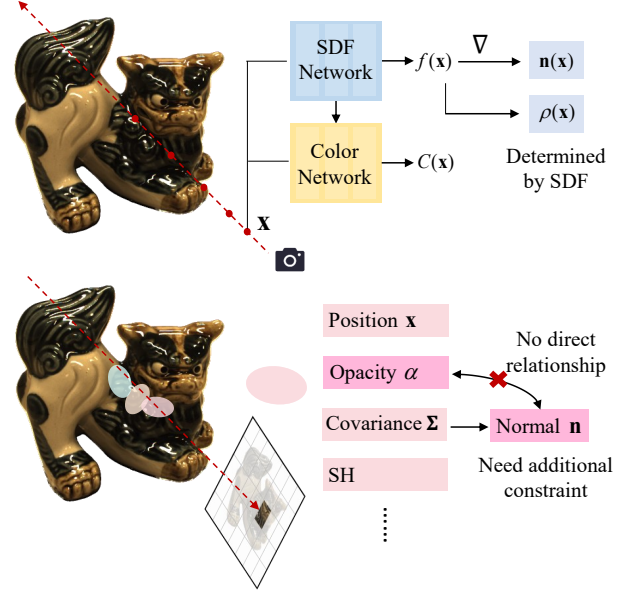


Figure 2. Comparison between SDF-based and 3DGS-based method on geometry representation. **(Top)** The surface normal of a point has a strong relationship with its opacity in NeuS [37]. **(Bottom)** The surface normal of a point is dependent on its opacity in Gaussian Surfels [7].

the half vector $\mathbf{h} = \frac{\omega + \mathbf{v}}{\|\omega + \mathbf{v}\|_2}$. For the sake of simplification in our model, we assume that ϕ_h is equivalent to ϕ_n . The Microfacet model incorporates the normal distribution and shadowing terms represented by D and G [35].

4. Proposed method

Our PolGS framework integrates multi-view polarized images, their corresponding masks, and camera pose information to produce a rich output that includes diffuse and specular components represented as Stokes vectors across various views, a reconstructed geometric mesh, and estimated environment light. In this section, we will analyze the surface normal representations employed in SDF-based approaches and 3DGS methods. Then we introduce a theoretical foundation for our polarization-guided reconstruction pipeline.

4.1. Analysis of surface normal representation

In implicit neural networks, they use signed distance to determine if points are on the surface and the surface normals are obtained by calculating gradients of the SDF. However, 3DGS methods often treat the surface normal as inherent properties typically. In this section, we analyze the differences between the surface normal representations in SDF-based methods and 3DGS, explaining the reason we use polarimetric cues in the Gaussian Splatting method.

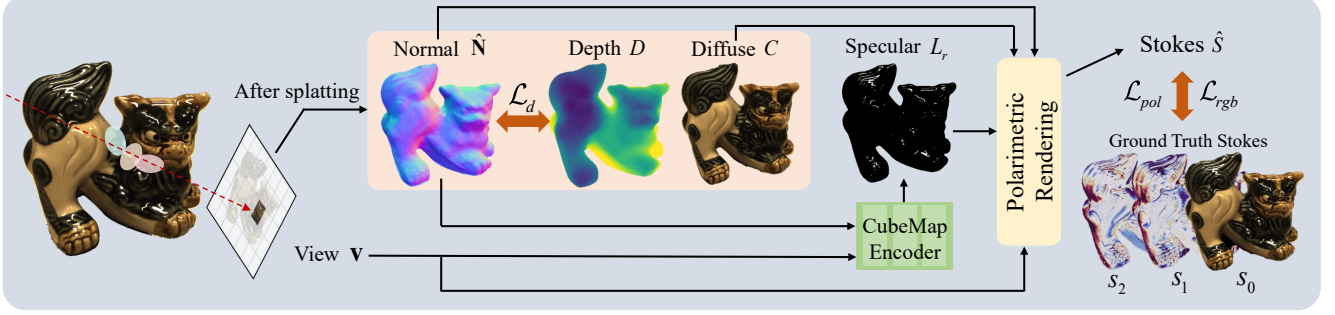


Figure 3. Pipeline of PolGS. We re-rendered Stokes vectors $\hat{s}$ by using the diffuse color C from 3DGS and specular color L_r from Cubemap encoder module, which is supervised by the ground truth Stokes information.

SDF-based representation The implicit neural network approach, exemplified by NeuS [37], enhances the constraints on the reconstructed surface by incorporating a Signed Distance Function (SDF). As illustrated at the top of Fig. 2, the SDF value $f(\mathbf{x})$, generated by the SDF Network, can be used to compute the normal vector $\mathbf{n}(\mathbf{x})$ at point $\mathbf{x}$. According to the definition, the opacity $\rho(\mathbf{x})$ of a point $\mathbf{x}$ is given by:

$$\rho(\mathbf{x}) = \max \left(\frac{-\frac{d\Phi_s}{dx}(f(\mathbf{x}))}{\Phi_s(f(\mathbf{x}))}, 0 \right), \quad (13)$$

where Φ_s is a sigmoid function.

During optimization, the surface normal and the opacity at a given point influence each other. Consequently, after optimization, points farther from the surface tend to have lower opacity values, while the final normal vector is predominantly determined by points located on the surface. This mutual interaction results in a more precise and accurate representation of the surface.

3DGS-based representation In the case of Gaussian Surfels [7], surface normal is derived from the vector perpendicular to the plane of a 2D ellipsoid. As shown at the bottom of Fig. 2, for a single Gaussian kernel, there is no inherent relationship between opacity and the surface normal.

The pixel-level constraint in 3DGS creates a probabilistic representation where no single Gaussian kernel is definitively assigned to a surface. Consequently, multiple Gaussian kernel configurations can potentially represent equivalent surface geometries, introducing inherent reconstruction ambiguity. The optimization mechanism effectively blends contributions from multiple Gaussian points, which offers reconstruction adaptability but compromises the precision of individual point geometric characterizations.

It is obvious that using the normal prediction model as a prior constraint in Gaussian Splatting benefits object reconstruction, as demonstrated by Gaussian Surfels [7]. However, the normal prediction model is not always reliable especially in reflective cases, highlighting the need for a more robust method to provide prior information.

Polarimetric information for reflective surfaces Reflective surface reconstruction is challenging due to the view-dependent appearance. Unlike traditional RGB image inputs, polarimetric information can effectively constrain the surface normal during the rendering of Stokes Vectors with the pBRDF model [2]. Specifically, as demonstrated in Eq. (12), the diffuse and specular components of s_1 and s_2 exhibit a strong correlation with the object’s surface normals. Leveraging this property, we incorporate polarimetric information as a prior in shape reconstruction and use the 3DGS-based method to accelerate this process.

4.2. PolGS

Network structure In this section, we introduce our proposed PolGS, which is a novel 3D reconstruction method that integrates the Gaussian Splatting method with polarimetric information. The network structure of PolGS is shown in Fig. 3. The framework comprises two primary components: the Gaussian Surfels module and the CubeMap Encoder module. Initially, the Gaussian Surfels module is employed to estimate the diffuse component of the object, seems like Gaussian Surfels [7]. Subsequently, we utilize the CubeMap Encoder to assess the specular component, akin to the approach taken in 3DGS-DR [44]. While the CubeMap Encoder does not provide the roughness component, it effectively handles reflective or rough surfaces and maintains high computational efficiency due to its CUDA-based implementation. To enhance the rendering process, we incorporate a pBRDF model into the rendering formulation. This addition introduces a polarimetric constraint that further refines the Gaussian Splatting method, enabling more accurate and realistic 3D reconstructions. The final rendering formulation model following Eq. (12) can be represented as:

$$\hat{S} = C \begin{bmatrix} T_o^+ \\ T_o^- \cos(2\phi_n) \\ -T_o^- \sin(2\phi_n) \\ 0 \end{bmatrix} + L_r \begin{bmatrix} R^+ \\ R^- \cos(2\phi_h) \\ -R^- \sin(2\phi_h) \\ 0 \end{bmatrix}, \quad (14)$$

where C is the diffuse color after Gaussian Surfels rendering according to Eq. (3) and L_r is the specular color after CubeMap rendering $E(\cdot)$, which can be represented as $L_r = E(2(\mathbf{v} \cdot \mathbf{n})\mathbf{n} - \mathbf{v})$.

Adjustment of spherical harmonic coefficients Due to Eq. (14), the diffuse color rendered by Gaussian Surfels must remain consistent across different viewing directions. To satisfy this constraint, we adjust the spherical harmonic (SH) coefficients of the Gaussian Surfels to zero-order, ensuring the splatting results are view-independent.

4.3. Training

The overall training loss in PolGS formulated as a comprehensive weighted sum of multiple loss components:

$$\mathcal{L} = \mathcal{L}_{rgb} + \lambda_1 \mathcal{L}_{pol} + \lambda_2 \mathcal{L}_m + \lambda_3 \mathcal{L}_o + \lambda_4 \mathcal{L}_d, \quad (15)$$

where we set $\lambda_1 = 1$, $\lambda_2 = 0.1$, $\lambda_3 = 0.01$, $\lambda_4 = 0.01 + 0.1 \cdot (\text{iteration}/15000)$ to balance the loss function.

Rendering Stokes loss $\mathcal{L}_{rgb}$ and $\mathcal{L}_{pol}$ The rendering Stokes loss is combined with the s_0 (unpolarized image) rendering loss as 3DGS [14] and the s_1, s_2 rendering loss as PANDORA [8] [2]. These two loss functions can be represented as:

$$\mathcal{L}_{rgb} = 0.8 \cdot L_1(s_0, \hat{s}_0) + 0.2 \cdot L_{DSSIM}(s_0, \hat{s}_0), \quad (16)$$

$$\mathcal{L}_{pol} = L_1(s_1, \hat{s}_1) + L_1(s_2, \hat{s}_2). \quad (17)$$

Mask loss $\mathcal{L}_m$ The mask loss is used to make the rendering results of the object more accurate, which can be represented as:

$$\mathcal{L}_m = \Sigma \text{BCE}(\mathbf{M}, \hat{\mathbf{M}}). \quad (18)$$

Opacity loss $\mathcal{L}_o$ The opacity loss follows Gaussian Surfels [7] to encourage the opacity of the Gaussian points to be close to 1 or 0. It can be represented as:

$$\mathcal{L}_o = \Sigma \exp(-20(o_i - 0.5)^2). \quad (19)$$

Depth-normal consistency loss $\mathcal{L}_d$ The depth-normal consistency loss follows Gaussian Surfels [7] to make the rendered depth and normal of the object to be more consistent. It can be represented as:

$$\mathcal{L}_d = 1 - \hat{\mathbf{N}} \cdot N(V(\hat{\mathbf{D}})), \quad (20)$$

where $V(\cdot)$ transforms each pixel and its depth to a 3D point and $N(\cdot)$ calculates the normal from neighboring points using the cross product.

5. Experiments

To evaluate the performance of our method, we conduct these experiments: 1) Shape reconstruction on synthetic dataset, 2) Shape reconstruction on real-world dataset, 3) Radiance decomposition and 4) Ablation study of polarimetric constraint adding.

Dataset We use three datasets to evaluate our method: the synthetic dataset SMVP3D [12], the real-world dataset RMVP3D [12], PANDORA [8] and PISR [5], where PANDORA [8] can only be used for qualitative evaluation due to the lack of ground truth.

Baselines We compare our method with state-of-the-art techniques, including SDF-based methods NeRO [23], MVAS [3], PANDORA [8], NeRSP [12], and 3DGS methods Gaussian Surfels [7], 3DGS-DR [44], and Re-3DGS [10]. All of these methods, except Gaussian Surfels [7], can handle reflective surfaces. MVAS [3], PANDORA [8] and NeRSP [12] utilize polarimetric information to reconstruct shapes. Specifically, we do not add the normal prior in Gaussian Surfels [7] among the whole experiments.

Evaluation metrics We use the Chamfer Distance (CD) to evaluate the shape recovery performance and the mean angular error (MAE) to evaluate the quality of surface normal estimations. 3DGS-based methods provide point cloud results and we use the Poisson surface reconstruction to generate the mesh for the evaluation.

Implementation details We conduct the experiments on an NVIDIA RTX 4090 GPU with 24GB memory. The training of our model is implemented in PyTorch 1.12.1 using the Adam optimizer. The specific learning rates of different components in Gaussian kernels are set the same as Gaussian Surfels [7]. We use the warm-up strategy to train the model, where adding the polarimetric information and deferred rendering after 1000 iterations.

5.1. Reconstruction results on synthetic dataset

In Fig. 4 and Table 2, we compare the shape recovery performance of various methods on the SMVP3D [12] dataset, which contains five objects with spatially varying and reflective properties. However, it is worth noting that SDF-based methods, such as NeRO [23], MVAS [3], PANDORA [8], and NeRSP [12], outperform 3DGS methods in terms of surface representation. This advantage is largely due to their superior ability to model complex surface details. Among the 3DGS methods, Gaussian Surfels [7] struggles with reflective surfaces, while 3DGS-DR [44] and Re-3DGS [10] still has difficulty in representation of surface normal accurately. In contrast, PolGS effectively leverages polarimetric information, significantly improving geometric surface performance. Due to the inadequate point cloud generation by other 3DGS methods, they fail to produce reasonable mesh results using Poisson surface reconstruction. Our method, however, achieves the lowest mean Chamfer Distance across the synthetic SMVP3D [12] dataset, reinforcing the trends observed in surface normal estimations and confirming that our approach delivers the best performance among the 3DGS techniques.

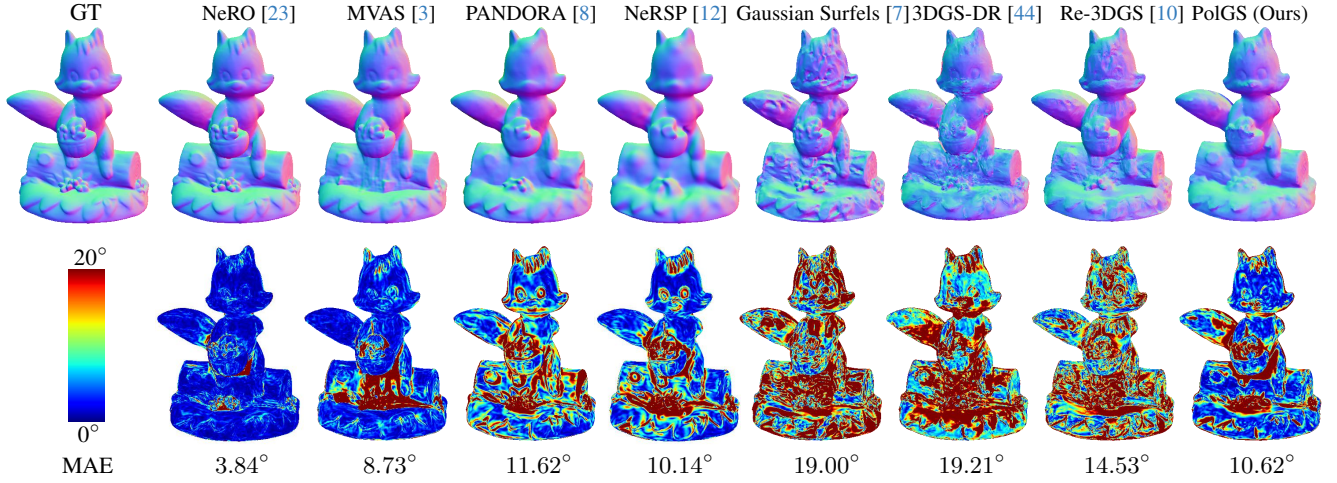


Figure 4. Qualitative comparisons on surface normal estimation of SQUIRREL in SMVP3D [12], where our 3DGS-based method can outperform existing methods based on the same representation and achieves comparable results with SDF-based methods such as NeRSP [12] and PANDORA [8].

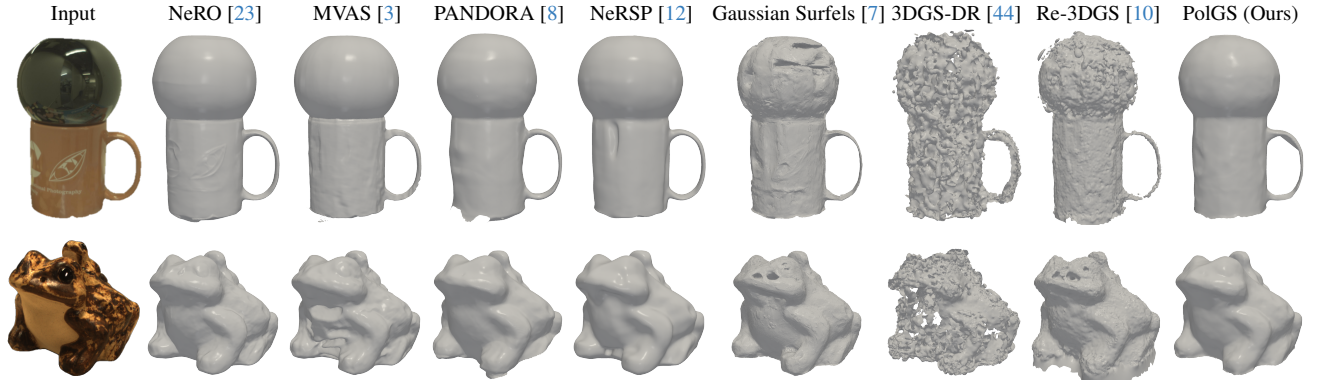


Figure 5. Qualitative comparison shape on PANDORA [8] and RMVP3D [12], where PolGS (Ours) produces similar quality of reconstruction mesh compared to SDF-based methods.

5.2. Reconstruction results on real-world dataset

We further evaluate the reconstruction performance on the RMVP3D [12] and PANDORA [8] datasets, with qualitative results illustrated in Fig. 5 and quantitative results detailed in Table 2. In real-world scenarios, our method produces reconstructions that align more closely with SDF-based approaches, while significantly outperforming other 3DGS-based methods in terms of reconstruction quality. For example, in the VASE case, our approach accurately estimates the shape of a ceramic surface in just 7 minutes, achieving results closer to those generated by SDF-based methods. Additionally, the FROG sample highlights our method’s ability to reconstruct objects with rough glossy surfaces, showing the robustness and generalization capability of our method. These results collectively demonstrate the effectiveness of our approach in handling diverse real-world objects with varying surface properties.

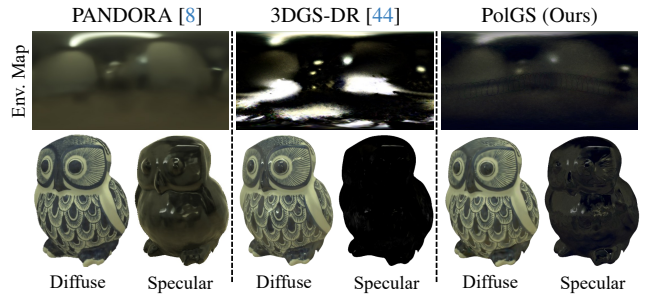


Figure 6. Separation of diffuse and specular components with PANDORA [8], 3DGS-DR [44] and PolGS (Ours).

5.3. Comparison of radiance decomposition

Figure 6 presents a comparison of diffuse and specular component decomposition generated by PANDORA [8], 3DGS-DR [44], and PolGS (Ours). Here, PANDORA [8] utilizes the IDE [34] structure to produce environment map results, whereas PolGS (Ours) adopts 3DGS-DR [44]’s methods by

Table 2. Comparisons on SMVP3D [12] and RMVP3D [12] evaluated by mean angular error (MAE) ($\downarrow$) with degree and Chamfer distance (CD) ($\downarrow$) in millimeter (mm), respectively. Best and second results in SDF-based and 3DGS-based methods (except PolGS w/o $\mathcal{L}_{pol}$) are highlighted as **1st** and **2nd**. The time consumed is shown on the far right side of the table.

Method	SMVP3D [12]										RMVP3D [12]						Mean		time
	HEDGEHOG		SQUIRREL		SNAIL		DAVID		DRAGON		SHISA		FROG		DOG				
	MAE	CD	MAE	CD	MAE	CD	MAE	CD	MAE	CD	MAE	CD	MAE	CD	MAE	CD	MAE	CD	
PANDORA [8]	9.41	9.50	10.85	5.88	8.08	10.97	14.75	4.88	16.33	4.78	12.93	11.29	15.86	7.88	20.11	10.19	13.54	8.17	10 h
NeRSP [12]	8.94	6.57	8.23	3.02	5.56	3.72	15.38	4.18	15.30	3.01	10.79	7.39	15.62	6.68	16.57	8.57	12.05	5.39	10 h
MVAS [3]	4.30	4.22	6.10	3.73	3.30	7.87	8.47	3.21	8.11	1.89	8.56	9.28	17.63	7.00	16.65	8.76	9.01	5.74	11 h
NeRO [23]	3.4	3.69	3.55	1.86	2.67	3.71	7.64	2.88	8.12	1.69	8.41	4.88	15.29	5.39	17.72	8.74	8.27	4.10	8 h
3DGS-DR [44]	12.28	12.66	17.18	11.20	11.42	20.7	20.56	7.91	26.20	9.56	19.53	15.87	17.08	27.67	24.85	11.00	19.76	13.32	0.2 h
Re-3DGS [10]	14.40	19.84	15.41	18.97	9.08	19.04	15.47	13.47	17.65	10.96	16.69	14.66	17.85	13.35	21.53	12.24	15.89	14.12	0.4 h
Gaussian Surfels [7]	16.50	8.82	21.65	9.53	19.05	14.04	21.56	7.39	21.41	6.98	12.79	9.09	16.19	7.01	19.30	9.56	18.56	10.30	0.2 h
PolGS (Ours)	10.83	7.62	11.42	6.28	9.64	10.85	13.99	5.30	24.23	7.61	10.88	7.76	15.03	7.48	18.80	7.71	14.35	7.57	0.1 h
PolGS w/o $\mathcal{L}_{pol}$	11.39	7.97	11.83	6.39	9.82	10.75	15.44	5.90	25.74	8.22	10.95	7.92	15.16	7.56	18.86	7.74	14.89	7.80	0.1 h

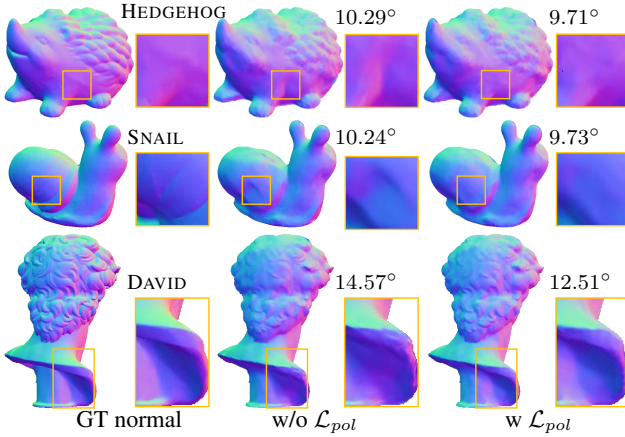


Figure 7. Ablation study on surface normal estimation with and without polarimetric information $\mathcal{L}_{pol}$. The MAE values are displayed on the top of each image.

employing a Cubemap encoder for the same purpose. Compared with 3DGS-DR [44], PolGS (Ours) leverages additional polarimetric information to effectively constrain and disambiguate diffuse and specular components. Notably, our results closely align with PANDORA [8]’s, demonstrating improved radiance decomposition ability.

5.4. Ablation study

In this section, we conduct an ablation study to evaluate the contribution of polarimetric information to surface reconstruction. As shown in Table 2 and Fig. 7, integrating $\mathcal{L}_{pol}$ consistently improves reconstruction accuracy across both synthetic and real-world datasets. The incorporation of $\mathcal{L}_{pol}$ particularly enhances geometric fidelity in challenging regions such as the HEDGEHOG’s bottom and the SNAIL’s back, where it effectively suppresses concave artifacts.

However, for texture-rich surfaces captured under dense views (>30 views), the existing photometric cues (RGB) provide sufficient constraints, leaving limited room for additional improvement from polarimetric information. This contrasts with texture-less surfaces like the DAVID statue, where $\mathcal{L}_{pol}$ contributes to reconstruction quality by providing additional geometric constraints.

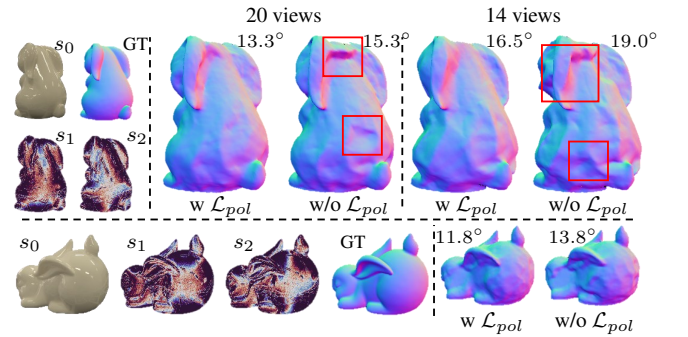


Figure 8. Stokes vectors can help surface normal reconstruction on texture-less surfaces.

Boosting on texture-less surfaces recovery with $\mathcal{L}_{pol}$ To further validate the benefits of $\mathcal{L}_{pol}$ on texture-less surfaces, we conduct additional experiments using a real-world dataset from PISR [5], as shown in Fig. 8. While the RGB appearance lacks discernible texture, the polarization channels s_1 and s_2 exhibit meaningful variations that provide additional geometric cues for resolving shape ambiguities. The results show that both concave and convex surface distortions are more accurately reconstructed when $\mathcal{L}_{pol}$ is employed. Moreover, as the number of input views decreases, the advantages of incorporating polarization become even more evident. This demonstrates the unique value of polarimetric information in challenging scenarios where conventional RGB inputs offer limited constraints.

6. Conclusion

We propose PolGS, a novel Polarimetric Gaussian Splatting method for fast and accurate 3D reconstruction of reflective surfaces. PolGS leverages polarimetric information for separating the diffuse and specular components based on Stokes field, which can help to constrain the surface normal in the gaussian splatting representation and finally improve the reflective surface reconstruction results. Extensive experiments demonstrate that PolGS outperforms state-of-the-art 3DGS methods in terms of accuracy and efficiency.

Acknowledgments

This work was supported by Hebei Natural Science Foundation Project No. 242Q0101Z, Beijing-Tianjin-Hebei Basic Research Funding Program No. F2024502017, National Natural Science Foundation of China (Grant No. 62472044, U24B20155, 62225601, U23B2052, 62136001), Beijing Natural Science Foundation Project No. L242025, BUPT Excellent Ph.D. Students Foundation. We thank open-bayes.com for providing computing resource.

References

- [1] Yunhao Ba, Alex Gilbert, Franklin Wang, Jinfa Yang, Rui Chen, Yiqin Wang, Lei Yan, Boxin Shi, and Achuta Kadambi. Deep shape from polarization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. 3
- [2] Seung-Hwan Baek, Daniel S Jeon, Xin Tong, and Min H Kim. Simultaneous acquisition of polarimetric svbrdf and normals. *ACM Transactions on Graphics*, 2018. 2, 3, 4, 5, 6
- [3] 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. 1, 2, 3, 6, 7, 8
- [4] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. TensorRF: Tensorial radiance fields. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 2
- [5] 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. 3, 6
- [6] Hanlin Chen, Chen Li, and Gim Hee Lee. NeuSG: Neural implicit surface reconstruction with 3d gaussian splatting guidance. *arXiv preprint arXiv:2312.00846*, 2023. 2
- [7] 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, 6, 7, 8
- [8] 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, 7, 8
- [9] Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. Deep polarization imaging for 3D shape and svbrdf acquisition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. 3
- [10] 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, 6, 7, 8
- [11] Antoine Guédon and Vincent Lepetit. SuGaR: Surface-aligned gaussian splatting for efficient 3D mesh reconstruction and high-quality mesh rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. 2
- [12] 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, 6, 7, 8
- [13] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2D Gaussian Splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH*, 2024. 2
- [14] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 2023. 1, 2, 3, 6
- [15] Chenyang Lei, Chenyang Qi, Jiaxin Xie, Na Fan, Vladlen Koltun, and Qifeng Chen. Shape from polarization for complex scenes in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 3
- [16] Chenhao Li, Trung Thanh Ngo, and Hajime Nagahara. Deep Polarization Cues for Single-shot Shape and Subsurface Scattering Estimation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2024. 3
- [17] Chenhao Li, Taishi Ono, Takeshi Uemori, Hajime Mihara, Alexander Gatto, Hajime Nagahara, and Yusuke Moriuchi. NeISF: Neural Incident Stokes Field for Geometry and Material Estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. 1, 3
- [18] Chenhao Li, Taishi Ono, Takeshi Uemori, Sho Nitta, Hajime Mihara, Alexander Gatto, Hajime Nagahara, and Yusuke Moriuchi. NeISF++: Neural incident stokes field for polarized inverse rendering of conductors and dielectrics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. 3
- [19] Jia Li, Lu Wang, Lei Zhang, and Beibei Wang. TensorSDF: Roughness-aware tensorial representation for robust geometry and material reconstruction. *ACM Transactions on Graphics*, 2024. 2
- [20] Rui Li, Simeng Qiu, Guangming Zang, and Wolfgang Heidrich. Reflection separation via multi-bounce polarization state tracing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. 3
- [21] Zhuoxiao Li, Haiyang Jiang, Mingdeng Cao, and Yinqiang Zheng. Polarized color image denoising. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 3
- [22] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 2
- [23] 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. 1, 2, 6, 7, 8

- [24] Xiaoyang Lyu, Yang-Tian Sun, Yi-Hua Huang, Xiuzhe Wu, Ziyi Yang, Yilun Chen, Jiangmiao Pang, and Xiaojuan Qi. 3DGSR: Implicit surface reconstruction with 3D gaussian splatting. *ACM Transactions on Graphics*, 2024. 2
- [25] Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, and Boxin Shi. Physics-guided reflection separation from a pair of unpolarized and polarized images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 3
- [26] Youwei Lyu, Lingran Zhao, Si Li, and Boxin Shi. Shape from polarization with distant lighting estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 3
- [27] Youwei Lyu, Heng Guo, Kailong Zhang, Si Li, and Boxin Shi. SiPUEL: Shape from polarization under unknown environment light. *Advances in Neural Information Processing Systems (NeurIPS)*, 2024. 3
- [28] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. 1, 2
- [29] Miyazaki, Tan, Hara, and Ikeuchi. Polarization-based inverse rendering from a single view. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2003. 3
- [30] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multi-resolution hash encoding. *ACM Transactions on Graphics*, 2022. 2
- [31] Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric rendering: Learning implicit 3D representations without 3D supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 2
- [32] Michael Oechsle, Songyou Peng, and Andreas Geiger. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021. 2
- [33] William AP Smith, Ravi Ramamoorthi, and Silvia Tozza. Height-from-polarisation with unknown lighting or albedo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018. 3
- [34] Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T Barron, and Pratul P Srinivasan. Ref-NeRF: Structured view-dependent appearance for neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 2, 7
- [35] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, 2007. 4
- [36] Haoyuan Wang, Wenbo Hu, Lei Zhu, and Rynson WH Lau. Inverse rendering of glossy objects via the neural plenoptic function and radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. 2
- [37] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction. *arXiv preprint arXiv:2106.10689*, 2021. 2, 4, 5
- [38] Xin Wang, Yong Zhang, and Yanchu Chen. Polarized reflection removal with dual-stream attention guidance. *Pattern Recognition*, 2025. 3
- [39] Yiqun Wang, Ivan Skorokhodov, and Peter Wonka. Hf-NeuS: Improved surface reconstruction using high-frequency details. *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. 2
- [40] Yiming Wang, Qin Han, Marc Habermann, Kostas Daniilidis, Christian Theobalt, and Lingjie Liu. NeuS2: Fast learning of neural implicit surfaces for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. 2
- [41] LI Yang, WU Ruizheng, LI Jiyong, and CHEN Ying-cong. GNeRP: Gaussian-guided neural reconstruction of reflective objects with noisy polarization priors. *arXiv preprint arXiv:2403.11899*, 2024. 3
- [42] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural Information Processing Systems (NeurIPS)*, 2020. 2
- [43] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021. 2
- [44] Keyang Ye, Qiming Hou, and Kun Zhou. 3D Gaussian Splatting with Deferred Reflection. In *ACM SIGGRAPH*, 2024. 1, 2, 3, 5, 6, 7, 8
- [45] Mulin Yu, Tao Lu, Linning Xu, Lihan Jiang, Yuanbo Xiangli, and Bo Dai. GSDF: 3DGS meets SDF for improved rendering and reconstruction. *arXiv preprint arXiv:2403.16964*, 2024. 2
- [46] Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient adaptive surface reconstruction in unbounded scenes. *ACM Transactions on Graphics*, 2024. 3
- [47] Jinyu Zhao, Yusuke Monno, and Masatoshi Okutomi. Polarimetric multi-view inverse rendering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 3
- [48] Chu Zhou, Yufei Han, Mingui Teng, Jin Han, Si Li, Chao Xu, and Boxin Shi. Polarization guided HDR reconstruction via pixel-wise depolarization. *IEEE Transactions on Image Processing*, 2023. 3
- [49] Matthias Zwicker, Hanspeter Pfister, Jeroen Van Baar, and Markus Gross. Surface splatting. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001. 3
- [50] Matthias Zwicker, Hanspeter Pfister, Jeroen Van Baar, and Markus Gross. Ewa splatting. *IEEE Transactions on Visualization and Computer Graphics*, 2002. 3

PolGS: Polarimetric Gaussian Splatting for Fast Reflective Surface Reconstruction

Supplementary Material

Yufei Han¹, Bowen Tie¹, Heng Guo^{1,2*}, Youwei Lyu¹, Si Li^{1*}, Boxin Shi^{3,4}, Yunpeng Jia¹, Zhanyu Ma¹

¹Beijing University of Posts and Telecommunications ²Xiong'an Aerospace Information Research Institute

³State Key Laboratory of Multimedia Information Processing, School of Computer Science, Peking University

⁴National 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	1
B Additional results on real-world datasets	1
B.1. Comparison on PANDORA [5]	2
B.2. Comparison on RMVP3D [7]	2
C Additional results on synthetic dataset	2
C.1. Comparison on SMVP3D [7]	2
D Analysis of SH coefficient adjustment	2
E Implementation details	4
E.1. Experiment settings	4
E.2. Training and Evaluation	4
F. Failure case	4
G Limitations	5

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:

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

The diffuse and specular Stokes components under a single

light source can be formulated respectively as:

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

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

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 $C = \int_{\Omega} \rho_d L(\omega) \omega^\top \mathbf{n} T_i^+ T_i^- d\omega$ and specular color $L_r = \rho_s \int_{\Omega} L(\omega) \frac{DG}{4\mathbf{n}^\top \mathbf{v}} d\omega$, and the final $\mathbf{s}_o(\mathbf{v})$ can be represented as:

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

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.

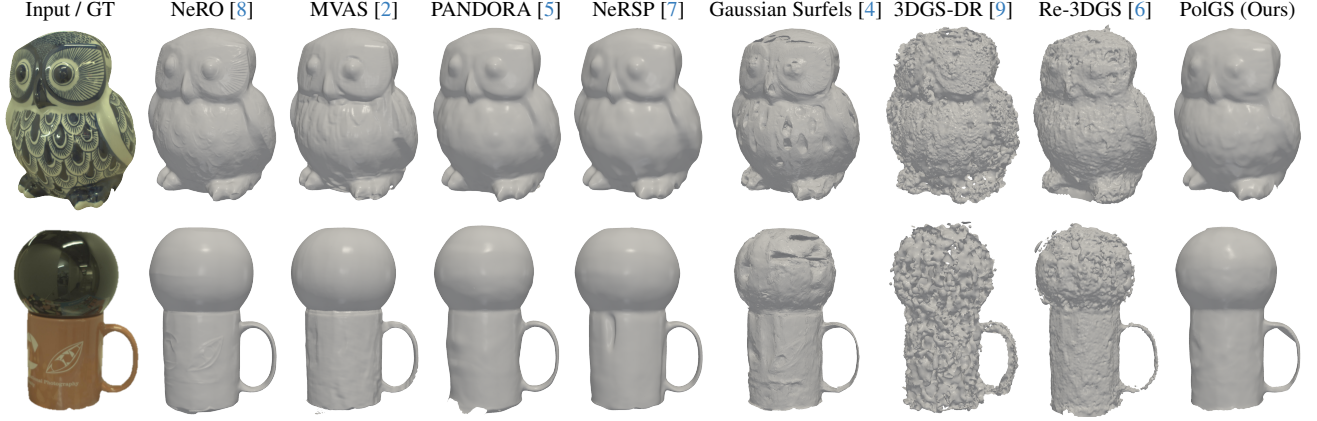


Figure 1. Qualitative shape recoveries on PANDORA [5].

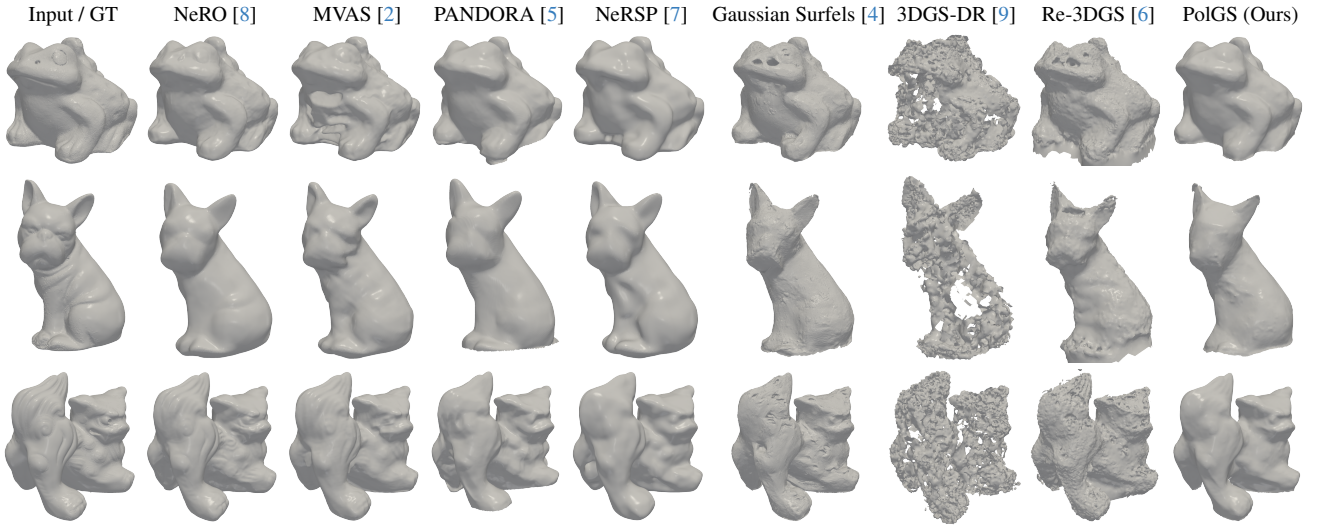


Figure 2. Qualitative shape recoveries on RMVP3D [7].

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-

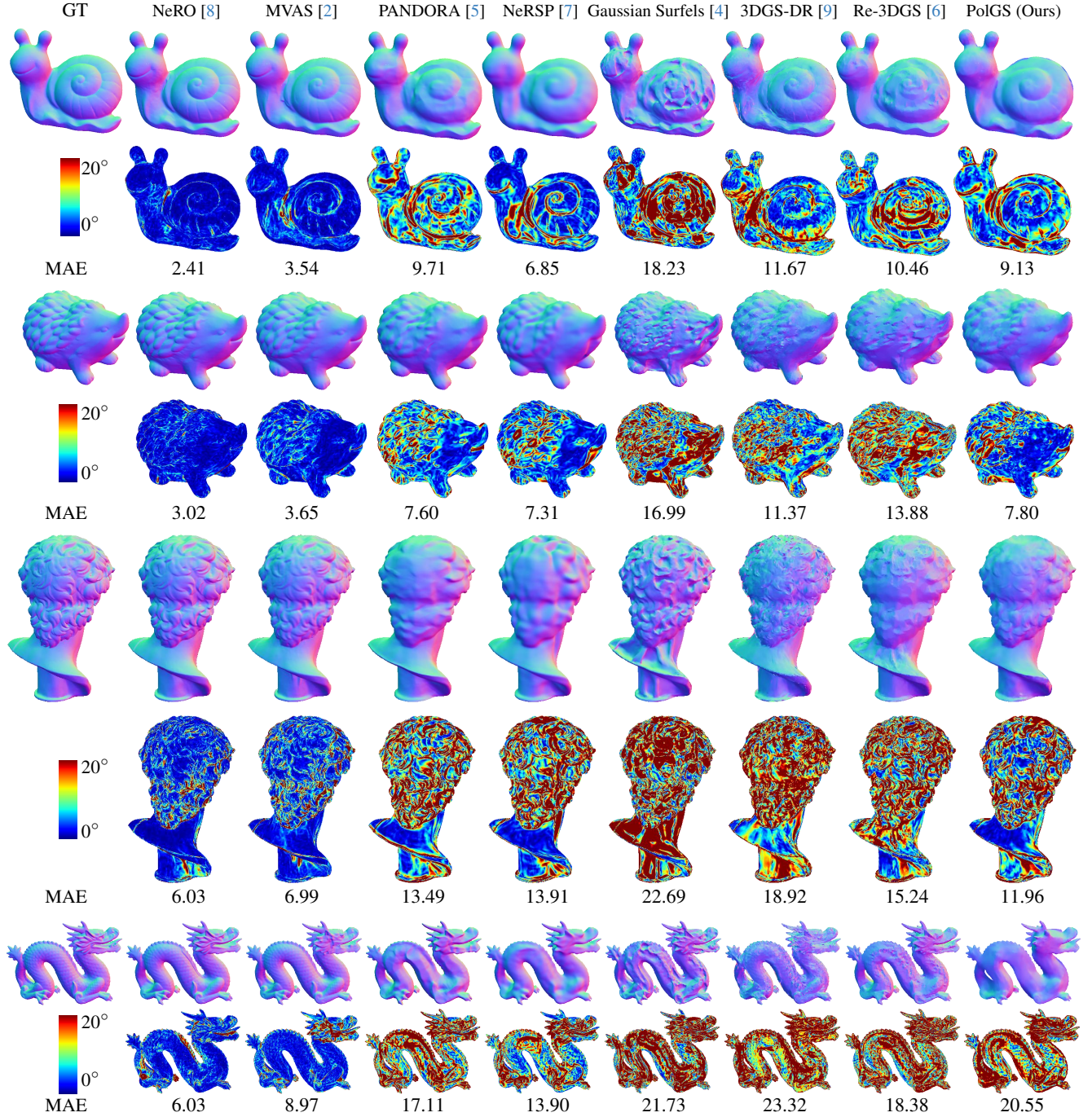


Figure 3. Qualitative comparisons on surface normal estimation of SQUIRREL in SMVP3D [7], where our 3DGS-based method can outperform existing methods based on the same representation and achieves comparable results with SDF-based methods such as NeRSP [7] and PANDORA [5] while with higher efficiency.

efficient 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 phe-

nomenon further demonstrates the reasonableness of our adjustment of the SH coefficients.

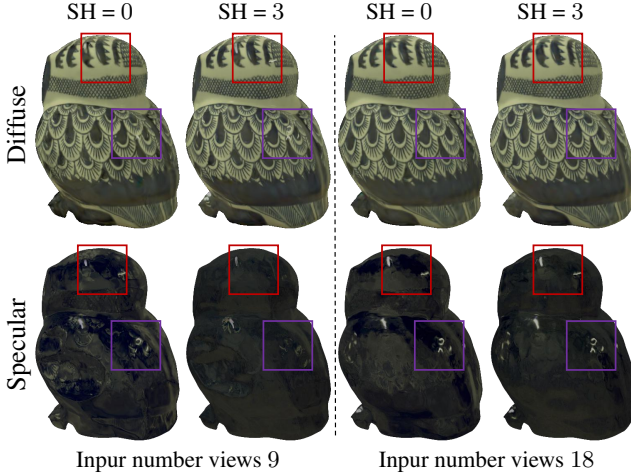


Figure 4. Radiance decomposition results obtained using different spherical harmonic (SH) coefficient orders, specifically 0, 1, and 3.

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,

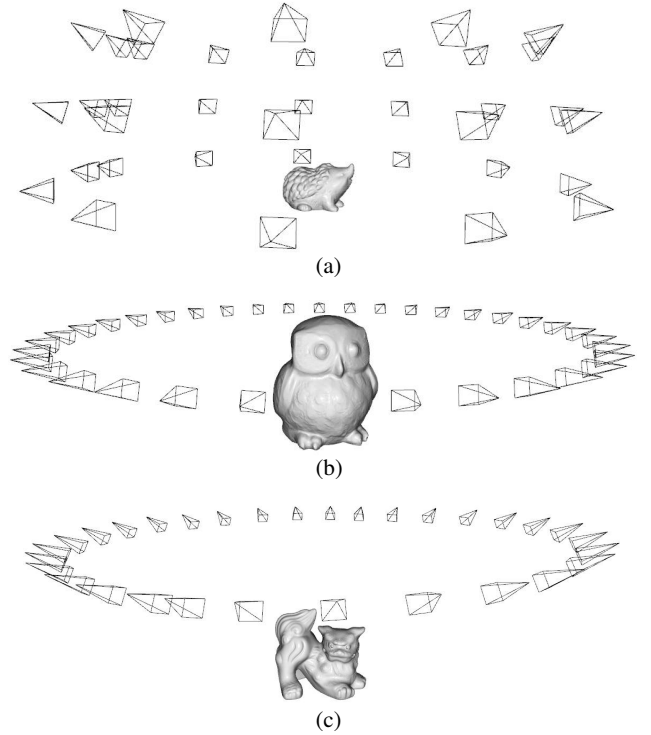


Figure 5. Visualization of view distributions of different datasets. (a) Input views display on SMVP3D [7]. (b) Input views display on PANDORA [5]. (c) Input views display on RMVP3D [7].

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.

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