Rapid and Physically-Based Gaussian Splatting of Unknown Space Objects in Low Earth Orbit

Tae Ha Park[1]

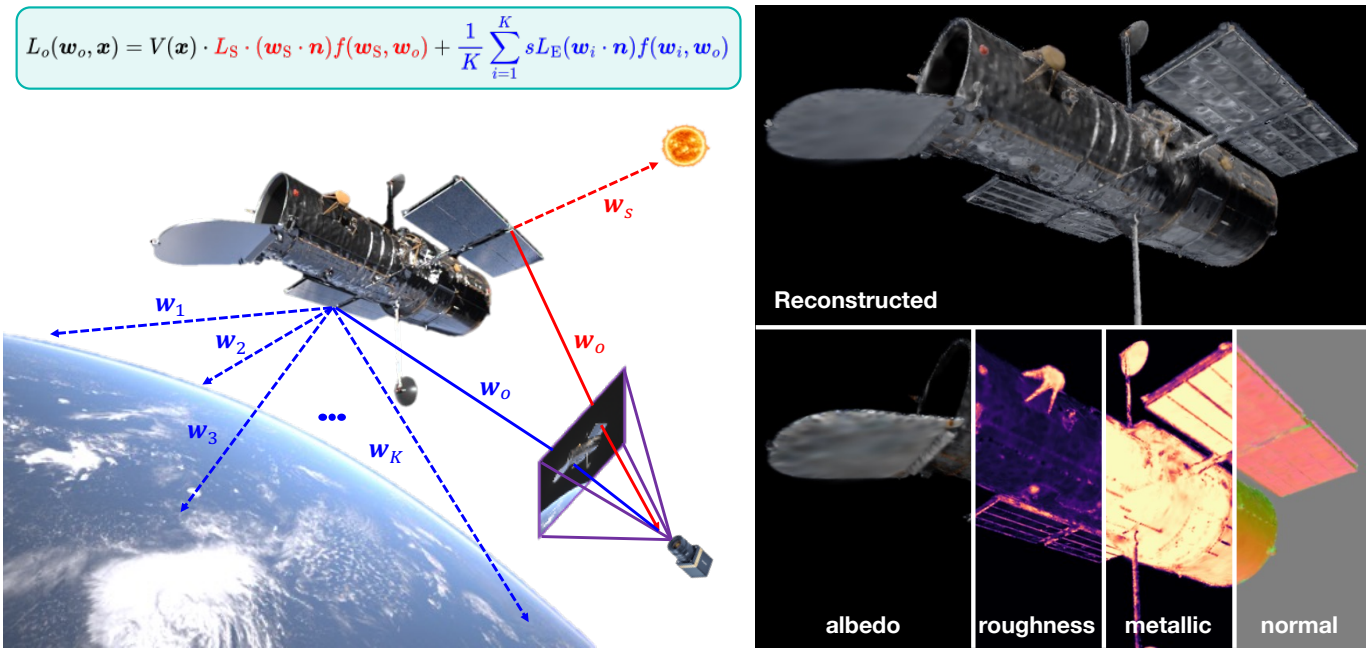

Fig. 1: We present a physically-based 2D Gaussian Splatting pipeline that leverages known Sun and Earth illumination in low Earth orbit. By modeling both directional solar lighting and diffuse Earth albedo, our model enables accurate material estimation and produces high-fidelity novel view synthesis and 3D reconstructions of spacecraft with complex geometry and appearance.

This paper presents a physically-based 2D Gaussian Splatting (2DGS) pipeline for novel view synthesis and 3D reconstruction of unknown spaceborne targets in Low Earth Orbit (LEO). In this setting, accurate reconstruction requires not only geometric fidelity but also photometric consistency, as appearance cues play a key role in estimating camera poses from image observations. However, existing approaches often rely on simplified lighting assumptions that limit their ability to model real space environments. To address this, the proposed method explicitly incorporates illumination priors by modeling the Sun as a directional light source and the Earth as a secondary, predominantly diffuse illumination term, capturing the dominant lighting effects in LEO. A deferred shading formulation, combined with an analytical Bidirectional Reflectance Distribution Function (BRDF) leveraging the inherent surface normals in 2DGS, enables efficient material estimation and rapid reconstruction under varying illumination conditions. Experiments on a high-fidelity synthetic dataset of the Hubble Space Telescope, featuring realistic and challenging materials, illumination, and sensor effects, demonstrate strong material consistency and rendering quality, particularly for challenging appearance effects such as Earth-induced reflections on specular surfaces.

T. H. Park is supported by the Technological Innovation R&D Program (RS-2024-00513881) funded by the Ministry of SMEs and Startups, Republic of Korea.

[1]Nara Space Technology Inc., Seoul, Republic of Korea. Email: thpark@naraspace.com

I. Introduction

Autonomous Guidance, Navigation and Control (GN&C) for non-cooperative spaceborne targets is a critical capability for Rendezvous and Proximity Operations (RPO), including on-orbit servicing and active debris removal [1]–[3]. Monocular vision-based approaches are particularly attractive due to their low Size, Weight, Power and Cost (SWaP-C) requirements, but require real-time estimation of the target's relative pose from image sequences acquired during RPO. Most existing works assume known targets with pre-existing 3D models, enabling the use of learning-based methods for single-frame pose estimation or temporally filtered tracking, often trained on synthetic datasets and validated with hardware-in-the-loop testbeds to ensure robustness to real-world conditions [4]–[7]. However, future missions such as debris removal demand relaxing this assumption, requiring the servicer to simultaneously estimate the target's motion, geometry, and inertial properties online, leading to a problem formulation analogous to Simultaneous Localization and Mapping (SLAM) in robotics [8].

Recent advances in vision-based 3D reconstruction and SLAM leverage radiance field models, notably Neural Radiance Fields (NeRF) [9] and Gaussian Splatting (GS) [10]. These

approaches represent a scene as a function that models how light is emitted and attenuated through space, by learning a view-dependent radiance field function $f(\boldsymbol{x}, \boldsymbol{d}) \rightarrow (\boldsymbol{c}, \sigma)$ that maps the 3D position $\boldsymbol{x}$ and viewing direction $\boldsymbol{d}$ to color $\boldsymbol{c}$ and volume density σ. Given this representation, novel view images are rendered by integrating radiance along camera rays using volumetric rendering, where colors and densities are composited via α-blending. Crucially, this rendering process is differentiable, enabling supervision through photometric reconstruction loss. NeRF and its variants parameterize the radiance field with a neural network, and optimize its parameters via differentiable volume rendering. However, this requires dense sampling along each ray, making it computationally expensive. In contrast, GS represents the scene as a set of anisotropic 3D Gaussians with learnable colors and opacities, and renders images via differentiable rasterization by splatting these primitives onto the image plane with α-compositing. This formulation allows direct optimization of Gaussian parameters using photometric loss, enabling efficient real-time rendering.

Notably, the differentiable rendering process can be leveraged not only to learn the scene representation but also to optimize camera extrinsics, enabling joint estimation of both camera poses and the global scene. This is achieved by minimizing photometric reconstruction loss, which implicitly constrains the camera poses such that accurate rendering requires correct pose alignment. Such a formulation has been successfully demonstrated in various SLAM systems based on both NeRF [11], [12] and Gaussian Splatting [13]–[15]. However, these approaches generally assume static scenes and often struggle in the presence of dynamic elements, where inconsistencies in photometric supervision can degrade both reconstruction quality and pose estimation. This limitation is particularly critical for an object-centric 3D reconstruction in highly dynamic in-orbit environments, e.g., RPO scenarios, where both the observer and target may undergo complex relative motion with respect to the Sun and Earth.

While dynamic illumination poses a significant challenge in in-orbit RPO scenarios, the dominant light source—the Sun—is both predictable and accurately observable via onboard Attitude and Orbit Control System (AOCS) and sun sensors, enabling its incorporation as a prior in the training of radiance field models. Park & D'Amico [16] leverage this illumination prior to compute real-time Gaussian visibility via shadow splatting [17], which is used to supervise a Multi-Layer Perceptron (MLP) to model dynamic appearance variation. However, using an MLP to capture dynamic view- and illumination-dependent appearances of individual Gaussians can become time-consuming due to its unconstrained nature. This could result in slow training and convergence which is detrimental to online SLAM during RPO. Moreover, they only consider the solar illumination and ignores Earth albedo which has non-negligible effect on spaceborne imagery acquired in Low Earth Orbits (LEO).

Building upon the work of Park & D'Amico [16], this paper proposes a physically-based 3D reconstruction framework that replaces learned appearance models with an analytical formulation of light transport. Specifically, a 2D Gaussian Splatting (2DGS) [18] framework is adopted with a deferred shading mechanism and an analytical Bidirectional Reflectance Distribution Function (BRDF), leveraging inherent surface normals for efficient and physically-consistent material estimation. To address the limitations of purely directional illumination in Low Earth Orbit (LEO), an explicit model of Earth albedo is introduced, along with an efficient sampling strategy that approximates incident albedo lights from sunlit regions of the Earth. This formulation enables a principled decomposition of space illumination into solar lighting, shadowing, and Earth albedo components. Experiments on a challenging dataset in LEO demonstrate the critical role of modeling Earth albedo for achieving physically consistent photometric reconstruction.

II. Preliminaries

A. Gaussian Splatting

The standard 3D Gaussian Splatting (3DGS) [10] models the scene as a set of anisotropic 3D Gaussian ellipsoids. Each Gaussian is characterized by the following learnable parameters: mean position $\boldsymbol{\mu}_i \in \mathbb{R}^3$, covariance $\boldsymbol{\Sigma}_i \in \mathbb{S}^3_+$, opacity $o_i \in [0, 1]$ and view-dependent color $\boldsymbol{c}_i$ represented as Spherical Harmonics (SH) coefficients. The density at point $\boldsymbol{x}$ is defined as

$$\mathcal{G}_i^{\text{3D}}(\boldsymbol{x}) = \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_i) \right) \tag{1}$$

During rendering, 3D Gaussians are first projected to 2D image space. The projected 2D Gaussians are affine approximated as $\boldsymbol{\Sigma}' = \boldsymbol{J}\boldsymbol{W}\boldsymbol{\Sigma}_i\boldsymbol{W}^\top\boldsymbol{J}^\top$ [19] where $\boldsymbol{W}$ is the camera view matrix and $\boldsymbol{J}$ is the Jacobian of the projective transformation. Then, each pixel color is rendered via α-blending the color of the depth-wise sorted Gaussians:

$$\hat{\boldsymbol{C}}(\boldsymbol{x}) = \sum_{i \in \mathcal{N}} \boldsymbol{c}_i \alpha_i \prod_{j=1}^{i-1}(1 - \alpha_j), \quad \alpha_i = o_i \mathcal{G}_i(\boldsymbol{x}) \tag{2}$$

where $\mathcal{G}_i(\boldsymbol{x})$ denotes density of projected Gaussians.

Despite the success of 3DGS, its volumetric representation lacks view consistency and accurate surface reconstructions. In response, 2D Gaussian Splatting (2DGS) [18] leverages a set of 2D Gaussian surfels which are equivalent to 3D Gaussians with a collapsed dimension. The covariance of a surfel centered at $\boldsymbol{\mu}$ is defined by the scaling factors (s_u, s_v) along local tangential vectors $(\boldsymbol{t}_u, \boldsymbol{t}_v)$. Then, the splatting process of 2DGS involves finding the exact ray-splat intersection $\boldsymbol{u} = (u, v)$ in the local tangent space, and the Gaussian density is simply given as

$$\mathcal{G}_i^{\text{2D}}(\boldsymbol{u}) = \exp\left(-\frac{u^2 + v^2}{2} \right) \tag{3}$$

which can be directly used in α-blending (Eq. 2).

In literature, normals are introduced to 3D Gaussian representations either as additional learnable parameters or approximated by enforcing one principal axis to be the shortest and treating it as the normal [20], [21]. However, such approaches are indirect and still depend on approximate covariance projections of 3D Gaussians. In contrast, 2D Gaussian surfels

inherently define surface normals via their local tangent vectors as $\boldsymbol{t}_u \times \boldsymbol{t}_v$, making them a more natural and geometrically consistent choice for performing physically-based rendering. Therefore, 2DGS is adopted as the core scene representation in this work.

B. Physically-Based Rendering

Physically-Based Rendering (PBR) has long history in gaming industries [22], [23] and has recently been explored in radiance field literature for inverse rendering, where the goal is to recover scene properties such as geometry, materials and lighting from images, enabling improved generalization to novel view synthesis under varying textures and illumination conditions [17], [21], [24]–[26]. At the core of these approaches is the rendering equation [27] which governs the amount of light reflected at a surface point $\boldsymbol{x}$:

$$L_o(\boldsymbol{w}_o, \boldsymbol{x}) = \int_{\Omega_+} L_i(\boldsymbol{w}_i, \boldsymbol{x})(\boldsymbol{w}_i \cdot \boldsymbol{n}) f(\boldsymbol{w}_i, \boldsymbol{w}_o, \boldsymbol{x}) d\boldsymbol{w}_i \quad (4)$$

where $\boldsymbol{w}_i$, $\boldsymbol{w}_o$ respectively denote the incoming and outgoing radiance directions, $\boldsymbol{n}$ is the surface normal direction at $\boldsymbol{x}$, $f(\boldsymbol{w}_i, \boldsymbol{w}_o, \boldsymbol{x})$ is the Bidirectional Reflectance Distribution Function (BRDF) function describing the reflection off the surface, $L_i(\boldsymbol{w}_i, \boldsymbol{x})$ is the incoming radiance at $\boldsymbol{x}$, and Ω_+ denotes the upper hemisphere where $\boldsymbol{w}_i \cdot \boldsymbol{n} > 0$.

In literature, many works adopt the Disney principled BRDF model [22] which decomposes f into a diffuse term and a Cook-Torrance microfacet-based specular term [28] given as

$$f(\boldsymbol{w}_i, \boldsymbol{w}_o, \boldsymbol{x}) = \frac{1-m}{\pi}\boldsymbol{a} + \frac{D(\boldsymbol{h})F(\boldsymbol{h}, \boldsymbol{w}_o)G(\boldsymbol{n}, \boldsymbol{w}_i, \boldsymbol{w}_o)}{4(\boldsymbol{n} \cdot \boldsymbol{w}_i)(\boldsymbol{n} \cdot \boldsymbol{w}_o)} \quad (5)$$

where $\boldsymbol{a} \in [0,1]^3$ is diffuse albedo color, $\rho \in [0,1]$ is roughness, and $m \in [0,1]$ the metallic term. Details on each term are provided in Appendix A.

III. Methodology

This section provides a detailed description of the proposed methodology and pipeline visualized in Fig. 1. Similar to Park & D'Amico [16], several favorable assumptions are made to isolate the problem of photometric reconstruction. Specifically, we assume known camera-to-target poses and the availability of binary masks to exclude the Earth background.

A. Leveraging Illumination Priors

In this work, we fully leverage the illumination priors available during space missions. In addition to the Sun which acts as the dominant source of illumination [16], we explicitly model the Earth's albedo as a secondary, predominantly diffuse illumination term within the splatting pipeline. To this end, we assume that the servicer spacecraft has access to accurate Sun vector measurements from onboard Sun sensors and the Earth's position from its AOCS module. Given the relative pose of the target (i.e., position and orientation) with respect to the chaser, we express both the Sun and Earth directions in the target body frame, which serves as the world frame.

While sunlight can be considered a directional light from an effectively infinite source, the Earth's atmosphere and surface scatter and reflect it back into space as a predominantly diffuse illumination. This allows the rendering equation in Eq. 4 to be decomposed as

$$\begin{aligned} L_o(\boldsymbol{w}_o, \boldsymbol{x}) = V(\boldsymbol{x}) \underbrace{L_{\mathrm{S}} \cdot (\boldsymbol{w}_{\mathrm{S}} \cdot \boldsymbol{n}) f(\boldsymbol{w}_{\mathrm{S}}, \boldsymbol{w}_o), \boldsymbol{x}}_{L_{\text{sun}}} \\ + \underbrace{\int_{\Omega_{\mathrm{E}}} L_i(\boldsymbol{w}_i)(\boldsymbol{w}_i \cdot \boldsymbol{n}) f(\boldsymbol{w}_i, \boldsymbol{w}_o, \boldsymbol{w}_i) d\boldsymbol{w}_i}_{L_{\text{earth}}} \end{aligned} \quad (6)$$

where $\boldsymbol{w}_{\mathrm{S}}$ denotes the Sun vector and Ω_{E} now represents the portion of the hemisphere subtended by the Earth that is illuminated by the Sun. We model the solar radiance as a learnable global scalar $L_{\mathrm{S}} \in \mathbb{R}_+$. Assuming white illumination, L_{S} is applied uniformly across all color channels.

In Eq. 6, the $V(\boldsymbol{x})$ term denotes the visibility of a Gaussian, accounting for self-occlusion due to the target's geometry. We consider occlusion only with respect to the Sun, as the Earth albedo originates from a spatially extended source and is thus less sensitive to fine-scale self-shadowing. Following GS^3 [17] and Park & D'Amico [16], we employ shadow splatting to compute $V(\boldsymbol{x})$, where Gaussian primitives are rasterized for visibility from the Sun's perspective using an orthographic projection. We refer the reader to GS^3 [17] for further details.

Inspired by NeILF [24] which employs Fibonacci sampling to obtain discrete samples of incident illumination ($L_i(\boldsymbol{w}_i)$) parameterized by a neural incident light field, we instead propose to sample K incident rays only from the sunlit portion of the Earth. The integral can then be approximated as

$$L_{\text{earth}} = \frac{1}{K} \sum_{i=1}^{K} s L_{\mathrm{E}} \cdot (\boldsymbol{w}_i \cdot \boldsymbol{n}) f(\boldsymbol{w}_i, \boldsymbol{w}_o) \quad (7)$$

where $L_{\mathrm{E}} \in \mathbb{R}^3_+$ denotes a learnable global Earth albedo radiance term, and the scalar attenuation factor s modulates this contribution based on the Sun-Earth geometry.

In order to accelerate training, the incident light rays are precomputed and saved in a Look-Up Table (LUT). Specifically, we first align the Earth with the z-axis of the world frame and vary the solar elevation angle $\theta \in [0°, 180°]$ within the xz-plane. For each discrete θ_i, we sample up to K incident rays that are both visible from the satellite at a given altitude and illuminated by the Sun at angle θ_i (see Fig. 2). During training, the saved samples are retrieved from the LUT and rotated according to the actual Sun and Earth directions.

B. Deferred Shading

The proposed method adopts a deferred shading mechanism [25] in order to enable more accurate estimation of the target's material properties, which is critical for high-fidelity novel view synthesis under varying illumination. In conventional GS, view-dependent colors $\boldsymbol{c}_i$ are computed per Gaussian prior to

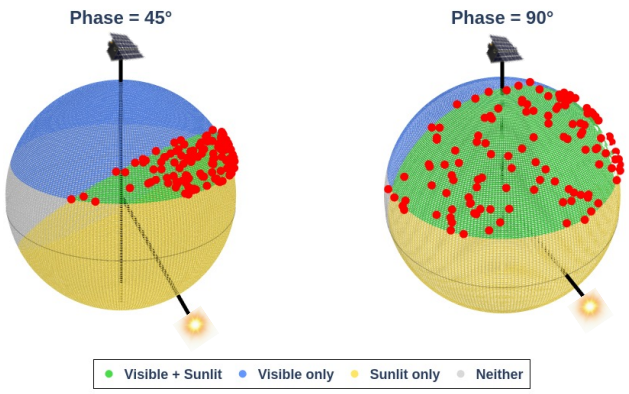

Fig. 2: Visualization of $\boldsymbol{w}_i$ samples for different Sun-Earth geometries. During training, the samples shown in red are retrieved then rotated to match the actual Earth and Sun positions with respect to the world frame.

rasterization and subsequently composited via α-blending. This tightly couples geometry, appearance and lighting, limiting the model's ability to generalize across illumination changes.

In contrast, our approach decouples material estimation from shading. This is done by first rasterizing intrinsic material properties, $(\boldsymbol{a}, \rho, m)$, along with surface normals $\boldsymbol{n}$ using Eq. 2, resulting in pixel-wise accumulated material buffers (see Fig. 1). Shading is then performed in a separate stage by evaluating Eqs. 6 & 7 using both Sun and Earth illumination, enabling consistent rendering under novel lighting conditions and improved physical plausibility.

C. Optimization

Finally, given the predicted radiance $\hat{\boldsymbol{L}}$ from Eq. 6 and the observed ground-truth $\boldsymbol{L}$, the Gaussians are trained by minimizing the loss function. Since the proposed formulation operates directly in the radiance domain, rather than low dynamic range (LDR) color space commonly used in existing datasets, training is performed using high dynamic range (HDR) ground-truth images. This is consistent with onboard imaging pipelines, which typically produce high bit-depth raw measurements (e.g., 10–12 bits). This also avoids the need to learn nonlinear tone mapping with unknown camera response functions or exposure time [29], [30].

However, directly applying the pixel-wise ℓ_1 loss used in GS leads to the optimization being dominated by high-radiance regions, degrading reconstruction of low-radiance regions. Therefore, we adopt the reconstruction loss of HDRSplat [31] which scales the ℓ_1 using the predicted radiance:

$$\mathcal{L}_{\text{HDR}}(\hat{\boldsymbol{L}}, \boldsymbol{L}) = \frac{\|\hat{\boldsymbol{L}} - \boldsymbol{L}\|_1}{\text{sg}(\hat{\boldsymbol{L}}) + \varepsilon} \tag{8}$$

where $\text{sg}(\cdot)$ denotes the stop-gradient operation to ensure that the normalization term does not bias the gradient flow.

Then, the total loss is given as

$$\mathcal{L}_{\text{total}} = (1-\lambda)\mathcal{L}_{\text{HDR}} + \lambda\mathcal{L}_{\text{D-SSIM}} + \mathcal{L}_{\text{reg}} \tag{9}$$

where $\mathcal{L}_{\text{D-SSIM}}(\hat{\boldsymbol{L}}, \boldsymbol{L})$ represents the differentiable Structural Similarity Index Metric (SSIM) [32] loss, and $\lambda = 0.2$ is the weighting parameter [10]. The $\mathcal{L}_{\text{reg}}$ term aggregates additional regularization terms described below.

2DGS. 2DGS [18] introduces distortion and normal consistency losses, defined respectively as

$$\mathcal{L}_{\text{dist}} = \sum_{i,j} \omega_i \omega_j |z_i - z_j| \tag{10}$$

$$\mathcal{L}_{\text{norm}} = \sum_{i} \omega_i (1 - \boldsymbol{n}_i^\top \boldsymbol{N}) \tag{11}$$

where $\omega_i = \alpha_i \prod_{j=1}^{i-1}(1-\alpha_j)$ denotes the blending weight (see Eq. 2), and $\boldsymbol{N}$ is the pseudo-normal map computed from the gradient of the predicted depth. The distortion loss aims to concentrate the Gaussian surfels on the surfaces by minimizing the distance between ray-splat intersections, while the normal consistency term promotes alignment between the surfel normals and the underlying surface geometry.

Smoothness Loss. The bilateral smoothness loss [24] penalizes abrupt spatial variations in the composited material properties while preserving image edges, encouraging smoothness in homogeneous regions while allowing discontinuities aligned with image gradients. The loss is defined as

$$\mathcal{L}_{\text{smooth}} = \sum_{i} \|\nabla\lambda(\boldsymbol{p}_i)\| \exp(-\|\nabla \boldsymbol{L}(\boldsymbol{p}_i)\|) \tag{12}$$

where $\boldsymbol{p}_i$ denotes i^{th} pixel location and $\lambda \in \{\rho, m\}$ denotes the material properties to be regularized (roughness and metallic).

Binary Opacity. Given ground-truth binary masks ($\boldsymbol{M}$), we additionally optimize a binary cross-entropy loss on the accumulated opacity $\boldsymbol{O} = \sum_i T_i \alpha_i$ given as

$$\mathcal{L}_{\text{O}} = -\boldsymbol{M} \log \boldsymbol{O} - (1-\boldsymbol{M}) \log(1-\boldsymbol{O}) \tag{13}$$

which encourages the reconstruction of well-defined opaque surfaces expected for spaceborne targets.

IV. Experiments

A. Implementation

The pipeline is implemented based on the official 2DGS codebase [18], with several key modifications described below. We sample $K = 16$ incident Earth albedo rays $\boldsymbol{w}_i$ at 1° resolution and use them throughout training to evaluate L_{earth} via Eq. 7. To accelerate computation, the batched BRDF evaluation (Eq. 5) is implemented as a custom CUDA kernel, enabling efficient processing with arbitrary numbers of Gaussians and incident directions. All experiments are conducted on a single NVIDIA GeForce RTX 4090 GPU.

Backface Culling. The original 2DGS implementation flips surfels whose normals face away from the camera. However, we found that such dual-visibility degrades performance for targets with very thin structures such as solar panels. When the two sides exhibit different appearance or material properties, normal flipping leads to both sides being α-composited along the same viewing ray, resulting in physically inconsistent renderings or

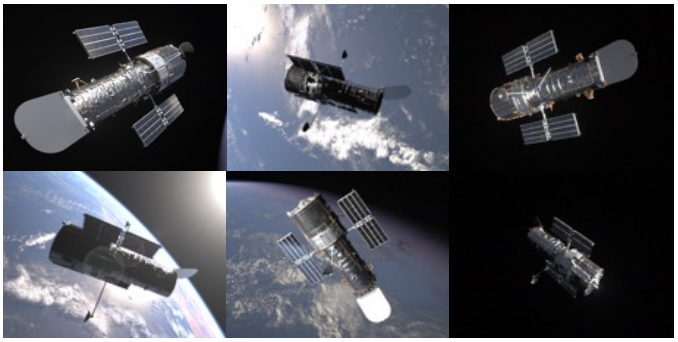

Fig. 3: Zoomed-in examples of DLVS^3-HST [34].

spurious surfels within thin structures. To address this issue, we instead cull back-facing surfels in Eq. 2.

MCMC Densification. The original 3DGS [10] employs a set of heuristics in which Gaussians are cloned, split and pruned based on positional gradient magnitudes and scales. This approach relies on multiple hand-tuned thresholds and does not inherently impose a bound on the number of Gaussians during training, which is undesirable for performance-critical space applications. While an upper bound can be enforced, doing so may interfere with these heuristic operations and lead to suboptimal distributions. Instead, we adopt the Markov Chain Monte Carlo (MCMC)-based strategy of Kheradmand et al. [33], which treats each Gaussian as an MCMC sample and reformulates densification as deterministic state transitions with a single opacity-based threshold. This naturally maintains a fixed number of Gaussians throughout training, enabling predictable resource usage compatible with spacecraft constraints. In this work, the number of Gaussians is fixed at 250K.

B. Dataset

The proposed method is trained and evaluated on DLVS^3-HST [34], which provides high-fidelity synthetic imagery of the Hubble Space Telescope (HST) (Fig. 3). Unlike existing vision-based RPO datasets that rely on simplified textures and idealized lighting, DLVS^3-HST incorporates realistic spacecraft materials, dynamic illumination (including Earth albedo via HDR environment maps), and sensor photoelectronic effects, resulting in highly realistic imagery. The HST model presents a particularly challenging target due to its complex geometry and appearance, including thin solar panels, a cylindrical hollow body, high-frequency MLI textures and both Earth-induced and self-reflections.

Training is performed on the first 1,000 HDR images, selected such that the Sun lies outside a 60° exclusion angle from the camera boresight, while evaluation is conducted on the subsequent 100 images. This setup reflects realistic mission constraints where cameras are not directed toward the Sun.

C. Metrics

In order to leverage the standard quantitative metrics that measure reconstruction quality–SSIM [32], Peak Signal-to-Noise Ratio (PSNR) and Learned Perceptual Image Patch Similarity (LPIPS) [35]–both HDR predictions and ground-truth images are tone-mapped to LDR in accordance with HDR-NeRF [30] using the following μ-law:

$$M(E) = \frac{\log(1+\mu E)}{\log(1+\mu)} \tag{14}$$

where μ = 5,000 denotes the amount of compression, and E is the HDR pixel value scaled to [0, 1]. For computing metrics, both ground-truth and predicted HDR images are normalized using the statistics of ground-truth for fair comparison.

TABLE I: Quantitative performances of different configurations. Arrows indicate the direction towards better performance. Colors denote the **best** , second best and third best performances.

Method	Shadow?	Earth?	SSIM (↑)	PSNR (↑)	LPIPS (↓)
Ref. [16]	✗	✗	0.9698	29.20	0.0260
Ref. [16]	✓	✗	0.9705	29.44	0.0257
Ours	✗	✗	0.9691	28.91	0.0273
Ours	✓	✗	0.9711	29.45	0.0273
Ours	✓	✓	**0.9747**	**31.19**	**0.0234**

D. Results

The reconstruction results are visualized in Fig. 4, where we present the full rendering along with its decomposed components–L_{sun}, shadow splatting (V), and L_{earth}. The examples highlight scenarios with strong Earth albedo effects, evident from the bluish tints and reflections on the HST's reflective surfaces. The L_{sun} column shows that a single BRDF evaluation with the known Sun direction already yields plausible reconstructions and material estimates, while shadow splatting captures self-occlusion. The L_{earth} component models Earth albedo by sampling K incident light directions, and its combination with $V \cdot L_{\text{sun}}$ recovers the joint effect of directional solar illumination and diffuse Earth reflection.

Table I further highlights the importance of incorporating Earth albedo into the rendering pipeline for analytical BRDF-based modeling as evidenced by consistent improvements observed across all evaluation metrics. The $V \cdot L_{\text{sun}}$ term alone is insufficient to explain the observed appearance, particularly the albedo effects on surfaces that are occluded or oriented away from the Sun. Performance significantly degrades without explicitly modeling these contributions.

When shadowing is not considered, the method of Park & D'Amico [16] outperforms the 2DGS + BRDF pipeline. This is primarily because their use of an MLP to model appearance provides greater flexibility to implicitly capture complex illumination effects, including partial occluhsion and Earthshine. In contrast, the analytical BRDF formulation imposes stronger physical constraints on light transport, limiting its ability to compensate for missing illumination terms. This underscores the necessity of explicitly modeling both shadowing and Earth albedo to achieve accurate and physically consistent reconstruction.

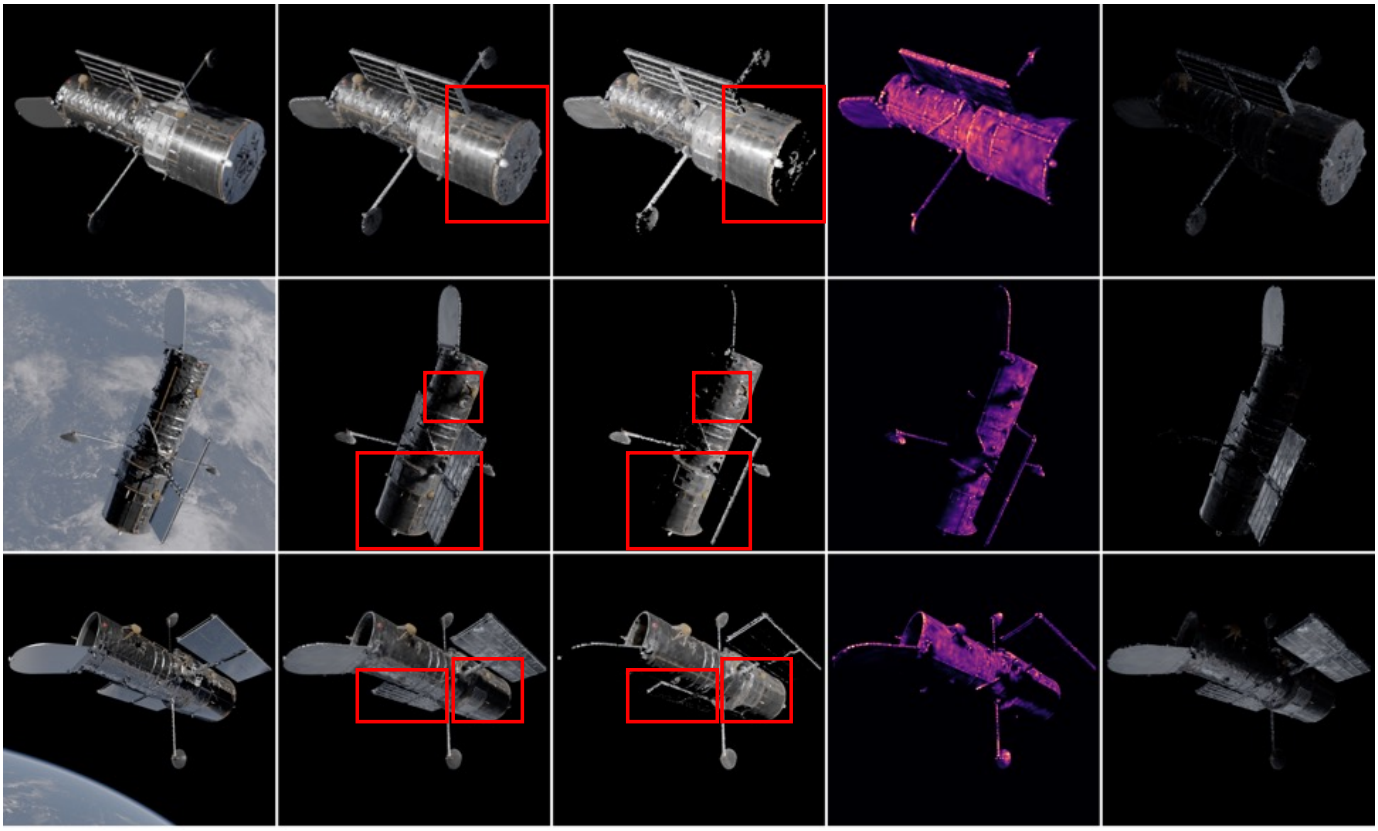

Fig. 4: Qualitative results of the proposed method. All images are tone-mapped from HDR values using μ-law. Red rectangles indicate regions of interest.

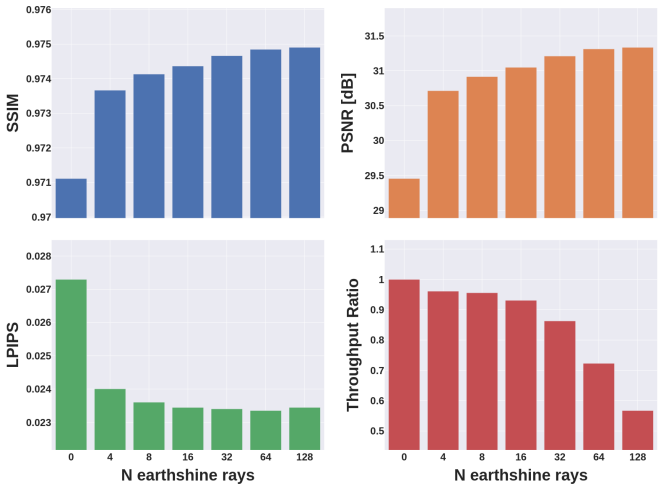

Fig. 5: Ablation study on different number of $\boldsymbol{w}_i$ samples used during training. Throughput ratio is with respect to $K = 0$, i.e., no Earth albedo modeling.

Finally, the number of incident Earth rays $\boldsymbol{w}_i$ is varied, and the performance is reported in Fig. 5. Notably, incorporating as few as four rays per training step yields a significant performance gain, while maintaining high training throughput (approximately $0.9\times$ up to 16 rays). Increasing the number of rays beyond this point leads to a substantial drop in efficiency with only marginal improvements in quantitative metrics.

V. Conclusions

This paper presented a physically-based 3D reconstruction framework for spaceborne targets in low Earth orbit (LEO) using 2D Gaussian Splatting. By leveraging an analytical formulation of light transport, the method enables an interpretable and photometrically consistent reconstruction under known illumination in LEO. In particular, explicitly modeling solar lighting, shadowing, and Earth albedo is shown to be critical for accurately capturing the appearance of complex spacecraft. The proposed framework achieves good material estimation and rendering quality, especially in scenarios involving strong reflections and shadowed regions, highlighting the importance of physically-based illumination modeling in space environments.

Future work includes integration into a simultaneous localization and mapping framework to validate the impact of improved photometric modeling on overall pose estimation performance and convergence behavior. Additionally, evaluating training time and memory efficiency on space-grade hardware is essential to assess the practicality of the proposed method for real-world deployment.

ACKNOWLEDGMENT

The author would like to thank Machine Intelligence Zrt for providing the RGB version of the DLVS^3-HST dataset.

REFERENCES

[1] B. B. Reed, R. C. Smith, B. J. Naasz, J. F. Pellegrino, and C. E. Bacon, "The Restore-L Servicing Mission," *AIAA Space 2016*, 2016.

[2] M. Pyrak and J. Anderson, "Performance of Northrop Grumman's Mission Extension Vehicle (MEV) RPO imagers at GEO," in *Autonomous Systems: Sensors, Processing and Security for Ground, Air, Sea and Space Vehicles and Infrastructure 2022*, M. C. Dudzik, S. M. Jameson, and T. J. Axenson, Eds., vol. 12115, International Society for Optics and Photonics. SPIE, 2022, p. 121150A.

[3] G. S. Aglietti, B. Taylor, S. Fellowes, S. Ainley, D. Tye, C. Cox, A. Zarkesh, A. Mafficini, N. Vinkoff, K. Bashford, T. Salmon, I. Retat, C. Burgess, A. Hall, T. Chabot, K. Kanani, A. Pisseloup, C. Bernal, F. Chaumette, A. Pollini, and W. H. Steyn, "RemoveDEBRIS: An in-orbit demonstration of technologies for the removal of space debris," *The Aeronautical Journal*, vol. 124, no. 1271, p. 1–23, 2020.

[4] S. Sharma and S. D'Amico, "Neural network-based pose estimation for noncooperative spacecraft rendezvous," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4638–4658, 2020.

[5] T. H. Park, M. Märtens, G. Lecuyer, D. Izzo, and S. D'Amico, "SPEED+: Next-generation dataset for spacecraft pose estimation across domain gap," in *2022 IEEE Aerospace Conference (AERO)*, 2022, pp. 1–15.

[6] P. F. Proença and Y. Gao, "Deep learning for spacecraft pose estimation from photorealistic rendering," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 6007–6013.

[7] M. A. Musallam, V. Gaudilliere, E. Ghorbel, K. A. Ismaeil, M. D. Perez, M. Poucet, and D. Aouada, "Spacecraft recognition leveraging knowledge of space environment: Simulator, dataset, competition design and analysis," in *2021 IEEE International Conference on Image Processing Challenges (ICIPC)*, 2021, pp. 11–15.

[8] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[9] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham: Springer International Publishing, 2020, pp. 405–421.

[10] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis, "3D gaussian splatting for real-time radiance field rendering," *ACM Trans. Graph.*, vol. 42, no. 4, Jul. 2023.

[11] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, "NICE-SLAM: Neural implicit scalable encoding for SLAM," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 12 776–12 786.

[12] A. Rosinol, J. J. Leonard, and L. Carlone, "NeRF-SLAM: Real-time dense monocular SLAM with neural radiance fields," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 3437–3444.

[13] H. Matsuki, R. Murai, P. H. J. Kelly, and A. J. Davison, "Gaussian splatting SLAM," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 18 039–18 048.

[14] C. Yan, D. Qu, D. Xu, B. Zhao, Z. Wang, D. Wang, and X. Li, "GS-SLAM: Dense visual SLAM with 3D gaussian splatting," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 19 595–19 604.

[15] K. R. Barad, A. Richard, J. Dentler, M. Olivares-Mendez, and C. Martinez, "Object-centric reconstruction and tracking of dynamic unknown objects using 3d gaussian splatting," in *2024 International Conference on Space Robotics (iSpaRo)*, 2024, pp. 202–209.

[16] T. H. Park and S. D'Amico, "Improved 3D gaussian splatting of unknown spacecraft structure using space environment illumination knowledge," 2025. [Online]. Available: https://arxiv.org/abs/2512.23998

[17] Z. Bi, Y. Zeng, C. Zeng, F. Pei, X. Feng, K. Zhou, and H. Wu, "GS3: Efficient relighting with triple gaussian splatting," in *SIGGRAPH Asia 2024 Conference Papers*, ser. SA '24. New York, NY, USA: Association for Computing Machinery, 2024.

[18] B. Huang, Z. Yu, A. Chen, A. Geiger, and S. Gao, "2D gaussian splatting for geometrically accurate radiance fields," in *ACM SIGGRAPH 2024 Conference Papers*. New York, NY, USA: Association for Computing Machinery, 2024.

[19] M. Zwicker, H. Pfister, J. van Baar, and M. Gross, "EWA volume splatting," in *Proceedings Visualization, 2001. VIS '01.*, 2001, pp. 29–538.

[20] A. Guédon and V. Lepetit, "SuGaR: Surface-aligned gaussian splatting for efficient 3D mesh reconstruction and high-quality mesh rendering," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 5354–5363.

[21] Y. Jiang, J. Tu, Y. Liu, X. Gao, X. Long, W. Wang, and Y. Ma, "GaussianShader: 3D gaussian splatting with shading functions for reflective surfaces," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 5322–5332.

[22] B. Burley, "Physically based shading at disney," 2012. [Online]. Available: https://media.disneyanimation.com/uploads/production/publication_asset/48/asset/s2012_pbs_disney_brdf_notes_v3.pdf

[23] B. Karis, "Real shading in unreal engine 4," 2013. [Online]. Available: https://blog.selfshadow.com/publications/s2013-shading-course/karis/s2013_pbs_epic_notes_v2.pdf

[24] Y. Yao, J. Zhang, J. Liu, Y. Qu, T. Fang, D. McKinnon, Y. Tsin, and L. Quan, "NeILF: Neural incident light field for physically-based material estimation," in *Computer Vision – ECCV 2022*, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds. Cham: Springer Nature Switzerland, 2022, pp. 700–716.

[25] K. Ye, Q. Hou, and K. Zhou, "3D gaussian splatting with deferred reflection," in *ACM SIGGRAPH 2024 Conference Papers*, ser. SIGGRAPH '24. New York, NY, USA: Association for Computing Machinery, 2024.

[26] D. Verbin, P. Hedman, B. Mildenhall, T. Zickler, J. T. Barron, and P. P. Srinivasan, "Ref-NeRF: Structured view-dependent appearance for neural radiance fields," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 5481–5490.

[27] J. T. Kajiya, "The rendering equation," *SIGGRAPH Comput. Graph.*, vol. 20, no. 4, 1986.

[28] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graph.*, vol. 1, no. 1, p. 7–24, Jan. 1982.

[29] Y. Cai, Z. Xiao, Y. Liang, M. Qin, Y. Zhang, X. Yang, Y. Liu, and A. Yuille, "HDR-GS: Efficient high dynamic range novel view synthesis at 1000x speed via gaussian splatting," in *Advances in Neural Information Processing Systems*, A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, Eds., vol. 37. Curran Associates, Inc., 2024, pp. 68 453–68 471. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2024/file/7e83fd2ff7ae58485d418685521c9608-Paper-Conference.pdf

[30] X. Huang, Q. Zhang, Y. Feng, H. Li, X. Wang, and Q. Wang, "HDR-NeRF: High dynamic range neural radiance fields," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 18 377–18 387.

[31] S. Singh, A. Garg, and K. Mitra, "HDRSplat: Gaussian splatting for high dynmaic range 3D scene reconstruction from raw images," in *35th British Machine Vision Conference 2024, BMVC 2024, Glasgow, UK, November 25-28, 2024*. BMVA, 2024. [Online]. Available: https://papers.bmvc2024.org/0022.pdf

[32] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[33] S. Kheradmand, D. Rebain, G. Sharma, W. Sun, Y.-C. Tseng, H. Isack, A. Kar, A. Tagliasacchi, and K. M. Yi, "3d gaussian splatting as markov chain monte carlo," in *Advances in Neural Information Processing Systems*, A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, Eds., vol. 37. Curran Associates, Inc., 2024, pp. 80 965–80 986. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2024/file/93be245fce00a9bb2333c17ceae4b732-Paper-Conference.pdf

[34] S. Velkei, C. Goldschmidt, and K. Vass, "A large-scale, physically-based synthetic dataset for satellite pose estimation," 2025. [Online]. Available: https://arxiv.org/abs/2506.12782

[35] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 586–595.

[36] C. Schlick, "An inexpensive brdf model for physically-based rendering," *Computer Graphics Forum*, vol. 13, no. 3, pp. 233–246, 1994.

APPENDIX A
COOK-TORRANCE BRDF IMPLEMENTATION

This section provides the adopted models for each term of the Cook-Torrance BRDF equation in Eq. 5. The terms here largely follow Karis [23] and Burley [22]. Recall that all these terms and the BRDF model are implemented in a custom CUDA kernel with analytical gradients in order to speed up the computation of multiple Earth albedo effects.

Normal Distribution Function (NDR). We use GGX/Trowbridge-Reitz specular NDF

$$D(\boldsymbol{h}) = \frac{\alpha^2}{\pi((\boldsymbol{n}\cdot\boldsymbol{h})^2(\alpha^2-1)+1)^2} \tag{15}$$

where $\alpha = \rho^2$ is the reparametrization adopted by Disney [22].

Fresnel Term. Schlick's approximation [36]

$$F = F_0 + (1-F_0)(1-\boldsymbol{h}\cdot\boldsymbol{w}_o)^5 \tag{16}$$

where

$$F_0 = 0.04(1-m) + m\boldsymbol{a} \tag{17}$$

Geometric Attenuation. We also use Schlick's approximation [36]

$$G_1(\boldsymbol{w}) = \frac{\boldsymbol{n}\cdot\boldsymbol{w}}{(\boldsymbol{n}\cdot\boldsymbol{w})(1-k)+k} \tag{18}$$

where $k = \rho^2/2$. Then, the specular geometric attenuation term is given as $G(\boldsymbol{w}_i, \boldsymbol{w}_o) = G_1(\boldsymbol{w}_i)G_1(\boldsymbol{w}_o)$.