RT-Splatting: Joint Reflection-Transmission Modeling with Gaussian Splatting

Ji Shi Xianghua Ying* Bowei Xing Ruohao Guo Wenzhen Yue

State Key Laboratory of General Artificial Intelligence
School of Intelligence Science and Technology
Peking University

Figure 1. **Photorealistic rendering and decomposition of real-world scene with coexisting reflection and transmission.** (Left) Compared to prior works, our method robustly handles semi-transparent surfaces, avoiding blurry reflections or overly occluded transmission. (Right) Our high-fidelity results are achieved by decomposing the scene radiance into Reflection and Transmission layers, enabled by a unified Gaussian representation that jointly captures surface geometry and scene volume.

Abstract

3D Gaussian Splatting (3DGS) enables real-time novel view synthesis with high visual quality. However, existing methods struggle with semi-transparent specular surfaces that exhibit both complex reflections and clear transmission, often producing blurry reflections or overly occluded transmission. To address this, we present ***RT-Splatting****, a framework that disentangles each Gaussian's geometric occupancy from its optical opacity. This factorization yields a unified surface-volume scene representation with a single set of Gaussian primitives. Our hybrid renderer interprets this representation both as a surface to capture high-frequency reflections and as a volume to preserve clear transmission. To mitigate the ambiguity in jointly optimizing reflection and transmission, we introduce Specular-Aware Gradient Gating, which suppresses misleading gradients from highly specular regions into the transmission branch, effectively reducing distracting floaters. Experiments on challenging semi-transparent scenes show that RT-Splatting achieves state-of-the-art performance, delivering high-fidelity reflections and clear transmission with real-time rendering. Moreover, our factorization naturally enables flexible scene editing. The project page is available at https://sjj118.github.io/RT-Splatting.*

1. Introduction

3D Gaussian Splatting (3DGS) [18] has revolutionized the field of novel view synthesis with its real-time rendering capabilities, achieved by representing a scene as a sparse set of 3D Gaussian primitives and rendering them efficiently via rasterization. Despite its success, 3DGS struggles to model semi-transparent specular surfaces where reflection and transmission coexist. To reproduce high-

*Corresponding author.

frequency specular highlights, standard 3DGS often hallucinates "floaters" behind the surface. These behind-surface floaters not only fail to faithfully capture the true reflected appearance, but also corrupt transmission by spuriously occluding background geometry that should remain visible through the surface.

Recent 3DGS variants address high-frequency view-dependent effects by replacing per-Gaussian SH with physically based shading that explicitly evaluates the rendering equation using scene geometry, material properties, and incident illumination [10, 12, 17, 25, 33]. Early work performed shading at the primitive level (per-Gaussian), whereas more recent methods adopt deferred shading [4, 6, 7, 22, 34, 40–45, 50, 51], which first rasterizes scene properties into G-buffers and then performs per-pixel shading. However, because the G-buffer stores only the properties of the nearest surface at each pixel, transparency is fundamentally difficult to handle in deferred shading [2]. For semi-transparent specular surfaces exhibiting both reflection and transmission, these methods either fail to aggregate the target surface's attributes needed for reflection modeling or simply treat the surface as opaque, completely occluding transmission. TransparentGS [15] instead adopts a multi-stage pipeline, where transparent objects are modeled with a separate set of Gaussian primitives on top of a background reconstructed in a separate stage using standard 3DGS while masking out transparent regions. Because the background is reconstructed without seeing through the transparent objects, the method struggles in scenes where the background is exclusively visible through the transparent surfaces, such as viewing a car's interior solely through its windows.

In this paper, we present *RT-Splatting*, a hybrid surface-volume rendering framework for jointly modeling reflection and transmission in real-world scenes containing thin semi-transparent surfaces. For each Gaussian primitive, we decouple its role as a surface element from its role in attenuating light along the ray. Specifically, we factorize its contribution into a geometric occupancy term and an optical opacity term, thereby enabling a unified surface-volume scene representation with a single set of Gaussians. The geometric occupancy determines how strongly the Gaussian participates as a surface element along a ray, while the optical opacity controls how much light is absorbed or scattered once that surface is hit. This unified representation naturally supports a hybrid rendering pipeline: geometric occupancy is used to aggregate first-hit surface attributes into G-buffers for deferred reflection shading, whereas optical opacity drives a volumetric forward pass that accumulates transmitted background radiance.

However, even with this surface-volume formulation, jointly optimizing reflection and transmission remains ambiguous. High-frequency specular reflections are inherently difficult to fit, and the residual errors tend to produce misleading gradients that leak into the transmission branch. This causes the transmission component to compensate by creating erroneous floaters that corrupt background clarity. To mitigate this issue, we introduce a *Specular-Aware Gradient Gating* mechanism that identifies pixels dominated by complex specular patterns and attenuates the corresponding gradients flowing to the transmission branch. This gating suppresses misleading supervision, substantially reduces distracting floaters, and improves the clarity of the transmitted background.

To summarize, our contributions are as follows:

- We introduce a unified surface-volume Gaussian scene representation for jointly modeling sharp specular reflections and clear transmission in real-world scenes containing thin semi-transparent surfaces.
- We propose Specular-Aware Gradient Gating to suppress misleading gradients from complex specular regions, substantially reducing floaters in the transmission branch.
- Extensive experiments demonstrate that RT-Splatting significantly outperforms prior methods while maintaining real-time rendering and enabling flexible scene editing.

2. Related Work

2.1. Reflective Scene Reconstruction

The reconstruction and rendering of reflective scenes is a long-standing challenge in novel view synthesis. Ref-NeRF [35] conditions outgoing radiance on the reflection direction, rather than the viewing direction, improving the capture of high-frequency specular effects. Subsequent works advance this idea by either strengthening directional encodings to better capture light-surface interactions [24, 26–28] or recovering more accurate surface geometry [11, 32, 38] to mitigate shape-radiance ambiguity [47]. To address the challenge of rendering consistent reflections of nearby content, NeRF-Casting [36] performs cone tracing along reflection paths and aggregates features before decoding, yielding high-fidelity inter-reflections. However, these approaches rely on dense implicit-field queries along rays during both training and inference, making real-time rendering impractical.

In contrast, recent works leveraging Gaussian Splatting have achieved real-time rendering capabilities for reflective scenes. GaussianShader [17] estimates per-Gaussian normals from the shortest axis and shades with a learnable environment map for efficient specular shading. 3DGS-DR [43] adopts a deferred pipeline that first rasterizes scene attributes into G-buffers and then performs per-pixel shading. Ref-GS [50] extends 2DGS [14] with a directional factorization for spatio-angular view-dependent effects. EnvGS [41] further employs a differentiable Gaussian ray tracer with environment Gaussians to capture near-field re-

flections in real time. While these methods excel at representing high-frequency specular effects, they still struggle with thin, semi-transparent surfaces whose appearance is a mixture of light transmitted through the surface and light reflected from the surface.

2.2. Transparent Object Reconstruction

While the native alpha-blending in volumetric methods like NeRF [29] and 3DGS [18] can simulate translucency, it does so by conflating the geometric presence of a surface with its optical transmissivity, preventing it from establishing a distinct surface geometry required for physically-based shading. To circumvent this ambiguity, one line of research [1, 16, 23] explores various first-surface extraction strategies to explicitly recover the surface of transparent object. Other works focus on the challenging task of reconstructing the complex view-dependent appearance on the transparent surface. To make this highly ill-posed problem tractable, a predominant strategy involves decoupling the object from its environment. This is typically achieved either by employing multi-stage pipelines to pre-reconstruct and freeze the opaque background [3, 9, 15], or by simplifying the background to an infinitely distant environment map [5, 37]. Such approaches, however, are not applicable to general, complex scenes where transparent object and diffuse background are photometrically entangled (*e.g.*, when the background is only visible through the transparent surface). Some approaches further impose strong constraints on the scene configuration, such as assuming simplified geometry like planar surfaces [19] or requiring controlled capture conditions like forward-facing camera arrangements [46].

The restrictions in these methods often stem from the inherent ill-posedness of disentangling reflection and refraction, since both phenomena are highly view-dependent and lack the multi-view photometric consistency. To avoid this challenge, a practical approach is to focus on ubiquitous thin semi-transparent surfaces, such as glass panes or plastic films, where the negligible refractive effect allows light transport to be approximated as straight-path transmission. Following this direction, recent works [8, 49] have shown promise in jointly modeling reflection and transmission, but their applicability remains limited to simple planar surfaces, failing to generalize to complex shapes.

3. Preliminaries

3.1. Gaussian Splatting

3D Gaussian Splatting (3DGS) [18] has recently emerged as a powerful technique for real-time, high-fidelity novel view synthesis. It represents a 3D scene with a collection of anisotropic 3D Gaussian primitives, each defined by its position, covariance, opacity α, and color represented by Spherical Harmonics (SH). During rendering, these 3D Gaussians are projected onto the 2D image plane and sorted by depth. The final color $\mathbf{C}$ for a pixel is then computed by alpha blending the Gaussians in front-to-back order:

$$\mathbf{C} = \sum_i w_i \mathbf{c}_i \,, \text{ where } \quad w_i = \alpha_i \mathcal{G}_i \prod_{j=1}^{i-1} (1 - \alpha_j \mathcal{G}_j) \,, \quad (1)$$

where $\mathbf{c}_i$ and α_i are the color and opacity of the i-th Gaussian, and $\mathcal{G}_i$ is the value of its projected 2D Gaussian kernel at the pixel center.

To better align the scene representation with surfaces, recent work has proposed 2D Gaussian Splatting (2DGS) [14]. Instead of 3D primitives, 2DGS models the scene as a set of 2D Gaussian surfels embedded in 3D space. This surface-aligned representation provides each primitive with a well-defined surface normal, typically derived from the orientation of the 2D disk. Furthermore, it mitigates the multi-view depth inconsistency issues that can arise from projecting 3D Gaussians, leading to a more geometrically accurate surface representation. Our work builds upon this 2DGS framework.

3.2. Deferred Shading

Deferred shading is a two-pass rendering technique that decouples geometry processing from lighting and material computations. In the first pass, known as the geometry pass, various attributes of the nearest surface, such as depth, normal, albedo, and roughness, are rendered into a set of intermediate 2D buffers, collectively called G-buffers. In the second pass, a shading program is executed for each pixel, using the information stored in the G-buffers to compute the final color. Recent works [4, 6, 7, 22, 34, 40–45, 50, 51] have successfully adapted this pipeline to Gaussian Splatting to efficiently render high-frequency, view-dependent effects. By performing complex shading calculations on a per-pixel basis rather than a per-Gaussian basis, deferred shading significantly enhances rendering quality and performance for complex materials.

4. Method

Our method is designed to reconstruct scenes with thin semi-transparent surfaces that exhibit both sharp reflections and clear transmission. We factorize the per-Gaussian opacity into geometric occupancy and optical opacity (Sec. 4.1), yielding a unified surface-volume representation that supports a hybrid pipeline for rendering reflections and transmission (Sec. 4.2). To suppress floaters caused by residual reflection errors, we introduce Specular-Aware Gradient Gating (Sec. 4.3), and we finally describe the optimization details in Sec. 4.4. An overview of the framework is shown in Fig. 2.

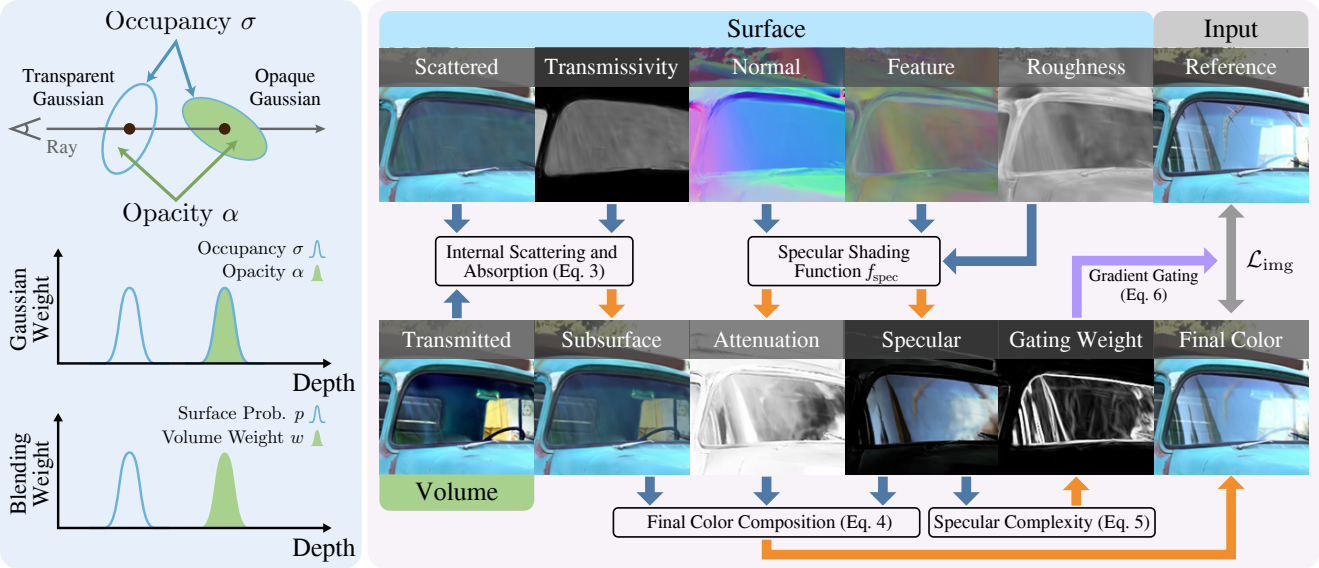

Figure 2. **Overview of RT-Splatting.** (Left) Transparent objects are represented by Gaussians with high geometric occupancy but low optical opacity, yielding strong contributions to surface aggregation while avoiding occlusion during volumetric compositing. (Right) Our hybrid rendering pipeline composites surface-based reflections from a deferred pass with volumetric transmission from a forward pass to produce the final color, and suppresses image-loss gradients flowing into the transmission branch according to the specular complexity of the corresponding pixels.

4.1. Occupancy-Opacity Factorization

The standard Gaussian Splatting pipeline uses a single opacity parameter for alpha blending, primarily to model optical occlusion. While recent deferred shading methods [41, 43, 50] have successfully repurposed this opacity to rasterize surface properties into G-buffers, this formulation conflates a Gaussian's geometric presence with its optical properties. This approximation is reasonable for opaque objects, but it fails fundamentally for semi-transparent surfaces, such as windows or plastic films. For these materials, the surface is geometrically solid (required for rendering sharp reflections) yet optically clear (allowing light transmission). A single opacity parameter cannot simultaneously satisfy these conflicting demands, leading to either blurry reflections or an opaque appearance.

To address this limitation, we factorize the standard per-Gaussian opacity into two physically motivated, learnable attributes. The *geometric occupancy* $\sigma \in [0, 1]$ encodes the probability that a ray interacts with the substance of the Gaussian. The *optical opacity* $\alpha \in [0, 1]$ then specifies the conditional probability that the ray is absorbed or scattered once such an interaction occurs. Their product $\alpha_{\text{eff}} = \sigma\alpha$ defines the effective opacity used for volumetric compositing in Eq. (1), meaning that optical attenuation only happens where the Gaussian is geometrically present. This factorization enables us to model transparent objects using Gaussians with high geometric occupancy but low optical opacity.

Our factorization naturally yields a probabilistic formulation for first-surface extraction, which is essential for deferred shading. Given a sequence of Gaussians along a ray sorted by depth, the expected value of any surface attribute $\mathbf{a}$ (*e.g.*, normal or roughness) is computed as:

$$\mathbf{A} = \sum_i p_i \mathbf{a}_i, \text{ where } \quad p_i = \sigma_i \mathcal{G}_i \prod_{j=1}^{i-1} (1 - \sigma_j \mathcal{G}_j) . \quad (2)$$

Here, p_i represents the probability that the i-th Gaussian is the first surface element with which the ray interacts.

While mathematically analogous to standard alpha blending, our formulation provides a crucial reinterpretation: we treat the collection of Gaussians not as discrete, semi-transparent surfels, but as a unified, probabilistic representation of a single surface. This physically-grounded view justifies the application of deferred shading for high-frequency reflection modeling in Gaussian Splatting.

4.2. Reflection-Transmission Modeling

To model the complex appearance of semi-transparent surfaces, which involves both high-frequency specular reflections and transmitted light, we propose a hybrid deferred-forward rendering framework.

Our framework begins with a deferred pass to handle high-frequency specular reflections on the first-hit surface. Leveraging our occupancy-opacity factorization, we first aggregate the expected surface properties into G-buffers

using the probabilistic formulation in Eq. (2). Once the G-buffers are populated, a specular shading function f_{spec} takes the view direction and a set of surface attributes, including normal $\mathbf{n}$, roughness ρ, and material feature $\mathbf{z}$, as input to compute the specular color $\mathbf{C}_{\text{spec}}$ for each pixel. This function is designed to reproduce complex, view-dependent specular effects, capturing reflections from the surrounding environment. For our implementation, we adopt a specular shading network architecture similar to that in Ref-GS [50], which has proven effective for this task.

To capture the intrinsic appearance of materials like colored glass, which involves internal scattering and absorption, we introduce two additional surface attributes for each Gaussian: an intrinsic scattered color $\mathbf{C}_{\text{scatter}}$ and a transmissivity ratio $\tau \in [0, 1]$. $\mathbf{C}_{\text{scatter}}$ represents light scattered back from within the material, while τ dictates the material's transmissivity by controlling the balance between this scattered light and the transmitted background light.

The background radiance itself, $\mathbf{C}_{\text{trans}}$, is computed with a concurrent forward pass. This pass operates like standard volumetric rendering, accumulating color from the opaque background scene. Crucially, it is accumulated with our effective opacity $\alpha_{\text{eff}} = \sigma\alpha$. This formulation allows the background scene to be correctly accumulated without being occluded by the transparent objects. We group all radiance that travels inside the material, including both transmitted and scattered components, into a subsurface-transport component in our formulation:

$$\mathbf{C}_{\text{sub}} = \tau\mathbf{C}_{\text{trans}} + (1 - \tau)\mathbf{C}_{\text{scatter}} \,. \tag{3}$$

Finally, we combine the specular reflection $\mathbf{C}_{\text{spec}}$ and the subsurface-transport component $\mathbf{C}_{\text{sub}}$ to produce the final pixel color. A purely physics-based blend using Fresnel equations is often broken in practice by tone-mapping and other nonlinear camera responses, and fails to capture our key perceptual observation: transmitted details are clearly visible through faint reflections but are suppressed or even masked by strong specular highlights. To model this dynamic effect, we augment our specular shading function to also output an attenuation factor $\beta \in [0, 1]$ that directly modulates the subsurface-transport component. The final color $\mathbf{C}$ is then computed as:

$$\mathbf{C} = \mathbf{C}_{\text{spec}} + \beta\mathbf{C}_{\text{sub}} \,. \tag{4}$$

Unlike previous methods [8, 13, 49] that modulate the reflection component, our approach attenuates the transmitted component, which provides a direct and stable mechanism to model the suppression of background light.

4.3. Specular-Aware Gradient Gating

While our hybrid deferred-forward rendering pipeline cleanly separates how reflection and transmission are rendered, jointly optimizing both branches remains challenging. High-frequency specular reflections are inherently difficult to model perfectly, leaving residual discrepancies between the rendered reflections and the ground-truth observations. During backpropagation, gradients induced by these residuals can be erroneously routed into the transmission branch, which then compensates by hallucinating spurious floaters behind the surface and degrading the clarity of the transmitted background.

To mitigate this erroneous gradient flow, we introduce a specular-aware gradient gating mechanism. Our key insight is that this incorrect compensation primarily occurs in image regions with high-frequency specular details. We identify these regions by using the local variance of the specular component, $\mathbf{C}_{\text{spec}}$, to estimate its complexity. For each pixel x, we compute a gating weight $g(x)$ over a small neighboring patch $\mathcal{N}(x)$:

$$g(x) = \exp\left(-k \cdot \text{Var}_{p\in\mathcal{N}(x)}[\mathbf{C}_{\text{spec}}(p)]\right) \,, \tag{5}$$

where $\text{Var}(\cdot)$ is the variance operator and k is a hyperparameter controlling the gate's sensitivity.

During the backward pass, this gating weight modulates the gradients flowing to the transmission branch. Specifically, we apply $g(x)$ to scale the gradient of the image loss $\mathcal{L}_{\text{img}}$ that backpropagates through the transmitted background color $\mathbf{C}_{\text{trans}}$:

$$\frac{\partial\mathcal{L}_{\text{img}}}{\partial\mathbf{C}_{\text{trans}}(x)} \leftarrow g(x) \cdot \frac{\partial\mathcal{L}_{\text{img}}}{\partial\mathbf{C}_{\text{trans}}(x)} \,. \tag{6}$$

In other words, this specular-aware gradient gating attenuates gradients at pixels dominated by complex specular patterns, but does not completely block supervision of the background scene behind the semi-transparent surface. At viewpoints and pixels where specular reflections are simple or weak, $g(x)$ remains close to one, so the transmitted background continues to receive full supervision through the transparent interface. This preserves a valid optimization path for the background geometry and appearance.

4.4. Optimization

Transparent mask regularization. Our occupancy-opacity factorization introduces a specific ambiguity: Gaussians with high geometric occupancy but near-zero optical opacity can exist anywhere in the scene without affecting the final rendered color. This is particularly problematic in diffuse regions lacking strong specular cues, where these unconstrained "ghost" geometries can accumulate, corrupting the surface representation and destabilizing the optimization process.

To resolve this ambiguity, we introduce a transparent mask loss that provides explicit supervision for the optical opacity of Gaussians. We leverage a transparent mask

Figure 3. **Qualitative comparisons on test-set views of real-world scenes.** Our method significantly improves rendering quality over previous approaches, simultaneously yielding sharper reflections and clearer transmissions in semi-transparent regions.

$\mathbf{M}$, obtained from the pre-trained SAM2 model [20, 31] to provide additional supervision. During the deferred pass, we aggregate the expected optical opacity α of the first-hit surface into G-buffers. We then supervise this opacity map with a binary cross-entropy (BCE) loss, encouraging it to match the inverted semantic mask:

$$\mathcal{L}_{\text{mask}} = \text{BCE}(1 - \mathbf{M}, \alpha)\,. \tag{7}$$

Joint optimization. We perform a joint optimization of all system components, simultaneously refining the Gaussian primitives, their factorized opacities, and the shading function. Unlike prior works [3, 15, 23] that use the transparent mask to segment the scene for separate processing, our approach integrates the mask solely as a regularization. This joint optimization is crucial as it makes our method applicable to complex scenes where the background is exclusively visible through the transparent surfaces.

5. Experiments

5.1. Implementation Details

We implement RT-Splatting in PyTorch [30] building upon the 2DGS framework [14]. The hyperparameters for the shading function in our deferred pass are kept consistent

Table 1. **Quantitative results on scenes from Ref-Real [35], NeRF-Casting [36], EnvGS [41] and T&T [21].** We report PSNR, SSIM [39], and LPIPS [48] on entire images and over transparent regions, together with rendering speed (FPS) and training time. ↑ (↓) indicates higher (lower) is better. We mark the best, the second best, and the third best results in each column.

Methods	Entire Image			Transparent Region			FPS ↑	Training Time ↓
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓		
3DGS [18]	26.493	0.816	0.181	37.673	0.990	0.012	**218.95**	**0.3h**
2DGS [14]	26.384	0.817	0.197	37.333	0.990	0.012	208.82	**0.3h**
GShader [17]	25.778	0.806	0.203	36.797	0.988	0.014	24.59	1.3h
3DGS-DR [43]	26.597	0.816	0.190	37.890	0.990	0.012	119.62	0.8h
Ref-GS [50]	26.599	0.819	0.188	37.761	0.989	0.013	38.41	0.8h
EnvGS [41]	27.141	0.821	0.182	37.953	0.990	0.012	18.31	2.9h
Ours	27.490	0.831	0.167	39.765	0.992	0.010	33.28	0.9h

Table 2. **Quantitative results on our self-captured scenes.**

Methods	Entire image			Transparent region		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
3DGS [18]	27.507	0.863	0.213	32.567	0.964	0.048
2DGS [14]	26.675	0.857	0.243	31.923	0.965	0.053
GShader [17]	21.133	0.801	0.367	26.995	0.944	0.084
3DGS-DR [43]	26.134	0.858	0.229	32.014	0.964	0.049
Ref-GS [50]	26.301	0.851	0.240	31.646	0.962	0.056
EnvGS [41]	26.847	0.847	0.260	31.726	0.963	0.057
Ours	28.780	0.871	0.197	35.490	0.970	0.042

with those in Ref-GS [50]. Please refer to the supplementary materials for more details.

5.2. Datasets and Metrics

We evaluate our method on several real-world datasets that prominently feature the coexistence of high-frequency specular reflections and clear transmission on semi-transparent surfaces. From public benchmarks, we select six scenes: *Sedan* and *Toycar* from Ref-Real [35], *Compact* and *Hatchback* from NeRF-Casting [36], *Audi* from EnvGS [41], and *Truck* from T&T [21]. We additionally captured two real-world scenes, *Van* and *Swab*, using a smartphone camera to capture 220 $\sim$ 240 views for each scene.

We report PSNR, SSIM [39], and LPIPS [48] measured on both entire images and transparent regions. Additional results are provided in the supplementary materials.

5.3. Baseline Comparisons

We conduct a comprehensive comparison of our method against several state-of-the-art Gaussian Splatting variants. The baselines include the foundational 3DGS [18], 2DGS [14], and methods specifically designed for reflective surfaces. Among these are GaussianShader [17], as well as the recent deferred shading-based approaches 3DGS-DR [43], Ref-GS [50], and EnvGS [41]. All baseline models are trained using their publicly available codebases and configurations.

We present quantitative results on both public benchmarks and our self-captured scenes in Tab. 1 and Tab. 2. The results consistently show that our method outperforms all baselines across all evaluated metrics. This performance advantage is particularly significant when evaluating the transparent regions, highlighting our model's enhanced capability in these challenging areas. Notably, our method achieves real-time rendering speeds and maintains a competitive training time, proving its efficiency and practical applicability.

Table 3. **Ablation study of our model in transparent regions.** We mark the **best** results in each column.

	PSNR ↑	SSIM ↑	LPIPS ↓
w/o occupancy	36.919	0.9885	0.0113
w/o joint optimization	36.288	0.9876	0.0120
w/o scattering	37.597	0.9897	0.0102
w/o attenuation	37.541	0.9897	0.0102
w/o gating	37.754	0.9899	0.0101
w/o $\mathcal{L}_{\text{mask}}$	37.167	0.9894	0.0106
Ours	**37.983**	**0.9901**	**0.0095**

Qualitative comparisons presented in Fig. 3 further demonstrate our approach's unique capability to faithfully render both sharp reflection details and clear transmitted light simultaneously. Existing methods struggle with the inherent ambiguity between reflection and transmission on semi-transparent surfaces. As illustrated, they often fail to reconstruct sharp reflection details, as the optimization is compromised by the underlying transmitted light. Conversely, attempts to model strong reflections typically result in the surface being rendered as opaque, thereby sacrificing transmission clarity and occluding the background scene entirely.

5.4. Ablation Studies

We validate the effectiveness of our key components through ablation studies on *Sedan* and *Truck*. Quantitative and qualitative comparisons are presented in Tab. 3 and Fig. 5, respectively.

Occupancy-opacity factorization. The "w/o occupancy" variant removes our occupancy-opacity factorization, forcing a single opacity parameter to model both geometric

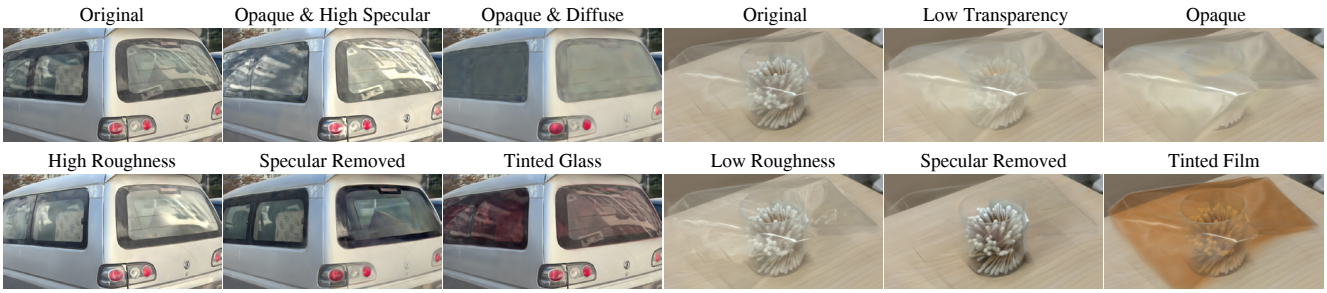

Figure 4. **Scene Editing.** Left: edited appearances of car windows. Right: edited appearances of a plastic film.

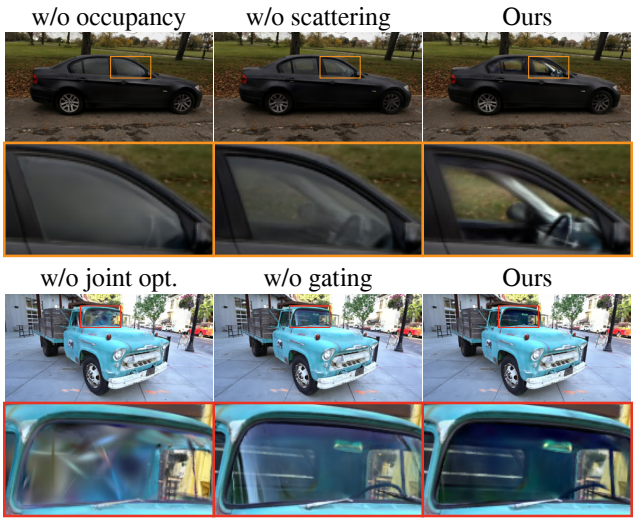

Figure 5. **Decomposed transmission components across ablation settings.**

occupancy and optical opacity. As shown in Tab. 3 and Fig. 5, this creates a conflict where achieving sharp reflections requires high opacity, which in turn severely compromises transmission clarity, leading to a more occluded background.

Joint optimization. The "w/o joint optimization" variant separates the training of the reflection and transmission components. As shown in Tab. 3 and Fig. 5, this completely prevents the reconstruction of the interior of the truck, which is exclusively visible through its windows.

Internal scattering and absorption. The "w/o scattering" variant removes the components responsible for the material's intrinsic appearance, namely the scattered color $\mathbf{C}_{\text{scatter}}$ and the transmissivity ratio τ. This forces the model to bake the material's intrinsic appearance into the volumetric background scene, resulting in a much darker transmission, as shown in Fig. 5 and Tab. 3.

Subsurface attenuation. The "w/o attenuation" variant removes the learnable attenuation factor β. This fails to model the view-dependent suppression of subsurface component, resulting in reduced rendering quality as demonstrated in Tab. 3.

Specular-aware gradient gating. The "w/o gating" variant disables our Specular-Aware Gradient Gating mechanism discussed in Sec. 4.3. This causes the transmission branch to generate visual artifacts near the transparent surface as shown in Fig. 5 and Tab. 3.

Transparent mask regularization. The "w/o $\mathcal{L}_{\text{mask}}$" variant removes the transparent mask regularization. This leads to unstable optimization and degraded surface quality as demonstrated in Tab. 3.

5.5. Applications

Our method's reflection-transmission decomposition naturally facilitates a variety of powerful scene editing applications. As illustrated in Fig. 4, we can independently manipulate surface attributes by adjusting roughness, changing its transparency level, removing specular reflections, or even altering the material's tint. This demonstrates the effectiveness and intuitive control offered by our decoupled representation.

6. Conclusion

We presented RT-Splatting, a framework for jointly modeling high-fidelity reflections and clear transmissions on semi-transparent surfaces. By disentangling each Gaussian primitive's geometric occupancy from its optical opacity, a single set of Gaussians supports a hybrid renderer that simultaneously interprets the scene as a reflective surface for sharp specular highlights and as a transmissive volume for clear background content. Furthermore, we introduce a specular-aware gradient gating mechanism to mitigate optimization ambiguities between the reflection and transmission components. Together, these designs enable RT-Splatting to achieve state-of-the-art performance on challenging scenes where reflection and transmission are strongly coupled.

A limitation of RT-Splatting is that it is designed for thin semi-transparent surfaces, as it does not explicitly model refraction or multiple light bounces. Future work could explore extending our framework to handle thicker refractive media and multi-bounce light transport, such as in water or solid glass objects.

Acknowledgments

This work was supported by the Beijing Natural Science Foundation under Grant No. L247029, and National Natural Science Foundation of China (NSFC) under Grant No. 62371009.

References

[1] Aviral Agrawal, Ritaban Roy, Bardienus Pieter Duisterhof, Keerthan Bhat Hekkadka, Hongyi Chen, and Jeffrey Ichnowski. Clear-splatting: Learning residual gaussian splats for transparent object manipulation. In *RoboNerF: 1st Workshop On Neural Fields In Robotics at ICRA 2024*, 2024. 3

[2] Ali Deniz Aladağlı. Deferred shading of transparent surfaces with shadows and refraction. Master's thesis, Middle East Technical University (Turkey), 2015. 2

[3] Mojtaba Bemana, Karol Myszkowski, Jeppe Revall Frisvad, Hans-Peter Seidel, and Tobias Ritschel. Eikonal fields for refractive novel-view synthesis. In *ACM SIGGRAPH 2022 Conference Proceedings*, New York, NY, USA, 2022. Association for Computing Machinery. 3, 6

[4] Hongze Chen, Zehong Lin, and Jun Zhang. GI-GS: global illumination decomposition on gaussian splatting for inverse rendering. In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24-28, 2025*, 2025. 2, 3

[5] Xiaoxue Chen, Junchen Liu, Hao Zhao, Guyue Zhou, and Ya-Qin Zhang. Nerrf: 3d reconstruction and view synthesis for transparent and specular objects with neural refractive-reflective fields. *arXiv preprint arXiv:2309.13039*, 2023. 3

[6] Jan-Niklas Dihlmann, Arjun Majumdar, Andreas Engelhardt, Raphael Braun, and Hendrik P.A. Lensch. Subsurface scattering for gaussian splatting. In *Advances in Neural Information Processing Systems*, pages 121765–121789. Curran Associates, Inc., 2024. 2, 3

[7] Kang Du, Zhihao Liang, Yulin Shen, and Zeyu Wang. Gs-id: Illumination decomposition on gaussian splatting via adaptive light aggregation and diffusion-guided material priors. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 26220–26229, 2025. 2, 3

[8] Chen Gao, Yipeng Wang, Changil Kim, Jia-Bin Huang, and Johannes Kopf. Planar reflection-aware neural radiance fields. In *SIGGRAPH Asia 2024 Conference Papers*, New York, NY, USA, 2024. Association for Computing Machinery. 3, 5

[9] Fangzhou Gao, Lianghao Zhang, Li Wang, Jiamin Cheng, and Jiawan Zhang. Transparent object reconstruction via implicit differentiable refraction rendering. In *SIGGRAPH Asia 2023 Conference Papers*, New York, NY, USA, 2023. Association for Computing Machinery. 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 *European Conference on Computer Vision*, pages 73–89. Springer, 2024. 2

[11] Wenhang Ge, Tao Hu, Haoyu Zhao, Shu Liu, and Ying-Cong Chen. Ref-neus: Ambiguity-reduced neural implicit surface learning for multi-view reconstruction with reflection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4251–4260, 2023. 2

[12] Yijia Guo, Yuanxi Bai, Liwen Hu, Ziyi Guo, Mianzhi Liu, Yu Cai, Tiejun Huang, and Lei Ma. Prtgs: Precomputed radiance transfer of gaussian splats for real-time high-quality relighting. In *Proceedings of the 32nd ACM International Conference on Multimedia*, page 5112–5120, New York, NY, USA, 2024. Association for Computing Machinery. 2

[13] Yuan-Chen Guo, Di Kang, Linchao Bao, Yu He, and Song-Hai Zhang. Nerfren: Neural radiance fields with reflections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18409–18418, 2022. 5

[14] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 Conference Papers*, New York, NY, USA, 2024. Association for Computing Machinery. 2, 3, 6, 7, 1, 4

[15] Letian Huang, Dongwei Ye, Jialin Dan, Chengzhi Tao, Huiwen Liu, Kun Zhou, Bo Ren, Yuanqi Li, Yanwen Guo, and Jie Guo. Transparentgs: Fast inverse rendering of transparent objects with gaussians. *ACM Trans. Graph.*, 44(4), 2025. 2, 3, 6

[16] Jeffrey Ichnowski, Yahav Avigal, Justin Kerr, and Ken Goldberg. Dex-nerf: Using a neural radiance field to grasp transparent objects. In *Proceedings of the 5th Conference on Robot Learning*, pages 526–536. PMLR, 2022. 3

[17] Yingwenqi Jiang, Jiadong Tu, Yuan Liu, Xifeng Gao, Xiaoxiao Long, Wenping Wang, and Yuexin Ma. Gaussianshader: 3d gaussian splatting with shading functions for reflective surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5322–5332, 2024. 2, 7, 3, 4

[18] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4), 2023. 1, 3, 7, 4

[19] Wooseok Kim, Taiki Fukiage, and Takeshi Oishi. Ref2-nerf: Reflection and refraction aware neural radiance field. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7196–7203, 2024. 3

[20] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. *arXiv:2304.02643*, 2023. 6

[21] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Tanks and temples: benchmarking large-scale scene reconstruction. *ACM Trans. Graph.*, 36(4), 2017. 7, 3

[22] Georgios Kouros, Minye Wu, and Tinne Tuytelaars. Rgs-dr: Reflective gaussian surfels with deferred rendering for shiny objects. *arXiv preprint arXiv:2504.18468*, 2025. 2, 3

[23] Mingwei Li, Pu Pang, Hehe Fan, Hua Huang, and Yi Yang. Tsgs: Improving gaussian splatting for transparent surface reconstruction via normal and de-lighting priors. In *Proceedings of the 33rd ACM International Conference on Multime-*

dia, page 7220–7229, New York, NY, USA, 2025. Association for Computing Machinery. 3, 6

[24] Ruofan Liang, Huiting Chen, Chunlin Li, Fan Chen, Selvakumar Panneer, and Nandita Vijaykumar. Envidr: Implicit differentiable renderer with neural environment lighting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 79–89, 2023. 2

[25] Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. Gs-ir: 3d gaussian splatting for inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21644–21653, 2024. 2

[26] 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 Trans. Graph.*, 42(4), 2023. 2

[27] Li Ma, Vasu Agrawal, Haithem Turki, Changil Kim, Chen Gao, Pedro Sander, Michael Zollhöfer, and Christian Richardt. Specnerf: Gaussian directional encoding for specular reflections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21188–21198, 2024.

[28] Alexander Mai, Dor Verbin, Falko Kuester, and Sara Fridovich-Keil. Neural microfacet fields for inverse rendering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 408–418, 2023. 2

[29] 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. *Commun. ACM*, 65(1):99–106, 2021. 3

[30] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019. 6

[31] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, and Christoph Feichtenhofer. Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*, 2024. 6

[32] Ji Shi, Xianghua Ying, Ruohao Guo, Bowei Xing, and Wenzhen Yue. Normal-nerf: Ambiguity-robust normal estimation for highly reflective scenes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(7):6869–6877, 2025. 2

[33] Yahao Shi, Yanmin Wu, Chenming Wu, Xing Liu, Chen Zhao, Haocheng Feng, Jian Zhang, Bin Zhou, Errui Ding, and Jingdong Wang. Gir: 3d gaussian inverse rendering for relightable scene factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–12, 2025. 2

[34] Jiajun Tang, Fan Fei, Zhihao Li, Xiao Tang, Shiyong Liu, Youyu Chen, Binxiao Huang, Zhenyu Chen, Xiaofei Wu, and Boxin Shi. Spectre-gs: Modeling highly specular surfaces with reflected nearby objects by tracing rays in 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16133–16142, 2025. 2, 3

[35] 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)*, pages 5491–5500, 2022. 2, 7, 3

[36] Dor Verbin, Pratul P. Srinivasan, Peter Hedman, Ben Mildenhall, Benjamin Attal, Richard Szeliski, and Jonathan T. Barron. Nerf-casting: Improved view-dependent appearance with consistent reflections. In *SIGGRAPH Asia 2024 Conference Papers*, New York, NY, USA, 2024. Association for Computing Machinery. 2, 7, 3

[37] Dongqing Wang, Tong Zhang, and Sabine Süsstrunk. Nemto: Neural environment matting for novel view and relighting synthesis of transparent objects. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 317–327, 2023. 3

[38] Fangjinhua Wang, Marie-Julie Rakotosaona, Michael Niemeyer, Richard Szeliski, Marc Pollefeys, and Federico Tombari. Unisdf: Unifying neural representations for high-fidelity 3d reconstruction of complex scenes with reflections. In *Advances in Neural Information Processing Systems*, pages 3157–3184. Curran Associates, Inc., 2024. 2

[39] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. 7, 1

[40] Tong Wu, Jia-Mu Sun, Yu-Kun Lai, Yuewen Ma, Leif Kobbelt, and Lin Gao. Deferredgs: Decoupled and editable gaussian splatting with deferred shading. *arXiv preprint arXiv:2404.09412*, 2024. 2, 3

[41] Tao Xie, Xi Chen, Zhen Xu, Yiman Xie, Yudong Jin, Yujun Shen, Sida Peng, Hujun Bao, and Xiaowei Zhou. Envgs: Modeling view-dependent appearance with environment gaussian. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5742–5751, 2025. 2, 4, 7, 1, 3

[42] Yuxuan Yao, Zixuan Zeng, Chun Gu, Xiatian Zhu, and Li Zhang. Reflective gaussian splatting. In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24-28, 2025*, 2025.

[43] Keyang Ye, Qiming Hou, and Kun Zhou. 3d gaussian splatting with deferred reflection. In *ACM SIGGRAPH 2024 Conference Papers*, New York, NY, USA, 2024. Association for Computing Machinery. 2, 4, 7, 1, 3

[44] Keyang Ye, Qiming Hou, and Kun Zhou. Progressive radiance distillation for inverse rendering with gaussian splatting. *arXiv preprint arXiv:2408.07595*, 2024.

[45] Kai Ye, Chong Gao, Guanbin Li, Wenzheng Chen, and Baoquan Chen. Geosplatting: Towards geometry guided gaussian splatting for physically-based inverse rendering. In

Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 28991–29000, 2025. 2, 3

[46] Yifan Zhan, Shohei Nobuhara, Ko Nishino, and Yinqiang Zheng. Nerfrac: Neural radiance fields through refractive surface. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 18402–18412, 2023. 3

[47] Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen Koltun. Nerf++: Analyzing and improving neural radiance fields. *arXiv preprint arXiv:2010.07492*, 2020. 2

[48] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 7

[49] Rui Zhang, Tianyue Luo, Weidong Yang, Ben Fei, Jingyi Xu, Qingyuan Zhou, Keyi Liu, and Ying He. Refgaussian: Disentangling reflections from 3d gaussian splatting for realistic rendering. *arXiv preprint arXiv:2406.05852*, 2024. 3, 5

[50] Youjia Zhang, Anpei Chen, Yumin Wan, Zikai Song, Junqing Yu, Yawei Luo, and Wei Yang. Ref-gs: Directional factorization for 2d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 26483–26492, 2025. 2, 3, 4, 5, 7, 1

[51] Zuoliang Zhu, Beibei Wang, and Jian Yang. Gs-ror2: Bidirectional-guided 3dgs and sdf for reflective object relighting and reconstruction. *ACM Trans. Graph.*, 45(1), 2025. 2, 3