

Confidence-Based Mesh Extraction from 3D Gaussians

Lukas Radl^{1*,†}, Felix Windisch^{1*}, Andreas Kurz^{1*}, Thomas Köhler¹,
Michael Steiner¹, and Markus Steinberger^{1,2}

¹ Graz University of Technology, Austria

² Huawei Technologies, Austria

<https://r4dl.github.io/CoMe/>

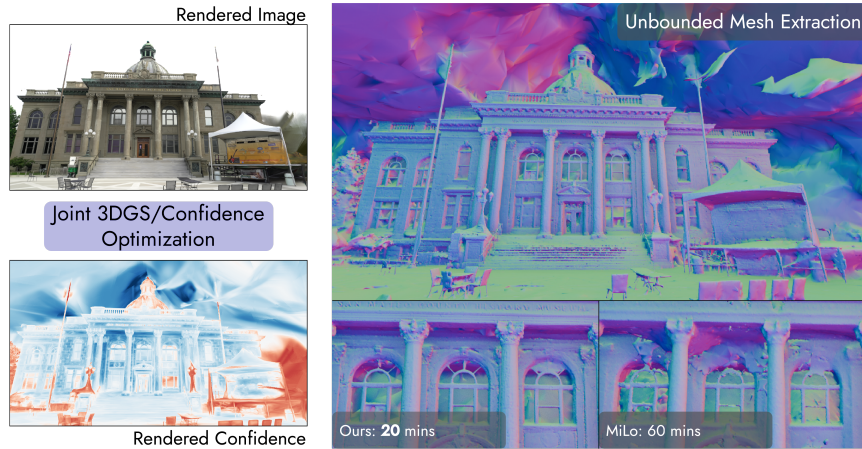


Fig. 1: Teaser: We propose a novel, confidence-based method to extract meshes from 3D Gaussians. Each Gaussian is equipped with additional confidence values that balance photometric and geometric losses in a self-supervised manner. Compared to related work, our final meshes exhibit finer details and fewer artifacts.

Abstract. Recently, 3D Gaussian Splatting (3DGS) greatly accelerated mesh extraction from posed images due to its explicit representation and fast software rasterization. While the addition of geometric losses and other priors has improved the accuracy of extracted surfaces, mesh extraction remains difficult in scenes with abundant view-dependent effects. To resolve the resulting ambiguities, prior works rely on multi-view techniques, iterative mesh extraction, or large pre-trained models, sacrificing the inherent efficiency of 3DGS. In this work, we present a simple and efficient alternative by introducing a self-supervised confidence framework to 3DGS: within this framework, learnable confidence values dynamically balance photometric and geometric supervision. Extending our confidence-driven formulation, we introduce losses which penalize per-primitive color and normal variance and demonstrate their benefits to

* denotes equal contribution.

† Correspondence lukas.radl@tugraz.at

surface extraction. Finally, we complement the above with an improved appearance model, by decoupling the individual terms of the D-SSIM loss. Our final approach delivers state-of-the-art results for unbounded meshes while remaining highly efficient.

Keywords: Mesh Extraction · 3D Gaussian Splatting · Surface Reconstruction

1 Introduction

Accurate reconstruction of high-quality surfaces from multi-view captures is a long-standing problem in Computer Vision. Recent progress in this field has mostly been driven by Novel View Synthesis methods such as Neural Radiance Fields (NeRF) [41] and, more recently, 3D Gaussian Splatting (3DGS) [25]. Most prior art optimizes a scene representation driven by photometric losses, with additional geometric priors during optimization; subsequently, a tailored mesh extraction algorithm is applied post-hoc to extract the final surface.

The improvements in speed and accuracy delivered by 3DGS have led to a recent surge in surface reconstruction methods [4, 17, 19, 69]. The main difficulty to overcome in this specific setting is the inherent *coupling* of geometry and appearance within 3DGS (*i.e.*, geometric primitives are endowed with appearance features): This implies that the appearance may be modified directly or indirectly by updates to the underlying geometry. In challenging scenarios with high-frequency view-dependent appearance, updating the geometry may, in fact, represent the only viable way to truly minimize the photometric error; particularly since Spherical Harmonics (SH), as currently employed, are limited to low-frequency components [10]. For instance, certain high-frequency view-dependent effects are often represented as highly opaque Gaussians behind a semi-transparent surface, which are only visible from certain viewpoints.

Recent work has already identified these challenges and proposed various solutions. Some methods [4, 31, 74] rely on multi-view geometric and photometric constraints, which effectively constrain the geometry of Gaussians. Other works [5, 30, 36] rely on large, pre-trained monocular models for either normals or depth, enabling stronger supervision compared to the commonly used geometric constraints. Another, more recent line of work [16] proposes mesh-aware optimization strategies, enforcing bidirectional consistency between Gaussians and the extracted mesh during training. However, all of these approaches add inherent overhead that negatively impacts computational efficiency.

In this work, we pose the question: How far can we push surface extraction without relying on any of the previously introduced techniques? The answer comes in the form of a self-supervised confidence framework inspired by recent work on Bayesian Networks [24] and feed-forward models [58], where learnable per-primitive confidence values dynamically balance photometric and geometric losses. Further, we introduce confidence-steered densification, which mitigates the over-densification problem, where small Gaussians in difficult-to-reconstruct regions are cloned repeatedly. Within this self-supervised framework, we identify

blending variance as a primary indicator of geometric uncertainty. To address this, we introduce color and normal variance losses, further boosting the resulting mesh quality. To support the optimization of a consistent 3D scene representation, we investigate commonly used appearance embeddings that compensate for varying illumination and propose a novel D-SSIM decoupling mechanism that improves mesh quality driven by a more consistent 3D scene representation.

Our *contributions* can be summarized as follows:

- We introduce a self-supervised confidence framework to 3D Gaussian Splatting, along with densification adaptations, improving surface reconstruction
- We demonstrate how reducing variance in color- and normal blending resolves visual and geometric ambiguities, thereby improving the accuracy of surfaces
- Motivated by an analysis of photometric losses in 3DGS, we propose a novel decoupled appearance embedding, improving reconstruction in challenging real-world datasets

Together, these purely self-supervised contributions allow us to achieve state-of-the-art performance for unbounded mesh extraction, while strictly avoiding the computational overhead of prior-dependent methods to maintain a highly competitive runtime (*cf.* Fig. 1).

2 Related Work

Radiance Field Methods. Kerbl *et al.* revolutionized novel view synthesis through 3D Gaussian Splatting (3DGS) [25], which usurped previous implicit and hybrid methods [2, 41, 44]. Recent work has extended this approach by mitigating various rendering artifacts [18, 48, 54, 57, 68], improved densification strategies [27, 39, 50, 65], and enhanced view-dependent appearance [10, 35, 64]. Other methods focused on obtaining more lightweight representations [12, 39, 47], or extending 3DGS to support other camera models [43, 62]. 3DGS has also inspired a wave of follow-up works, focused on improving the underlying 3D representation itself [15, 37, 55].

Uncertainty in Radiance Fields. The optimization of radiance fields given imperfect, real-world data is inherently underconstrained. To quantify the resulting uncertainty, existing works leverage Bayesian frameworks: Variational inference has been used to learn probability distributions over radiance fields [32, 52, 53]. Alternatively, Bayes’ Rays [14] estimates the uncertainty after optimization via Laplace approximation, while other works explore density-aware or self-ensembling techniques [56, 75]. While these methods have proven effective for active reconstruction — where agents are guided towards high-uncertainty regions [22, 23, 63] — applications to surface reconstruction remain underexplored: Both UA-GS [36] and VCR-GauS [5] use confidence to balance the influence of inconsistent, independently predicted pseudonormals. In contrast, our self-supervised confidence loss balances the influence of photometric and geometric losses during optimization, without relying on large, pre-trained networks.

Compensating for Varying Appearance in Radiance Fields. Compensating for camera-specific image sensor processing (ISP) is central to all 3D reconstruction tasks from images. NeRF in the Wild [40] repurposes generative latent optimization [3] to compensate for variations in capture time; this approach was adapted by follow-up works extending 3DGS to in-the-wild data [7, 29, 71]. For less extreme variations, lighter models may also be employed. Both VastGaussian [34] and Hierarchical 3DGS [26] compensate for exposure variations for large-scale data; interestingly, the appearance model from VastGaussian has been widely adopted in surface reconstruction methods [16, 49, 69]. Wang *et al.* extend learnable affine mappings with bilateral grids [59] to compensate for varying illumination. ADOP [51] models a camera compensation module as a learnable post-process, whereas NExF [45] elevates exposure prediction to 3D. Concurrently, PPISP [9] proposed a physically-motivated camera compensation module with a custom controller module, allowing for inference of ISP properties during inference. In contrast, we analyze and improve the commonly used appearance embeddings for mesh extraction, which we demonstrate to vastly improve surface reconstruction metrics.

Surface Extraction from 3D Gaussians. Extracting meshes from 3DGS is usually achieved by incorporating geometric priors during optimization and extracting a mesh as a post-hoc process. Such geometric priors may come in the form of surface-alignment [17], multi-view constraints [4, 31], monocular priors [5, 30, 31], or the use of Gaussian surfels directly [8, 19]. These baseline approaches have been extended with rasterized depth [70] or quadric surfaces [74]. Some works establish an opacity field in 3D to use marching tetrahedra for unbounded meshes [49, 69]. Another recent line of work establishes bidirectional consistency between meshes and Gaussians [16]. Alternatively, some works jointly optimize a neural SDF for mesh extraction [33, 38, 67, 72], while others leverage sparse voxel rasterization [30, 55]. Another class of techniques explores stochastic frameworks for surface reconstruction [6, 21, 42]. Building on these prior works, we demonstrate state-of-the-art unbounded mesh extraction.

3 Method

3.1 Preliminaries

3DGS [25] represents scenes using N anisotropic 3D Gaussians. Each primitive is characterized by a position $\boldsymbol{\mu}_i \in \mathbb{R}^3$, orientation $\mathbf{R}_i \in SO(3)$, diagonal scaling matrix $\mathbf{S}_i \in \text{diag}(\mathbb{R}_+^3)$, opacity $o_i \in [0, 1]$, and spherical harmonics (SH) coefficients $\boldsymbol{\theta}_i$. The covariance matrix is given by $\boldsymbol{\Sigma}_i = \mathbf{R}_i \mathbf{S}_i \mathbf{S}_i^T \mathbf{R}_i^T$.

To render a pixel, Gaussians are sorted front-to-back along the corresponding ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$ and accumulated via alpha-blending:

$$\hat{\mathbf{I}}(\mathbf{r}) = \sum_{i=0}^{N-1} w_i(\mathbf{r}) \text{sh}(\boldsymbol{\theta}_i, \mathbf{d}), \quad (1)$$

where $w_i(\mathbf{r}) = \alpha_i(\mathbf{r}) \prod_{j=1}^{i-1} (1 - \alpha_j(\mathbf{r}))$. Following recent works [18, 54, 69], we evaluate the opacity $\alpha_i(\mathbf{r})$ directly in 3D at the point $\mathbf{x}_i^*$ of maximum contribution along the ray:

$$\alpha_i(\mathbf{r}) = o_i \exp \left(-\frac{1}{2} (\mathbf{x}_i^* - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x}_i^* - \boldsymbol{\mu}_i) \right), \quad (2)$$

where $\mathbf{x}_i^*$ is analytically found by minimizing the Mahalanobis distance between the ray and the Gaussian center:

$$\mathbf{x}_i^* = \mathbf{o} + \frac{\mathbf{d}^T \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{\mu}_i - \mathbf{o})}{\mathbf{d}^T \boldsymbol{\Sigma}_i^{-1} \mathbf{d}} \mathbf{d}. \quad (3)$$

The parameters are optimized to minimize a photometric reconstruction loss, which is a combination of $\mathcal{L}_1$ and D-SSIM [60] terms:

$$\mathcal{L}_{\text{rgb}} = (1 - \lambda_{\text{rgb}}) \mathcal{L}_1(\mathbf{I}, \hat{\mathbf{I}}) + \lambda_{\text{rgb}} \mathcal{L}_{\text{D-SSIM}}(\mathbf{I}, \hat{\mathbf{I}}), \quad (4)$$

with $\lambda_{\text{rgb}} = 0.2$. Here, $\mathcal{L}_{\text{D-SSIM}}$ is the dissimilarity, which can be re-written as

$$\mathcal{L}_{\text{D-SSIM}}(\mathbf{I}, \hat{\mathbf{I}}) = 1 - l(\mathbf{I}, \hat{\mathbf{I}}) \cdot c(\mathbf{I}, \hat{\mathbf{I}}) \cdot s(\mathbf{I}, \hat{\mathbf{I}}), \quad (5)$$

with l, c, s the luminance, contrast, and structure terms, respectively. We refer the interested reader to [46] for an in-depth analysis of SSIM.

Appearance Modeling. The appearance modeling approach of VastGaussian [34] consists of per-image latents $\boldsymbol{\rho}_i \in \mathbb{R}^{64}$, which are concatenated to the downsampled version of $\hat{\mathbf{I}}$. A convolutional neural network F_Θ then predicts a full-resolution, per-channel corrective mapping

$$\mathbf{M}_i = F_\Theta \left([\text{ds}_{32}(\hat{\mathbf{I}}), \boldsymbol{\rho}_i] \right), \quad (6)$$

with $\mathbf{M}_i \in \mathbb{R}^{H \times W \times 3}$, $\text{ds}_{32}(\cdot)$ denoting downsampling by a factor of 32 and $[\cdot, \cdot]$ denoting channel-wise concatenation, where $\boldsymbol{\rho}_i$ is first spatially broadcast. The rendered image is then transformed according to

$$\hat{\mathbf{I}}^{\text{app}} = \hat{\mathbf{I}} \odot \sigma(\mathbf{M}_i), \quad (7)$$

with $\sigma(\cdot)$ denoting the sigmoid activation.³ Importantly, the modified image $\hat{\mathbf{I}}^{\text{app}}$ is only used for the $\mathcal{L}_1$ -loss in Eq. (4).

3.2 Improved Decoupled Appearance

Compensating for the camera-specific image sensor processing is central to obtaining a coherent 3D scene representation. Previous mesh extraction pipelines [16, 49, 69] have addressed this by adopting the decoupled appearance model from

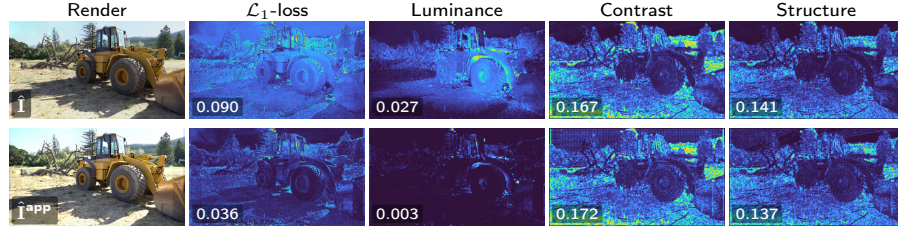


Fig. 2: Analysis of the Photometric Loss Components: The luminance term of D-SSIM is highly dependent on the illumination; in this example, the luminance error is reduced by a factor of $9\times$ by considering the appearance embedding.

VastGaussian [34], or by directly learning per-image transformation during training [4]. However, both of the aforementioned approaches only consider the modified image for the $\mathcal{L}_1$ -loss and use the original rendered image for $\mathcal{L}_{\text{D-SSIM}}$, arguing that SSIM is inherently more structural.

Upon revisiting Eq. (5), we find that the luminance term in particular is not structural, and dominated by variations in illumination; see Fig. 2 for a visualization. Therefore, we use $\hat{\mathbf{I}}^{\text{app}}$ for the luminance term, effectively modifying Eq. (5) to

$$\mathcal{L}_{\text{D-SSIM}}^{\text{dec}} = 1 - l(\mathbf{I}, \hat{\mathbf{I}}^{\text{app}}) \cdot c(\mathbf{I}, \hat{\mathbf{I}}) \cdot s(\mathbf{I}, \hat{\mathbf{I}}). \quad (8)$$

As we will demonstrate, using the modified image *only in the luminance-term* delivers superior results; intuitively, as the contrast and structure terms are illumination-invariant, using the modified image for all SSIM terms may lead to the appearance model compensating for incorrect renderings. Moreover, when considering the $\hat{\mathbf{I}}^{\text{app}}$ for these terms, artifacts due to CNN upsampling may pollute gradient flow.

We additionally improved the original VastGaussian implementation [34], unlocking a minor performance boost in itself; cf. Supplementary S.1.1 for details.

3.3 Confidence-Aware Gaussian Splatting

Previous approaches for surface extraction struggle with balancing photometric and geometric losses; in regions with high-frequency view-dependent appearance, photometric losses dominate, and large gradient magnitudes lead to over-densification, where small Gaussians are repeatedly cloned. In these high-primitive density regions, 3DGS is particularly prone to faking view-dependent appearance with geometry.

To compensate for this, we take inspiration from recent feed-forward reconstruction models [58] and weigh our photometric loss by a learned confidence value $\hat{C} \in \mathbb{R}_+^{H \times W}$:

$$\mathcal{L}_{\text{conf}} = \mathcal{L}_{\text{rgb}} \cdot \hat{C} - \beta \cdot \log \hat{C}, \quad (9)$$

³ The use of the sigmoid activation is following the re-implementation from GOF [69]; note that the source code for VastGaussian [34] was never released.

where the hyperparameter $\beta \in \mathbb{R}_+$ defines the exact photometric error value at which $\mathcal{L}_{\text{conf}} = \mathcal{L}_{\text{rgb}}$. When $\hat{C} = 1$, the loss reduces to $\mathcal{L}_{\text{rgb}}$. Intuitively, this allows for trading photometric loss for uncertainty when $\hat{C} < 1$. Importantly, the geometric losses remain untouched by this formulation; thus, the rendered confidence can be used to balance both terms.

Eq. (9) also directly influences parameter updates: As $\partial \mathcal{L}_{\text{conf}} / \partial \mathcal{L}_{\text{rgb}} = \hat{C}$, the gradients for each primitive are scaled by the rendered confidence. Hence, updates from ambiguous regions are downweighted intrinsically by using this loss formulation, balancing larger image errors. Conversely, updates for confident regions are scaled by a factor larger than 1; in practice, this is not an issue as gradients for regions of high confidence are typically small.



Fig. 3: Confidence-Driven Gaussian Splatting: We show rendered images from our trained method, accompanied by the rendered confidence maps $\hat{C}$. The confidence maps effectively isolate reflective surfaces, thin foliage, or rarely observed areas (such as the roof for Barn). Importantly, the low-confidence regions are still well-reconstructed.

Rendering Confidence. We endow each primitive with a learnable scalar value γ_i , initialized to 0, from which we derive the confidence $\tilde{\gamma}_i$ using

$$\tilde{\gamma}_i = \exp(\gamma_i), \quad (10)$$

ensuring an initial value of 1. We derive the final rendered confidence map $\hat{C}$ via alpha-blending (*cf.* Fig. 3 for a visualization):

$$\hat{C}(\mathbf{r}) = \sum_{i=0}^{N-1} w_i(\mathbf{r}) \tilde{\gamma}_i. \quad (11)$$

Confidence-Aware Densification. Clearly, the confidence-based balancing from Eq. (9) changes the magnitude of the loss, and may even lead to negative values for $\mathcal{L}_{\text{conf}}$, when confidence is large and $\mathcal{L}_{\text{rgb}}$ approaches 0. Because the densification scheme is built on the magnitudes of per-Gaussian gradients [65, 69], it needs to be adapted.

In particular, we want to prevent over-densification for difficult-to-reconstruct regions; these are usually characterized by low $\hat{C}$. Within gradient densification, primitives are usually split or cloned when the accumulated positional gradients exceed a threshold $\tau_{\text{grad}} \in \mathbb{R}_+$. To discourage unconfident Gaussians from splitting, we divide this value by the per-primitive confidence:

$$\bar{\tau}_{\text{grad}} = \frac{\tau_{\text{grad}}}{\min(\tilde{\gamma}_i, 1)}. \quad (12)$$

Importantly, we clamp the denominator by 1 to ensure that confident Gaussians are not over-densified (*i.e.* $\bar{\tau}_{\text{grad}} \geq \tau_{\text{grad}}$). In practice, we find that these modifications slightly reduce the overall primitive count when β is set to a reasonable value.



Fig. 4: Effect of our proposed Variance Losses: Rendering color and normals only for the first-hit Gaussians shows how the variance losses align individual Gaussians better to the true object surface, in both color and orientation.

3.4 Variance Reducing Losses for Blending

While confidence-steered optimization already improves mesh quality, we still observe cases where view-dependent appearance is represented with geometry; most frequently, this is characterized by highly opaque primitives just behind surfaces, which are blocked in different amounts by semitransparent primitives depending on the viewing angle. This is in part enabled by the use of alpha-blending (see Eq. (1)): Since only the blended final color is used to compute the loss, individual primitives have no incentive to correctly model radiance.

Previous work [73] replaces the loss function of NeRF to apply the photometric loss to the individual color samples along the ray instead of the blended result, which effectively constrains such extreme cases. We find that naïvely applying this method to 3DGS significantly deteriorates reconstruction quality, because the vital SSIM loss cannot be applied per-primitive. Instead, we introduce an additional weighted per-primitive $\mathcal{L}_2$ -loss with the ground truth pixel color $\mathbf{I}$ to complement the standard 3DGS training loss:

$$\mathcal{L}_{\text{color-var}} = \sum_{i=0}^{N-1} w_i(\mathbf{r}) \|\text{sh}(\boldsymbol{\theta}_i, \mathbf{d}) - \mathbf{I}\|_2^2. \quad (13)$$

Under the assumption that $\mathbf{I}$ represents the rendered color, this is equivalent to minimizing the blended color variance along the ray. While this formulation of uncertainty has already been explored in radiance fields [11], the key difference is our employment of this formulation as a loss term, within the specific context of surface reconstruction.

The per-pixel normals used in the depth-normal consistency loss [19, 49, 69] are also alpha-blended, resulting in poor alignment of individual Gaussians to the surface that can be remedied with the introduction of a normal variance loss

$$\mathcal{L}_{\text{normal-var}} = \sum_{i=0}^{N-1} w_i(\mathbf{r}) \|\mathbf{n}_i - \mathbf{N}\|_2^2 = 1 - \|\mathbf{N}\|_2^2, \quad (14)$$

with blended pixel normal $\mathbf{N} = \sum_{i=0}^{N-1} w_i(\mathbf{r}) \mathbf{n}_i$.

The resulting improved color and normal consistency is demonstrated in Fig. 4. We note that the commonly used *depth distortion* loss [1] can also be interpreted as a variance loss on the blended depth.

4 Results and Evaluation

Implementation Details. Due to its relative speed, we build our solution on top of the SOF codebase [49]. For mesh extraction, we also use Marching Tetrahedra with binary search [69] to extract unbounded meshes, using the faster implementation from Radl *et al.* [49]. Our final loss function is:

$$\mathcal{L} = \mathcal{L}_{\text{conf}} + \mathcal{L}_{\text{geom}} + \lambda_{\text{color-var}} \mathcal{L}_{\text{color-var}} + \lambda_{\text{normal-var}} \mathcal{L}_{\text{normal-var}}, \quad (15)$$

where $\mathcal{L}_{\text{geom}}$ denotes the geometry losses as used in SOF (see [49] for details). We set $\lambda_{\text{color-var}} = 5 \times 10^{-1}$ and $\lambda_{\text{normal-var}} = 5 \times 10^{-3}$ across all scenes.

For the confidence values γ_i , we use a learning rate of 2.5×10^{-4} , and set $\beta = 0.075$ for all scenes. We also report results for $\beta = 0.05$, which represents a lighter variant of our model. We enable $\mathcal{L}_{\text{conf}}$ from iteration 500, right at the start of densification; empirically, we have not observed improved training stability when enabling confidence learning later. Further implementation details are included in Supplemental S.1.

Datasets and Comparisons. To benchmark unbounded mesh extraction, we primarily rely on the Tanks & Temples dataset [28]. For bounded scenes, we include a more rigorous baseline for real-world surface extraction by conducting extensive evaluations on the high-fidelity ScanNet++ dataset [66], moving beyond traditional small-scale object benchmarks [20]. Finally, we evaluate novel view synthesis and mesh extraction on the Mip-NeRF 360 dataset [1].

Our main comparison points are GOF [69], SOF [49], RaDe-GS [70] and MILO [16] for unbounded meshes; we also include QGS [74] and PGSR [4] as representative bounded, multi-view baselines. We reproduced all results using the latest available codebase. To ensure a fair and reproducible comparison,

we standardized the evaluation protocol across all methods. Specifically, we adjusted the training and evaluation settings (*e.g.* matching image resolutions and train/test splits) for several baselines (see the supplementary material for details). Consequently, the reported results do not always reflect the numbers reported by the authors. Importantly, we excluded methods relying on pre-trained monocular models [5, 30, 31] from this comparison, and refer the reader to the supplementary material for a dedicated discussion.

4.1 Surface Reconstruction

Table 1: Full Geometry Reconstruction results for the Tanks & Temples dataset [28]. We evaluate the F1-score (higher is better) and report runtimes, measured on an RTX 4090. Overall, our method achieves the best reconstruction quality of all tested methods (without leveraging multi-view constraints), while remaining efficient.

	Barn	Caterpillar	Courthouse	Ignatius	Meetingroom	Truck	Average	Runtime
PGSR	0.548	0.437	0.238	0.728	0.367	0.658	0.496	28min
QGS	0.536	0.374	0.183	0.733	0.374	0.645	0.474	41min
GOF	0.484	0.402	0.288	0.674	0.275	0.596	0.453	40min
SOF	0.535	0.408	0.297	0.736	0.309	0.558	0.474	17min
RaDe-GS	0.529	0.464	0.316	0.536	0.343	0.576	0.461	8min
MILo	0.541	0.389	0.322	0.757	0.281	0.617	0.485	60min
Ours	0.534	0.472	0.333	0.782	0.372	0.634	0.521	18min
Ours ($\beta = 0.05$)	0.519	0.451	0.315	0.752	0.358	0.637	0.505	16min

Unbounded Meshes. Our main quantitative results concern mesh reconstruction for the Tanks & Temples dataset [28], which are presented in Tab. 1. We report the F1-score for each scene, the overall average, as well as the average optimization time across methods. Our approach delivers superior overall accuracy, achieving a leading average F1-score of **0.521**, outperforming all evaluated baselines with a significant margin. Specifically, our method demonstrates exceptional robustness in challenging scenarios: our decoupled appearance module effectively resolves the varying illumination present in Caterpillar and Ignatius, obtaining the best scores for these scenes. Our confidence framework and variance losses result in robustness for particularly challenging scenes, such as Courthouse (where we have > 1100 input images). For Meetingroom and Truck, our method attains the highest scores out of all unbounded methods; here, bounded baselines [4, 74] rely on heavy multi-view constraints to regularize underconstrained regions such as the textureless floor. In summary, our method attains state-of-the-art performance, extracts detailed unbounded meshes, and demonstrates highly competitive runtimes (optimization in < 20 minutes).

Bounded Meshes. We additionally evaluated bounded mesh extraction on a 6-scene subset of the ScanNet++ dataset [66] (see the supplementary material for selection details). We present our quantitative results in Tab. 2, where we report

Table 2: Full Geometry Reconstruction results for our ScanNet++-v2 subset [66], where we report the F1-score. Our method attains the highest scores across all scenes, underscoring our real-world applicability.

	5a269ba6fe	08bbbdcc3d	39f36da05b	c263dfbf0	ef18cf0708	fb564c935d	Average
PGSR	0.668	0.666	0.647	0.630	0.542	0.632	0.631
QGS	0.646	0.622	0.571	0.522	0.504	0.572	0.573
GOF	0.620	0.670	0.611	0.672	0.504	0.661	0.623
SOF	0.599	0.681	0.616	0.675	0.497	0.621	0.615
RaDe-GS	0.642	0.628	0.618	0.600	0.557	0.634	0.613
MiLo	0.622	0.675	0.621	0.626	0.542	0.660	0.624
Ours	0.670	0.729	0.657	0.715	0.551	0.684	0.668
Ours ($\beta = 0.05$)	0.678	0.715	0.656	0.716	0.555	0.678	0.667

the per-scene and average F1-score. Our full method outperforms other works by a significant margin and attains the best performance for all scenes. Crucially, these results demonstrate the applicability of our method in unconstrained, real-world scenarios: ScanNet++ is known to suffer from exposure variation, inconsistent lighting, and occasional blur, which our decoupled appearance module faithfully compensates for.

Qualitative Comparison. Fig. 5 visualizes qualitative mesh comparisons against baseline approaches. When compared to bounded surface extraction methods [4, 74], our meshes are significantly more complete. These baselines rely on multi-view constraints that tend to aggressively over-smooth or entirely cull fine, complex geometry (*e.g.*, thin structures and foliage).

Conversely, the state-of-the-art unbounded method MiLo [16] suffers from similar artifacts. This limitation can be directly attributed to its mesh-aware bidirectional consistency loss, which is most easily minimized by generating overly smooth geometry. Consequently, the resulting meshes lack intricate surface details. Finally, compared to our direct baseline, SOF [49], our extracted meshes exhibit finer details while successfully avoiding the majority of visual artifacts. These enhancements are driven by our proposed confidence framework, which provides a dedicated mechanism to explicitly resolve severe photometric ambiguities. Finally, we demonstrate mesh rendering of challenging unbounded scenes from Mip-NeRF 360 [1] in Fig. 6.

4.2 Ablation Studies

Individual Components. We investigate the impact of our contributions in Tab. 3, where we considered the Tanks & Temples as well as ScanNet++ dataset. Adding our improved appearance module leads to better mesh quality across both datasets; on its own, it outperforms MiLo [16], the current state-of-the-art method for unbounded meshes. Adding $\mathcal{L}_{\text{conf}}$ leads to further improvements, with a significant **0.03** gain for ScanNet++. This substantial boost can be attributed to the inherent ambiguities of indoor scenes, such as untextured walls,

Table 3: Ablation Study for our Method: We analyze the effects of our individual contribution on both the Tanks & Temples [28] and the ScanNet++ dataset [66]. As we can see, all our proposed components result in notable improvements.

Method	Tanks & Temples			ScanNet++		
	Precision [↑]	Recall [↑]	F1-Score [↑]	Precision [↑]	Recall [↑]	F1-Score [↑]
SOF [49] (baseline)	0.542	0.438	0.474	0.565	0.677	0.615
+ Improved Appearance	0.557	0.459	0.493	0.577	0.685	0.625
+ $\mathcal{L}_{\text{conf}}$	0.568	0.479	0.509	0.605	0.715	0.655
+ $\mathcal{L}_{\text{color-var}}$	0.571	0.491	0.519	0.606	0.722	0.658
+ $\mathcal{L}_{\text{normal-var}}$	0.580	0.491	0.521	0.613	0.734	0.668

Table 4: Ablation Study on the Decoupled Appearance Module: We compare appearance embeddings via surface reconstruction on the Tanks & Temples dataset [28]. Our proposed decoupled configuration yields the strongest results.

Method	Baselines (Standard)			Method	Ours (Ablations)		
	Precision [↑]	Recall [↑]	F1-Score [↑]		Precision [↑]	Recall [↑]	F1-Score [↑]
VastG. [34]	0.541	0.440	0.475	VastG. (Imp.)	0.557	0.445	0.484
PGSR [4]	0.545	0.441	0.478	w/o $l(\cdot)$	0.559	0.441	0.483
H3DGS [26]	0.553	0.448	0.484	w/o SSIM Dec.	0.559	0.453	0.490
PPISP [9]	0.556	0.456	0.489	Ours	0.557	0.459	0.493

inconsistent lighting, and occasional blur. The color variance loss further improves performance by constraining extreme scenarios, in which view-dependent appearance is represented with geometry; naturally, this yields the strongest gains on Tanks & Temples, where such high-frequency effects are abundant.

Conversely, $\mathcal{L}_{\text{normal-var}}$ leads to notable improvements for both datasets, with a stronger impact for ScanNet++. This perfectly aligns with the indoor nature of the dataset, which is heavily dominated by planar structures (*e.g.*, walls, tables); in such environments, enforcing consensus between per-primitive normals acts as a highly effective regularizer. Overall, our techniques lead to F1-score improvements of **0.047/0.053**.

Decoupled Appearance. We evaluate multiple variants of our decoupled appearance module in Tab. 4, validating our SSIM-decoupled design; here, we disabled all other components to isolate the effects of the appearance embeddings. As baselines, we consider the appearance embeddings as used in VastGaussian [34], PGSR [4], Hierarchical-3DGS (H3DGS) [26], as well as the concurrent work PPISP [9]. Our architectural enhancements to the VastGaussian baseline (VastG. (Imp.)) already yield a notable improvement in F1-score. Disregarding the luminance term (w/o $l(\cdot)$) attains similar results compared to VastG. (Imp.), demonstrating that this loss term alone does not benefit scene reconstruction unless appearance is also explicitly compensated for. Using the $\hat{\mathbf{I}}^{\text{app}}$ for $\mathcal{L}_{\text{D-SSIM}}$ (w/o SSIM Decoupling) greatly improves surface extraction results, however, we ob-

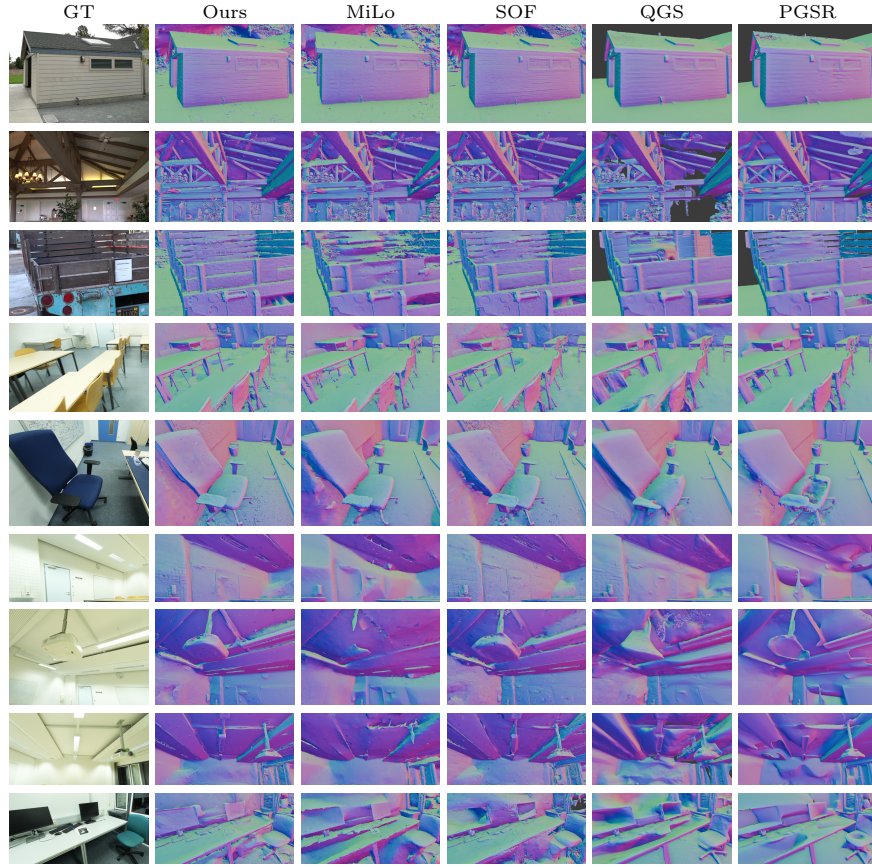


Fig. 5: Qualitative Mesh Comparison: Our approach produces more detailed meshes with fewer artifacts than other unbounded methods [16, 49]. Additionally, it achieves higher completeness and finer detail than bounded extraction works [4, 74].



Fig. 6: Unbounded Mesh Rendering: Our method reconstructs both fine details and high-quality backgrounds on challenging Mip-NeRF 360 scenes [1].

tain even stronger results when using $\hat{\mathbf{I}}$ for contrast and structure terms (**Ours**). Both of these variants slightly outperform PPISP, which targets inference; conversely, we optimize for 3D coherence and faithful mesh extraction.

Confidence. Finally, we ablate our choice of β , which is the primary hyperparameter in our implementation. To this end, we present results for the Tanks & Temples dataset on surface reconstruction, using $\beta \in [0, 0.2]$, cf. in Fig. 7. Here, we only used our improved appearance model while removing our variance losses to keep the confidence isolated. As can be seen, using $\beta = 0$ provides no incentive to the model to maximize $\hat{C}$, as $\mathcal{L}_{\text{conf}}$ reduces to $\mathcal{L}_{\text{rgb}} \cdot \hat{C}$, resulting in under-reconstruction with very few points. Using a value of $\beta = 0.2$ results in the opposite effects, where the penalty is too harsh to allow for any uncertainty; this results in a huge number of primitives, as the large $\mathcal{L}_{\text{rgb}}$ causes massive absolute gradients, negatively impacting densification. Note that β represents a specific photometric error, where even small perturbations have large effects on visual quality. Interestingly, a choice of $\beta = 0.05$ may represent a *light* variant of our model; achieving strong mesh extraction results, while keeping the number of primitives low. Ultimately, $\beta = 0.075$ achieves the highest F1-score, while using slightly fewer primitives than the baseline (w/o confidence). Our evaluation further demonstrates the robustness of β , achieving SOTA results across 4 diverse datasets without any tuning. We refer to our supplementary material for a more thorough discussion, and more visualizations.

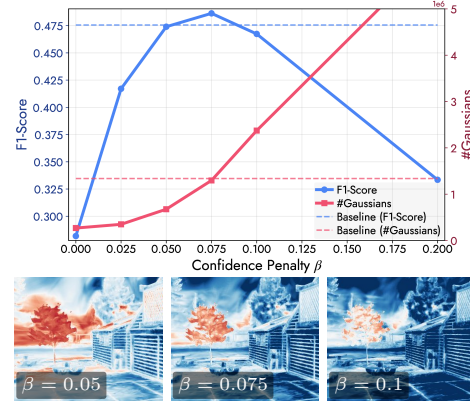


Fig. 7: β -Ablation Study: We evaluate surface reconstruction quality for varying choices of β . Our chosen value of 0.075 achieves the best performance.

5 Discussion, Limitations and Future Work

Our proposed method effectively extracts high-fidelity surfaces, but it is not without limitations. Fundamentally, our framework relies on the assumption of a moderately dense capture setting. While our confidence-driven losses successfully resolve photometric ambiguities and greatly reduce artifacts, occasional holes or cavities may still manifest in severely occluded or unobserved regions. We believe that for such scenarios, or specifically in sparse-view capture settings, integrating multi-view geometric constraints [4] or monocular priors [5, 30] effectively compensates for the lack of densely captured data.

Furthermore, our method struggles to reconstruct detailed geometry for distant background regions. Due to the insufficient capture density and lack of “close-up” captures, the 3DGS reconstruction itself is insufficiently densified in these areas. We believe that pre-trained multi-view diffusion models [13, 61],

along with specialized densification recipes, may aid in supplying additional details for such regions.

Finally, our confidence framework opens exciting avenues for broader downstream applications. We hypothesize that this formulation could readily be adapted into a heuristic-free densification framework for general 3DGS: As demonstrated in Fig. 7, the confidence penalty parameter β already acts as a hyperparameter that balances mesh quality and primitive count. We also believe that this approach may benefit other downstream 3DGS applications, such as 3D inpainting or super-resolution training.

6 Conclusion

In this work, we presented a confidence-steered framework tailored for high-fidelity surface extraction from 3D Gaussian Splatting. By introducing learnable, per-primitive confidence values alongside color and normal variance losses, our method dynamically mitigates the ambiguities caused by high-frequency view-dependent appearance. Furthermore, based on a rigorous analysis of the popular D-SSIM loss, we proposed an architecturally improved, decoupled appearance embedding that explicitly compensates for unconstrained real-world illumination changes. Ultimately, our contributions allow us to achieve state-of-the-art results for unbounded mesh extraction, all while strictly avoiding the inherent computational overhead of prior-dependent methods to maintain a highly competitive runtime (optimization in < 20 minutes). Our full implementation is publicly available at <https://github.com/r4dl/CoMe>.

Acknowledgements

This research was supported by the Austrian Science Fund FWF [10.55776/I6663]. We thank Diego Gomez and Ziyi Zhang for their help with evaluation, as well as Nissim Maruani for his help with illustrations.

References

1. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. In: CVPR (2022)
2. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Zip-NeRF: Anti-Aliased Grid-Based Neural Radiance Fields. In: ICCV (2023)
3. Bojanowski, P., Joulin, A., Lopez-Paz, D., Szlam, A.: Optimizing the Latent Space of Generative Networks. In: ICML (2017)
4. Chen, D., Li, H., Ye, W., Wang, Y., Xie, W., Zhai, S., Wang, N., Liu, H., Bao, H., Zhang, G.: PGSR: Planar-Based Gaussian Splatting for Efficient and High-Fidelity Surface Reconstruction. IEEE TVCG **31**(09) (2025)
5. Chen, H., Wei, F., Li, C., Huang, T., Wang, Y., Lee, G.H.: VCR-GauS: View Consistent Depth-Normal Regularizer for Gaussian Surface Reconstruction. In: NeurIPS (2024)

6. Chen, H., Miller, B., Gkioulekas, I.: 3D Reconstruction with Fast Dipole Sums. *ACM TOG* **43**(6) (2024)
7. Dahmani, H., Bennehar, M., Piasco, N., Roldao, L., Tsishkou, D.: SWAG: Splatting in the Wild images with Appearance-conditioned Gaussians. In: *ECCV* (2024)
8. Dai, P., Xu, J., Xie, W., Liu, X., Wang, H., Xu, W.: High-quality Surface Reconstruction using Gaussian Surfels. In: *SIGGRAPH Asia* (2024)
9. Deutsch, I., Moënné-Loccoz, N., State, G., Gojcic, Z.: PPISP: Physically-Plausible Compensation and Control of Photometric Variations in Radiance Field Reconstruction. *arXiv preprint arXiv:2601.18336* (2026)
10. Di Sario, F., Rebain, D., Verbin, D., Grangetto, M., Tagliasacchi, A.: Spherical Voronoi: Directional Appearance as a Differentiable Partition of the Sphere. *arXiv preprint arXiv:2512.14180* (2025)
11. Ewen, P., Chen, H., Isaacson, S., Wilson, J., Skinner, K.A., Vasudevan, R.: These Magic Moments: Differentiable Uncertainty Quantification of Radiance Field Models. *arXiv preprint arXiv:2503.14665* (2025)
12. Fang, G., Wang, B.: Mini-Splatting: Representing Scenes with a Constrained Number of Gaussians. In: *ECCV* (2024)
13. Fischer, T., Bulò, S.R., Yang, Y.H., Keetha, N., Porzi, L., Müller, N., Schwarz, K., Luiten, J., Pollefeys, M., Kotschieder, P.: FlowR: Flowing from Sparse to Dense 3D Reconstructions. In: *ICCV* (2025)
14. Goli, L., Reading, C., Sellán, S., Jacobson, A., Tagliasacchi, A.: Bayes’ Rays: Uncertainty Quantification in Neural Radiance Fields. In: *CVPR* (2024)
15. Govindarajan, S., Rebain, D., Yi, K.M., Tagliasacchi, A.: Radiant Foam: Real-Time Differentiable Ray Tracing. In: *ICCV* (2025)
16. Guédon, A., Gomez, D., Maruani, N., Gong, B., Drettakis, G., Ovsjanikov, M.: MILO: Mesh-In-the-Loop Gaussian Splatting for Detailed and Efficient Surface Reconstruction. *ACM TOG* **44**(6) (2025)
17. Guédon, A., Lepetit, V.: SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering. In: *CVPR* (2024)
18. Hahlbohm, F., Friederichs, F., Weyrich, T., Franke, L., Kappel, M., Castillo, S., Stamminger, M., Eisemann, M., Magnor, M.: Efficient Perspective-Correct 3D Gaussian Splatting Using Hybrid Transparency. *Comput. Graph. Forum* **44**(2) (2025)
19. Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S.: 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. In: *SIGGRAPH* (2024)
20. Jensen, R., Dahl, A., Vogiatzis, G., Tola, E., Aanaes, H.: Large Scale Multi-View Stereopsis Evaluation. In: *CVPR* (2014)
21. Jiang, K., Sivaram, V., Peng, C., Ramamoorthi, R.: Geometry Field Splatting with Gaussian Surfels. In: *CVPR* (2025)
22. Jiang, W., Lei, B., Daniilidis, K.: FisherRF: Active View Selection and Uncertainty Quantification for Radiance Fields using Fisher Information. In: *ECCV* (2024)
23. Jin, L., Zhong, X., Pan, Y., Behley, J., Stachniss, C., Popović, M.: Activegs: Active scene reconstruction using gaussian splatting. *IEEE Robotics and Automation Letters* (2025)
24. Kendall, A., Gal, Y.: What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? In: *NeurIPS* (2017)
25. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM TOG* **42**(4) (2023)
26. Kerbl, B., Meuleman, A., Kopanas, G., Wimmer, M., Lanvin, A., Drettakis, G.: A Hierarchical 3D Gaussian Representation for Real-Time Rendering of Very Large Datasets. *ACM TOG* **43**(4) (2024)

27. Kheradmand, S., Rebain, D., Sharma, G., Sun, W., Tseng, Y.C., Isack, H., Kar, A., Tagliasacchi, A., Yi, K.M.: 3D Gaussian Splatting as Markov Chain Monte Carlo. In: NeurIPS (2024)
28. Knapitsch, A., Park, J., Zhou, Q.Y., Koltun, V.: Tanks and Temples: Benchmarking Large-Scale Scene Reconstruction. ACM TOG **36**(4) (2017)
29. Kulhanek, J., Peng, S., Kukelova, Z., Pollefeys, M., Sattler, T.: WildGaussians: 3D Gaussian Splatting in the Wild. In: NeurIPS (2024)
30. Li, J., Zhang, J., Zhang, Y., Bai, X., Zheng, J., Yu, X., Gu, L.: GeoSVR: Taming Sparse Voxels for Geometrically Accurate Surface Reconstruction. In: NeurIPS (2025)
31. Li, Q., Feng, H., Gong, X., Liu, Y.S.: VA-GS: Enhancing the Geometric Representation of Gaussian Splatting via View Alignment. In: NeurIPS (2025)
32. Li, R., Cheung, Y.m.: Variational Multi-scale Representation for Estimating Uncertainty in 3D Gaussian Splatting. In: NeurIPS (2024)
33. Li, S., Liu, Y.S., Han, Z.: GaussianUDF: Inferring Unsigned Distance Functions through 3D Gaussian Splatting. In: CVPR (2025)
34. Lin, J., Li, Z., Tang, X., Liu, J., Liu, S., Liu, J., Lu, Y., Wu, X., Xu, S., Yan, Y., Yang, W.: VastGaussian: Vast 3D Gaussians for Large Scene Reconstruction. In: CVPR (2024)
35. Liu, R., Sun, D., Chen, M., Wang, Y., Feng, A.: Deformable Beta Splatting. In: SIGGRAPH (2025)
36. Liu, S., Wu, J., Wu, W., Chu, L., Liu, X.: Uncertainty-Aware Gaussian Splatting with View-Dependent Regularization for High-Fidelity 3D Reconstruction. In: Eurographics Symposium on Rendering (2025)
37. von Lütow, N., Nießner, M.: LinPrim: Linear Primitives for Differentiable Volumetric Rendering. In: NeurIPS (2025)
38. Lyu, X., Sun, Y.T., Huang, Y.H., Wu, X., Yang, Z., Chen, Y., Pang, J., Qi, X.: 3DGSr: Implicit Surface Reconstruction with 3D Gaussian Splatting. ACM TOG **43**(6) (Nov 2024)
39. Mallick, S., Goel, R., Kerbl, B., Vicente Carrasco, F., Steinberger, M., De La Torre, F.: Taming 3DGS: High-Quality Radiance Fields with Limited Resources. In: SIGGRAPH Asia (2024)
40. Martin-Brualla, R., Radwan, N., Sajjadi, M.S., Barron, J.T., Dosovitskiy, A., Duckworth, D.: NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. In: CVPR (2021)
41. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In: ECCV (2020)
42. Miller, B., Chen, H., Lai, A., Gkioulekas, I.: Objects as volumes: A stochastic geometry view of opaque solids. In: CVPR (2024)
43. Moënné-Loccoz, N., Mirzaei, A., Perel, O., de Lutio, R., Esturo, J.M., State, G., Fidler, S., Sharp, N., Gojcic, Z.: 3D Gaussian Ray Tracing: Fast Tracing of Particle Scenes. ACM TOG **43**(6) (2024)
44. Müller, T., Evans, A., Schied, C., Keller, A.: Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. ACM TOG **41**(4) (2022)
45. Niemeyer, M., Manhardt, F., Rakotosaona, M.J., Oechsle, M., Tsalicoglou, C., Tateno, K., Barron, J.T., Tombari, F.: Learning Neural Exposure Fields for View Synthesis. In: NeurIPS (2025)
46. Nilsson, J., Akenine-Möller, T.: Understanding SSIM. arXiv preprint arXiv:2006.13846 (2020)

47. Papantonakis, P., Kopanas, G., Kerbl, B., Lanvin, A., Drettakis, G.: Reducing the Memory Footprint of 3D Gaussian Splatting. *Proc. ACM Comput. Graph. Interact. Tech* **7**(1) (2024)
48. Radl, L., Steiner, M., Parger, M., Weinrauch, A., Kerbl, B., Steinberger, M.: StopThePop: Sorted Gaussian Splatting for View-Consistent Real-time Rendering. *ACM TOG* **43**(4) (2024)
49. Radl, L., Windisch, F., Deixelberger, T., Hladky, J., Steiner, M., Schmalstieg, D., Steinberger, M.: SOF: Sorted Opacity Fields for Fast Unbounded Surface Reconstruction. In: *SIGGRAPH Asia* (2025)
50. Rota Bulò, S., Porzi, L., Kotschieder, P.: Revising Densification in Gaussian Splatting. In: *ECCV* (2024)
51. Rückert, D., Franke, L., Stamminger, M.: ADOP: Approximate Differentiable One-Pixel Point Rendering. *ACM TOG* **41**(4) (2022)
52. Shen, J., Agudo, A., Moreno-Noguer, F., Ruiz, A.: Conditional-Flow NeRF: Accurate 3D Modelling with Reliable Uncertainty Quantification. In: *ECCV* (2022)
53. Shen, J., Ruiz, A., Agudo, A., Moreno-Noguer, F.: Stochastic Neural Radiance Fields: Quantifying Uncertainty in Implicit 3D Representations. In: *3DV* (2021)
54. Steiner, M., Köhler, T., Radl, L., Windisch, F., Schmalstieg, D., Steinberger, M.: AAA-Gaussians: Anti-Aliased and Artifact-Free 3D Gaussian Rendering. In: *ICCV* (2025)
55. Sun, C., Choe, J., Loop, C., Ma, W.C., Wang, Y.C.F.: Sparse Voxels Rasterization: Real-time High-fidelity Radiance Field Rendering. In: *CVPR* (2025)
56. Sünderhauf, N., Abou-Chakra, J., Miller, D.: Density-aware NeRF Ensembles: Quantifying Predictive Uncertainty in Neural Radiance Fields. In: *ICRA* (2022)
57. Tu, X., Radl, L., Steiner, M., Steinberger, M., Kerbl, B., de la Torre, F.: VRsplat: Fast and Robust Gaussian Splatting for Virtual Reality. *Proc. ACM Comput. Graph. Interact. Tech* **8**(1) (2025)
58. Wang, S., Leroy, V., Cabon, Y., Chidlovskii, B., Revaud, J.: DUST3R: Geometric 3D Vision Made Easy. In: *CVPR* (2024)
59. Wang, Y., Wang, C., Gong, B., Xue, T.: Bilateral Guided Radiance Field Processing. *ACM TOG* **43**(4), 1–13 (2024)
60. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE TIP* **13**(4), 600–612 (2004)
61. Wu, J.Z., Zhang, Y., Turki, H., Ren, X., Gao, J., Shou, M.Z., Fidler, S., Gojcic, Z., Ling, H.: Difx3D+: Improving 3D Reconstructions with Single-Step Diffusion Models. In: *CVPR* (2025)
62. Wu, Q., Martinez Esturo, J., Mirzaei, A., Moënné-Loccoz, N., Gojcic, Z.: 3DGUT: Enabling Distorted Cameras and Secondary Rays in Gaussian Splatting. In: *CVPR* (2025)
63. Xue, S., Dill, J., Mathur, P., Dellaert, F., Tsiotras, P., Xu, D.: Neural Visibility Field for Uncertainty-Driven Active Mapping. In: *CVPR* (2024)
64. Yang, Z., Gao, X., Sun, Y., Huang, Y., Lyu, X., Zhou, W., Jiao, S., Qi, X., Jin, X.: Spec-Gaussian: Anisotropic View-Dependent Appearance for 3D Gaussian Splatting. In: *NeurIPS* (2024)
65. Ye, Z., Li, W., Liu, S., Qiao, P., Dou, Y.: AbsGS: Recovering Fine Details in 3D Gaussian Splatting. In: *ACM MM* (2024)
66. Yeshwanth, C., Liu, Y.C., Niekner, M., Dai, A.: ScanNet++: A High-Fidelity Dataset of 3D Indoor Scenes. In: *ICCV* (2023)
67. Yu, M., Lu, T., Xu, L., Jiang, L., Xiangli, Y., Dai, B.: GSDF: 3DGS Meets SDF for Improved Rendering and Reconstruction. In: *NeurIPS* (2024)

68. Yu, Z., Chen, A., Huang, B., Sattler, T., Geiger, A.: Mip-Splatting: Alias-free 3D Gaussian Splatting. In: CVPR (2024)
69. Yu, Z., Sattler, T., Geiger, A.: Gaussian Opacity Fields: Efficient Adaptive Surface Reconstruction in Unbounded Scenes. ACM TOG **43**(6) (2024)
70. Zhang, B., Fang, C., Shrestha, R., Liang, Y., Long, X.X., Tan, P.: RaDe-GS: Rasterizing Depth in Gaussian Splatting. arXiv preprint arXiv:2406.01467 (2024)
71. Zhang, D., Wang, C., Wang, W., Li, P., Qin, M., Wang, H.: Gaussian in the Wild: 3D Gaussian Splatting for Unconstrained Image Collections. In: ECCV (2024)
72. Zhang, W., Liu, Y.S., Han, Z.: Neural Signed Distance Function Inference through Splatting 3D Gaussians Pulled on Zero-Level Set. In: NeurIPS (2024)
73. Zhang, Z., Roussel, N., Muller, T., Zeltner, T., Nimier-David, M., Rousselle, F., Jakob, W.: Radiance Surfaces: Optimizing Surface Representations with a 5D Radiance Field Loss. In: SIGGRAPH (2025)
74. Zhang, Z., Huang, B., Jiang, H., Zhou, L., Xiang, X., Shen, S.: Quadratic Gaussian Splatting: High Quality Surface Reconstruction with Second-order Geometric Primitives. In: ICCV (2025)
75. Zhao, C., Wang, X., Zhang, T., Javed, S., Salzmänn, M.: Self-Ensembling Gaussian Splatting for Few-Shot Novel View Synthesis. In: ICCV (2025)