Eurographics Symposium on Rendering (2025)
B. Wang and A. Wilkie (Editors)

Uncertainty-Aware Gaussian Splatting with View-Dependent Regularization for High-Fidelity 3D Reconstruction

Shengjun Liu[1] , Jiangxin Wu[1] , Wenhui Wu[1] , Lixiang Chu[1] , Xinru Liu[1]

[1]Institute of Engineering Modeling and Scientific Computing, Central South University, Changsha, Hunan 410083, China

Abstract

3D Gaussian Splatting (3DGS) has emerged as a groundbreaking paradigm for explicit scene representation, achieving photorealistic novel view synthesis with real-time rendering capabilities. However, reconstructing geometrically consistent and accurate surfaces under complex real-world scenarios remains a critical challenge. Current 3DGS frameworks primarily rely on photometric loss optimization, which often results in multi-view geometric inconsistencies and inadequate handling of texture-less regions due to two inherent limitations: 1) the absence of explicit geometric constraints during Gaussian parameter optimization, and 2) the lack of effective mechanisms to resolve multi-view geometric ambiguities. To address these challenges, we propose ***Uncertainty-Aware Gaussian Splatting (UA-GS)****, a novel framework that integrates geometric priors with view-dependent uncertainty modeling to explicitly capture and resolve multi-view inconsistencies. For efficient optimization of Gaussian attributes, our approach introduces a spherical harmonics-based uncertainty representation that dynamically models view-dependent geometric variations. Building on this framework, we further design uncertainty-aware optimization strategies. Extensive experiments on real-world and synthetic benchmarks demonstrate that our method significantly outperforms state-of-the-art 3DGS-based approaches in geometric accuracy while retaining competitive rendering quality. Code and data will be made available soon.*

CCS Concepts

• ***Computing methodologies*** → *Collision detection;* • ***Hardware*** → *Sensors and actuators; PCB design and layout;*

1. Introduction

With the rapid development of augmented reality (AR), virtual reality (VR) [DHY*22, TLH*24, YCW*23], and autonomous driving systems [ZLS*24, YLZ*24, YIL*23], high-fidelity 3D reconstruction and photorealistic novel view synthesis have become increasingly crucial. Traditional multi-view stereo (MVS) methods [GCS06], though effective for surface reconstruction via multi-view depth estimation, often struggle with complex scene geometries and suffer from high computational costs. Neural implicit representations, particularly Neural Radiance Fields (NeRF) [MST*21], represent important advances in geometric modeling, but NeRF has low geometric reconstruction accuracy despite its outstanding visual quality. To enhance the geometric precision of NeRF-based approaches, numerous improved methods have emerged [MESK22, CLS*23, WLL*21, WHH*23, FXOT22, LME*23]. However, these methods still face long training cycles and huge computational overhead, limiting their efficiency and real-time performance in practical applications [WLL*21, WHH*23, FXOT22, LME*23].

In recent years, 3D Gaussian Splatting (3DGS) [KKLD23] has emerged as an explicit scene representation paradigm, enabling efficient 3D modeling and rendering through differentiable Gaussian primitives. It achieves fully-real-time rendering speeds while maintaining competitive visual quality. However, in complex real-world environments, its geometric reconstruction accuracy remains constrained by inherent limitations in photometric loss optimization strategies: (1) the lack of explicit geometric constraints, and (2) deficiencies in traditional normalized cross-correlation (NCC) [YH09] photometric optimization. Conventional NCC-based methods [FXOT22, CLY*24] primarily rely on reprojection error filtering to eliminate pixels with large errors; however, this single screening criterion exhibits significant drawbacks in high-uncertainty regions (e.g., occlusion boundaries, weakly textured areas). We visualize the reconstruction results using the traditional NCC method in Figure 6. When reprojection errors are coincidentally suppressed due to local noise or viewpoint variations, low-confidence geometric estimates may still be erroneously retained, leading to surface distortions or artifacts. This issue stems from the absence of explicit modeling for multi-view geometric consistency.

To address these critical limitations, inspired by advancements in depth estimation and normal prediction [BD24, XXYG24, YLSY19, YPN*22], we propose a novel uncertainty-aware 3DGS framework that integrates geometric priors with view-dependent uncertainty modeling to explicitly capture and resolve inter-view

DOI: 10.2312/sr.20251192

geometric inconsistencies. Specifically, we innovatively introduce a spherical harmonics-based uncertainty representation to dynamically model view-dependent geometric variations, enabling efficient optimization of Gaussian attributes.

Building upon this framework, we further develop an uncertainty-aware normal refinement method to enhance normal estimation. Additionally, during the NCC-based optimization, we augment the traditional reprojection error filtering mechanism with an uncertainty-conditioned pixel screening strategy. This dual approach effectively eliminates high-uncertainty pixels, preventing noise-induced optimization interference while ensuring more accurate geometric recovery in textureless regions and other high-ambiguity areas.

The main contributions of our work are summarized as follows:

- **Uncertainty-aware Geometric Reconstruction**: We propose a novel uncertainty quantification framework that jointly models depth and normal estimation uncertainties during optimization. Through renderable uncertainty maps, our method efficiently identifies geometrically ambiguous regions, enabling real-time optimization.
- **View-dependent Spherical Harmonics Representation**: We innovatively parameterize uncertainties using spherical harmonics, effectively capturing view-dependent geometric uncertainties to resolve multi-view inconsistencies while maintaining computational efficiency.
- **Uncertainty-aware Geometric Refinement**: We design a dynamic optimization strategy comprising:
  - Uncertainty-aware refinement of normal estimation;
  - An enhanced NCC photometric optimization that integrates uncertainty-conditioned pixel filtering with traditional reprojection error screening, effectively eliminating noisy constraints to improve geometric accuracy.

2. Related Work

2.1. Neural Rendering and Radiance Fields

Neural implicit representation has emerged as a core framework for novel view synthesis and surface reconstruction. NeRF [MST*21] employs MLPs to predict volume density and color through continuous volume rendering, achieving high-quality view synthesis. However, its computational overhead and data dependency limit applications in large-scale scenes. Subsequent improvements adopt implicit surface modeling by integrating volume rendering with geometric detail reconstruction, reducing reliance on 3D supervision while enhancing geometric fidelity [NMOG20, WLL*21, YGKL21].

To address the high computational cost of continuous network inference, researchers propose discretized scene representations: point clouds [XXP*22], voxels [LYZ*22, LGZL*20], and sparse encodings [LME*23, MESK22], improving processing efficiency through spatial decomposition. Further innovations include adaptive basis functions [CLS*23], geometric consistency constraints [FXOT22], and hierarchical representations [WHH*23, CXG*22, FKYT*22, TLY*21, GKJ*21, HSM*21], which significantly enhance inference speed without compromising rendering quality. Kilonerf [RPLG21] achieves accelerated inference via large-scale lightweight MLP clusters, providing new insights for real-time rendering.

Current challenges remain in high-precision 3D reconstruction and real-time rendering for complex scenes, particularly under dynamic lighting conditions where computational costs and extended training times persist.

2.2. 3D Gaussian Splatting

Recent advancements in 3D Gaussian Splatting [KKLD23] enable real-time radiance field rendering and high-fidelity surface reconstruction by discretizing continuous scenes into learnable 3D Gaussian primitives. Building on this, the surface-oriented 2D Gaussian Splatting [HYC*24] projects radiance fields onto differentiable planar surfaces, reducing computational complexity while jointly optimizing geometry and appearance.

Further extensions have expanded the geometric modeling capabilities. SuGaR [GL24] introduces geometry alignment constraints to reconstruct both implicit surfaces and explicit meshes simultaneously, and GOF [YSG24] implements adaptive Gaussian opacity fields that dynamically control density, preserving structural coherence in unbounded scenes and capturing fine surface details. To address the challenge of defining surface boundaries, GS-Pull [ZLH24] uses zero-level set concentration to align Gaussian distributions with implicit surface boundaries, enabling efficient signed distance function (SDF) inference. In addition, RaDe-GS [ZFS*24] incorporates depth-aware loss functions with differentiable rasterization to enforce perspective-consistent geometry, while PGSR [CLY*24] leverages structural symmetry priors to enhance reconstruction fidelity without additional computational cost.

2DGS-Room [ZXL*24] improves indoor reconstruction by introducing seed-guided Gaussian initialization with monocular geometric priors and multi-view consistency constraints, significantly enhancing structural accuracy in low-texture regions. GS-ID [DLW24] further extends 2DGS with intrinsic diffusion priors and parametric light modeling, enabling high-quality illumination decomposition and editable relighting via deferred physically-based rendering.

Generative paradigms have also emerged. Dreamgaussian [TRZ*23] employs a diffusion-based framework to learn Gaussian primitive distributions for high-quality 3D content synthesis, bridging physical reconstruction with creative design. Meanwhile, NeuSG [CLL23] combines explicit 3DGS geometric priors with implicit neural representations via joint optimization, achieving superior accuracy.

While recent 3DGS-based methods advance geometric alignment and multi-view consistency, they inadequately address sensor noise and geometric ambiguity, often compromising detail fidelity in weakly observed regions. Our work introduces uncertainty-aware optimization by encoding view-dependent confidence into Gaussian attributes, dynamically guiding geometry refinement to suppress noise propagation while preserving fine topology.

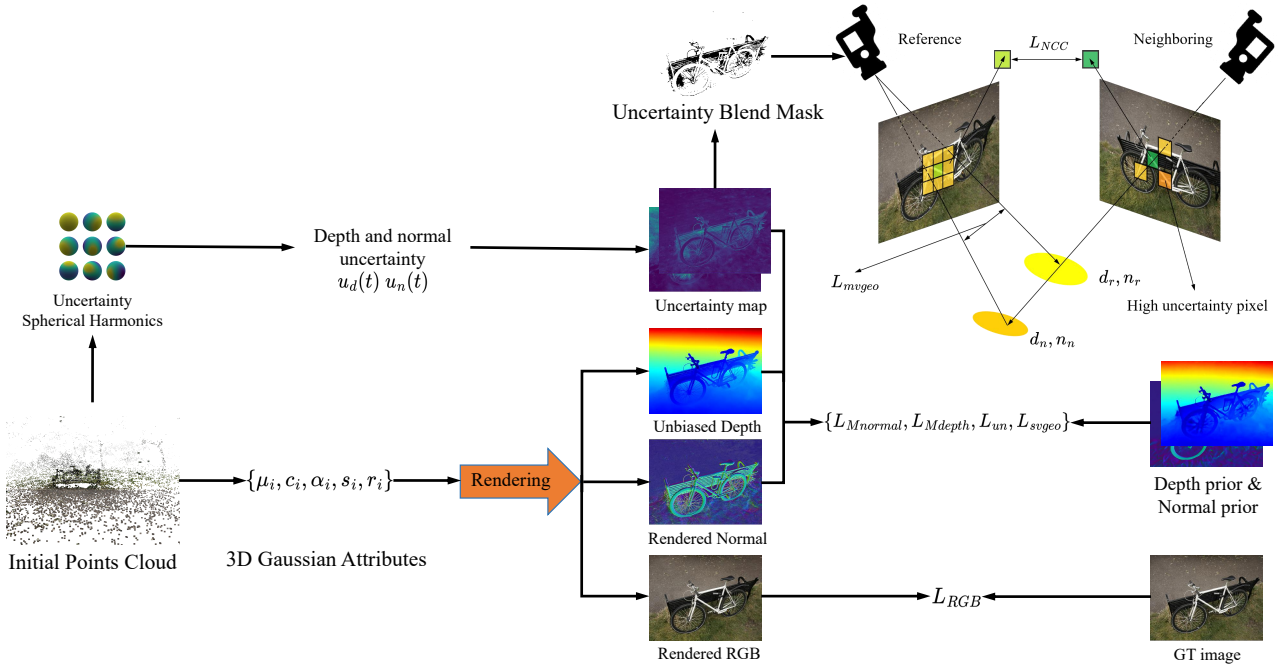

Figure 1: ***Framework Overview.*** *We use depth and normal priors obtained from pre-trained large models to enhance the reconstruction quality of Gaussian splatting. Using the depth and normal prior, we model the depth and normal uncertainty that measure multi-view inconsistency and use them to guide the reconstruction of the Gaussian splatting, as well as mesh the reconstruction results more accurately by adding geometric regularization.*

2.3. Uncertainty Modeling

In multi-view stereo, uncertainty-based adaptive depth sampling [SXT22] enhances probabilistic depth estimation by dynamically refining depth search ranges using confidence measures, reducing cost volume redundancy through staged refinement. Concurrently, visibility-aware fusion [ZYL*20] strengthens geometric consistency by weighting cost volume aggregation with pair-wise matching uncertainties, mitigating occlusion interference. These solutions respectively tackle uncertainty modeling and geometric reliability challenges. Complementary uncertainty-aware photometric stereo methods [KKO*22] selectively integrate reliable predictions from deep photometric stereo and multi-view stereo through Bayesian neural networks, discarding unreliable depth and normal predictions via confidence thresholds to improve surface reconstruction.

Bayesian frameworks provide solid theoretical foundations for uncertainty modeling, enabling the propagation of geometric uncertainty through differentiable pipelines [AAHL20]. Depth completion techniques enhance sparse-to-dense depth estimation by employing Bayesian deep basis fitting [QLT21] to model uncertainty, while leveraging the confidence propagation mechanism of normalized convolutions [EFHP20], which utilizes self-supervised input confidence learning and explicit uncertainty propagation to estimate output variance. Recent work also demonstrates the utility of sampling-based Bayesian inference for reflectance modeling, where stochastic particle optimization [ZZD*24] captures parameter and image-space uncertainties, guiding acquisition with entropy-based view selection. Alternatively, frequency-domain analysis offers a lightweight approximation of SVBRDF ambiguity [WPH*24], enabling fast uncertainty estimation and error-aware material acquisition in challenging visual conditions.

In neural implicit representations, addressing geometric bias and detail loss in reconstruction has shown significant promise [XXYG24]. This effort is complemented by conditional normalizing flows for uncertainty-aware radiance field modeling, which enhances the robustness of novel view synthesis [SAMNR22]. Additionally, active sampling strategies prioritizing novel viewpoints covering high-uncertainty regions optimize training efficiency and reconstruction robustness [PLSH22]. For surface normal estimation, modeling aleatoric uncertainty via a parameterized angular von Mises-Fisher distribution improves accuracy by adaptively weighting the loss based on input-dependent noise [BBC21].

Explicit 3D reconstruction techniques benefit from adaptive thin volume representations with uncertainty awareness [CXZ*20] and view-consistent depth-normal regularization for Gaussian surface reconstruction [CWL*24]. These methods improve reconstruction precision through variance-based uncertainty modeling and multi-view geometric consistency constraints, though they primarily address aleatoric uncertainty without explicit separation from epistemic uncertainty.

Existing uncertainty-aware methods face inherent trade-offs between computational efficiency and modeling precision: cost vol-

umes incur redundant computation, implicit neural fields lack explicit uncertainty propagation, and explicit Gaussian methods conflate sensor noise with geometric ambiguity. Our framework addresses these limitations by embedding view-dependent uncertainty directly into 3D Gaussian representations, enabling real-time uncertainty propagation and dynamic topology refinement. Unlike monolithic uncertainty modeling, our joint depth-normal confidence guidance suppresses artifacts in ambiguous regions while preserving computational efficiency.

3. Preliminary

3.1. 3D Gaussian Splatting

3D Gaussian Splatting (3DGS) is a Gaussian primitive-based 3D scene representation method that models scene geometry by distributing multiple Gaussian primitives in 3D space. Each Gaussian primitive is defined as a 3D Gaussian distribution with attributes such as position, scale, orientation, and color. In computer graphics, 3DGS demonstrates efficient scene reconstruction and rendering capabilities, particularly excelling in multi-view reconstruction tasks.

In 3DGS, a scene is composed of multiple 3D Gaussian primitives. Each Gaussian primitive is characterized by its position $\boldsymbol{\mu} \in \mathbb{R}^3$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{3\times3}$, with its distribution function defined as:

$$G(\boldsymbol{x}) = \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^3$ represents an arbitrary point in the scene, $\boldsymbol{\mu}$ is the center of the Gaussian primitive, and $\boldsymbol{\Sigma}$ is the covariance matrix. To optimize computation, the covariance matrix $\boldsymbol{\Sigma}$ is typically decomposed into a product of a scale matrix $\boldsymbol{S} \in \mathbb{R}^3$ and a rotation matrix $\boldsymbol{R} \in \mathbb{R}^{3\times3}$:

$$\boldsymbol{\Sigma} = \boldsymbol{RSS^TR^T}, \tag{2}$$

where $\boldsymbol{S}$ is a diagonal matrix encoding scale information, and $\boldsymbol{R}$ is a rotation matrix representing the orientation of the Gaussian primitive.

When rendering novel views, Gaussian primitives must be transformed into the camera coordinate system. Let $\boldsymbol{W}$ denote the view transformation matrix of the camera. The transformation applied to the position $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of a Gaussian primitive is implemented as follows:

$$\boldsymbol{\mu'} = \boldsymbol{KW}[\boldsymbol{\mu}, 1]^T, \tag{3}$$

$$\boldsymbol{\Sigma'} = \boldsymbol{JW\Sigma W^TJ^T}, \tag{4}$$

where $\boldsymbol{K}$ is the cameras intrinsic matrix, $\boldsymbol{J}$ is the Jacobian matrix of the perspective projection, and $\boldsymbol{W}$ is the world-to-camera rotation matrix. The transformed covariance matrix $\boldsymbol{\Sigma'}$ represents an approximation of the Gaussian primitive aligned with the projection in the camera coordinate system.

The rendering process in 3D Gaussian Splatting employs sorted alpha blending. The final pixel color is computed through opacity and accumulated transmittance, with each Gaussian primitives contribution determined via alpha blending. The blending coefficient α_i for the i-th Gaussian primitive is given by:

$$\alpha_i(\boldsymbol{x}) = o_i \cdot \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_i})^T\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_i})\right), \tag{5}$$

where o_i is the base opacity of the i-th Gaussian primitive, $\boldsymbol{x}$ is the 3D position, $\boldsymbol{\mu_i}$ is the center of the i-th Gaussian primitive, and $\boldsymbol{\Sigma_i}$ is its covariance matrix.

The final color $\boldsymbol{C}$ at a pixel is computed as the weighted sum of contributions from all Gaussian primitives along the ray:

$$\boldsymbol{C} = \sum_{i\in N} T_i\alpha_i\boldsymbol{c_i}, \quad T_i = \prod_{j=1}^{i-1}(1-\alpha_j), \tag{6}$$

where $\boldsymbol{c_i} \in \mathbb{R}^3$ denotes the color of the i-th Gaussian primitive (typically represented using Spherical Harmonics), α_i is its blending coefficient, and N is the total number of Gaussian primitives intersected by the ray.

3.2. Depth Estimation

To improve depth estimation accuracy, we adopt an unbiased depth estimation strategy inspired by PGSR. The rendered depth $\hat{D}(\boldsymbol{p})$ and normal $\hat{\boldsymbol{N}}(\boldsymbol{p})$ at pixel $\boldsymbol{p}$ are defined as:

$$\hat{D}(\boldsymbol{p}) = \frac{\sum_{i\in N} d_i\alpha_i\prod_{j=1}^{i-1}(1-\alpha_j)}{\hat{\boldsymbol{N}}(\boldsymbol{p})\mathbf{K}^{-1}\tilde{\mathbf{p}}}, \tag{7}$$

$$\hat{\boldsymbol{N}}(\boldsymbol{p}) = \sum_{i\in N}\mathbf{R}_c^T\boldsymbol{n}_i\alpha_i\prod_{j=1}^{i-1}(1-\alpha_j), \tag{8}$$

where $d_i = (\mathbf{R}_c^\top(\boldsymbol{\mu}_i - \mathbf{T}_c))^\top\mathbf{R}_c^\top\boldsymbol{n}_i$ represents the distance from the Gaussian primitive to the camera center, $\mathbf{T}_c$ is the camera center in the world, $\mathbf{R}_c$ is the rotation matrix from the camera to the world coordinate system, and $\boldsymbol{n}_i$ is the normal vector of the i-th Gaussian primitive. Here, $\boldsymbol{p} = [u,v]^T$ denotes the 2D pixel coordinate, $\tilde{\boldsymbol{p}}$ is its homogeneous form, and $\mathbf{K}$ is the camera intrinsic matrix.

To address the limited prediction range of depth estimation networks, we solve the problem by normalizing the depth prior using a least-squares-derived scale w and offset q [EPF14]:

$$D(\boldsymbol{p}) = wD'(\boldsymbol{p}) + q, \tag{9}$$

where $D'(\boldsymbol{p})$ is the depth prior predicted by a pre-trained depth estimation network. The monocular depth and normal regularization losses are defined as:

$$\mathcal{L}_{\text{depth}} = \sum_{\boldsymbol{p}\in\mathcal{R}}\|w\hat{D}(\boldsymbol{p}) + q - D(\boldsymbol{p})\|^2, \tag{10}$$

$$\mathcal{L}_{\text{normal}} = \sum_{\boldsymbol{p}\in\mathcal{R}}\|\hat{\boldsymbol{N}}(\boldsymbol{p}) - \boldsymbol{N}(\boldsymbol{p})\|_1 + \|1 - \hat{\boldsymbol{N}}(\boldsymbol{p})^\top\boldsymbol{N}(\boldsymbol{p})\|_1, \tag{11}$$

where $D(\boldsymbol{p})$ and $\boldsymbol{N}(\boldsymbol{p})$ are the depth and normal priors provided by the pre-trained Metric3D V2 model [HYZ*24].

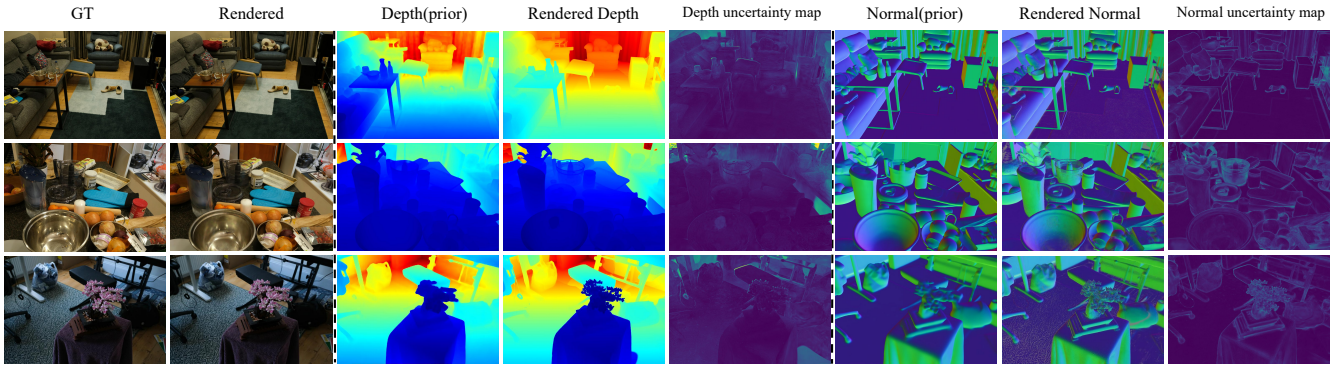

Figure 2: ***The rendering of depth, normal, and uncertainty maps for various scenes that we have achieved.*** *The rendered uncertainty map can localize regions where the monocular a prior map is inaccurate, which usually correspond to thin structures with a high level of detail. Thus, we can use the rendered uncertainty map to guide the optimization of the mesh structure.*

4. Method

In our UA-GS framework, we introduce uncertainty modeling for 3DGS. By applying adaptive masked depth and normal regularization, we effectively suppress the propagation of noise in regions of high uncertainty. To efficiently model view-dependent uncertainties and ensure that the uncertainty map retains the efficient rendering characteristics of Splatting, we define depth uncertainty and normal uncertainty attributes for each Gaussian primitive, using spherical harmonic functions to model the view-dependent uncertainties. This approach allows us to achieve more accurate normal estimation, refining the normal priors through the uncertainty map. To further ensure multi-view geometric consistency, we integrate multi-view geometric consistency with PGSR-based losses and optimize photometric consistency using uncertainty-aware NCC regularization.

4.1. Uncertainty Modeling for 3DGS

Modern 3D reconstruction systems must account for inherent sensor noise and observational uncertainties arising from dynamic occlusions, lighting variations, and geometric ambiguities. Our framework addresses these challenges through explicit uncertainty modeling integrated with 3DGS, enabling robust geometry reasoning while maintaining real-time rendering capabilities.

To model this uncertainty, we assign a depth uncertainty attribute u_{d_i} and a normal uncertainty attribute u_{n_i} to each Gaussian primitive. It is important to note that this uncertainty is view-dependent. To maintain rendering efficiency while capturing view-dependent variations, we model uncertainty fields using spherical harmonics, enabling seamless integration with standard 3DGS splatting operations. The technical details are explained in detail in Section 4.2. This formulation allows analytical uncertainty computation during differentiable rendering while preserving the parametric advantages of 3DGS.

However, merely parameterizing uncertainties does not inherently prevent their propagation during optimization. Uncertain regions with ambiguous geometry may introduce erroneous gradients that degrade reconstruction quality. To mitigate this, according to the method [XXYG24], we introduce adaptive regularization mechanisms to mitigate noise propagation from uncertain regions:
Masked Depth Regularization:

$$\mathcal{L}_{\text{Mdepth}} = \ln\left|\hat{U}_d(\boldsymbol{p})\right| + \frac{\Omega(\hat{U}_d(\boldsymbol{p}),\tau_d)\odot\left|\hat{D}(\boldsymbol{p}) - D(\boldsymbol{p})\right|}{\left|\hat{U}_d(\boldsymbol{p})\right|} \tag{12}$$

where $\hat{U}_d(\boldsymbol{p})$ is the depth uncertainty, $\hat{D}(\boldsymbol{p})$ is the predicted depth geometry prior, and $D(\boldsymbol{p})$ is the rendered depth map.$\Omega(U,\tau)$ serves as an adaptive gradient separator, detaching gradients from regions exceeding depth uncertainty threshold τ_d to prevent noise interference. To ensure consistency in scale, the depth in the formula also needs to undergo a least squares transformation.

Masked Normal Consistency:

$$\mathcal{L}_{\text{Mnormal}} = \ln\hat{U}_n^2(\boldsymbol{p}) + \frac{\Omega(\hat{U}_n(\boldsymbol{p}),\tau_n)\odot\|\hat{\boldsymbol{N}}(\boldsymbol{p}) - \boldsymbol{N}(\boldsymbol{p})\|_2}{\hat{U}_n^2(\boldsymbol{p})} \tag{13}$$

where $\hat{U}_n(\boldsymbol{p})$ is the normal uncertainty, $\hat{\boldsymbol{N}}(\boldsymbol{p})$ is the predicted normal geometry prior, and $\boldsymbol{N}(\boldsymbol{p})$ is the rendered normal map.This selectively enforces normal consistency in low-uncertainty regions ($\hat{U}_n < \tau_n$) while suppressing erroneous gradients from ambiguous areas.

Uncertainty Regularized Total Loss Thus the uncertainty regularized total loss $\mathcal{L}_{MM}$ is:

$$\mathcal{L}_{MM} = \lambda_1\mathcal{L}_{\text{Mnormal}} + \lambda_2\mathcal{L}_{\text{Mdepth}} \tag{14}$$

The unified framework dynamically regulates Gaussian densification and pruning based on uncertainty maps while maintaining real-time performance. This co-optimization of geometry and uncertainty fields enhances reconstruction fidelity, particularly for sparse viewpoints and weakly-textured surfaces where traditional methods struggle.

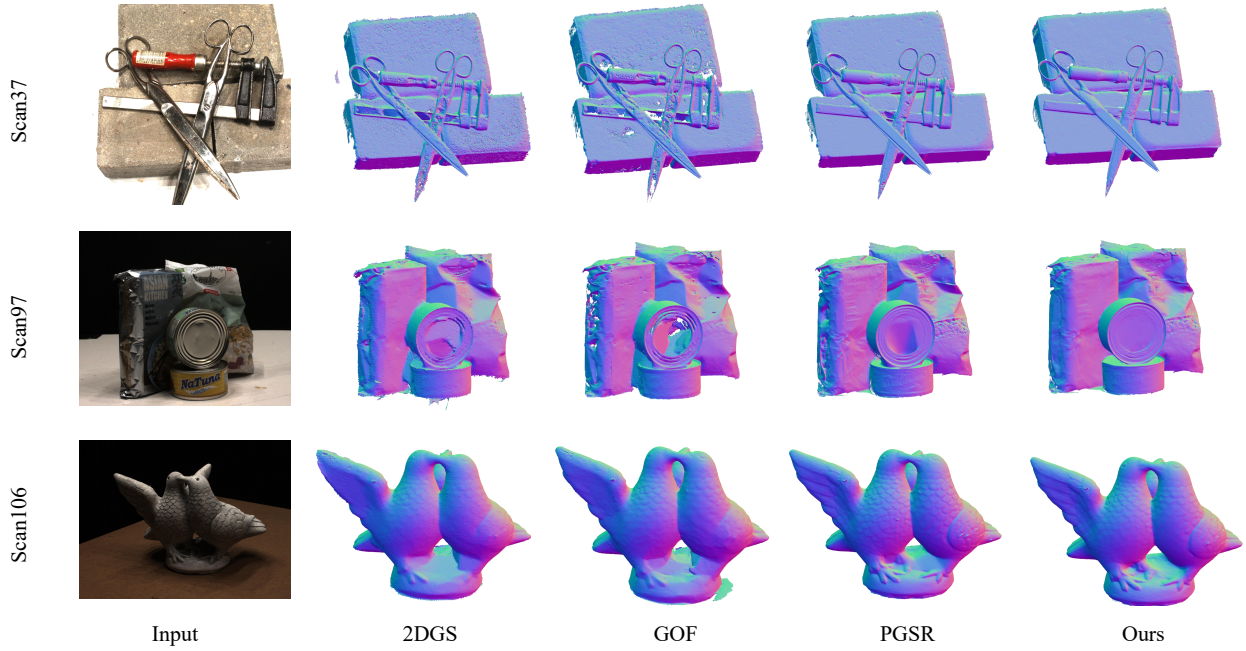

Figure 3: ***Qualitative comparison on DTU dataset.*** *Our approach achieves more complete and topologically correct surfaces.*

4.2. View-dependent Uncertainty-aware Prior Screening via Spherical Harmonics

Existing monocular depth and normal prior methods [HYZ*24, YZC*23, ESMZ21] often overlook the complexity of view dependency and uncertainty modeling. To address this, we propose a view-dependent uncertainty-aware prior screening framework based on spherical harmonics. Unlike previous approaches [XXYG24] that employ multilayer perceptrons(MLPs) to predict uncertainties, our method leverages the inherent ability of spherical harmonics to model view-dependent uncertainties with superior computational efficiency.

While prior works [XXYG24] employ MLPs to predict uncertainties through spatial coordinates $\boldsymbol{x}$ and view direction $\boldsymbol{v}$ inputs, such approaches suffer from slow training convergence and querying uncertainties for billions of Gaussian primitives significantly degrades real-time rendering performance. In contrast, spherical harmonics excel at directionally continuous uncertainty representation while maintaining high training efficiency.

Specifically, for a given view direction $\boldsymbol{v}$ parameterized by polar angle θ and azimuth ϕ and spatial position $\boldsymbol{x}$, we parameterize view-dependent uncertainty as:

$$U(\boldsymbol{x},\boldsymbol{v}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\phi), \tag{15}$$

where $Y_{lm}(\theta,\phi)$ denotes the standard spherical harmonic basis functions, a_{lm} are learnable coefficients. This formulation generates smooth, directionally sensitive uncertainty fields.

This representation enables direct integration with 3D Gaussian Splatting rasterization pipelines for real-time uncertainty map generation. The rendered depth and normal uncertainties at pixel $\boldsymbol{p}$ are computed through parallelizable splatting operations:

$$\hat{U}_d(\boldsymbol{p}) = \sum_{i\in N} T_i \alpha_i u_{d_i}, \quad \hat{U}_n(\boldsymbol{p}) = \sum_{i\in N} T_i \alpha_i u_{n_i}, \tag{16}$$

where u_{d_i} and u_{n_i} denote the depth and normal uncertainties at the intersection of the ray with the i-th Gaussian primitive, as parametrized and computed by Eq. 15 while T_i and α_i follow the definitions in Section 3.1.

4.3. Geometric Regularization

Geometric regularization plays a key role in ensuring that the geometry is reasonable. Our framework ensures that Gaussian primitives are aligned with the plane through scale regularization, while utilizing uncertainty-aware normal constraints to attenuate unreliable signals. In addition, a multi-view consistency term ensures that geometric information from different views remains consistent. Together, the components build a geometric regularization term that effectively solves the geometric ambiguity problem, reduces the error propagation in the uncertainty region, and harmonizes the local details with the overall structure, which significantly improves the reconstruction accuracy of the geometric structure and weakly textured regions.

Scale Regularization To enhance geometric alignment with scene surfaces, we use the scale regularization loss first proposed

Table 1: ***Quantitative Comparison on the DTU Dataset [JDV*14]****. We report the Chamfer Distance, Average Optimization Time, and Reconstruction Accuracy of different methods. Our method achieves the best accuracy and is more than 100 times faster than the NeRF-based SDF method.*

	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean	Time
NeRF [MST*21]	1.90	1.60	1.85	0.58	2.28	1.27	1.47	1.67	2.05	1.07	0.88	2.53	1.06	1.15	0.96	1.49	>12h
VolSDF [YGKL21]	1.14	1.26	0.81	0.49	1.25	0.70	0.72	1.29	1.18	0.70	0.66	1.08	0.42	0.61	0.55	0.86	>12h
NeuS [WLL*21]	1.00	1.37	0.93	0.43	1.10	0.65	0.57	1.48	1.09	0.83	0.52	1.20	0.35	0.49	0.54	0.84	>12h
Neuralangelo [LME*23]	0.37	0.72	0.35	0.35	0.87	0.54	0.53	1.29	0.97	0.73	0.47	0.74	0.32	0.41	0.43	0.61	>128h
SuGaR [GL24]	1.47	1.33	1.13	0.61	2.25	1.71	1.15	1.63	1.62	1.07	0.79	2.45	0.98	0.88	0.79	1.33	1h
2D GS [HYC*24]	0.48	0.91	0.39	0.39	1.01	0.83	0.81	1.36	1.27	0.76	0.70	1.40	0.40	0.76	0.52	0.80	0.32h
GOF [YSG24]	0.50	0.82	0.37	0.37	1.12	0.74	0.73	1.18	1.29	0.68	0.77	0.90	0.42	0.46	0.49	0.74	2h
GS-Pull [ZLH24]	0.51	0.56	0.46	0.39	0.82	0.67	0.85	1.37	1.25	0.73	0.54	1.39	0.35	0.88	0.42	0.75	0.36h
RaDe-GS [ZFS*24]	0.40	0.71	0.33	0.37	0.87	0.79	0.77	1.22	1.26	0.70	0.65	0.85	0.33	0.66	0.44	0.69	0.34h
PGSR [CLY*24]	0.34	0.58	0.29	0.29	0.78	0.58	0.54	1.01	0.73	0.51	0.49	0.69	0.31	0.37	0.38	0.53	0.6h
Ours	0.32	0.47	0.29	0.28	0.75	0.53	0.46	0.93	0.65	0.49	0.37	0.55	0.30	0.34	0.32	0.47	0.49h

by NeuSG [CLL23]:

$$\mathcal{L}_s = \|\min(s_1, s_2, s_3)\|_1, \tag{17}$$

which penalizes the minimal axis scale (s_1, s_2, s_3) of each Gaussian.

Geometric Consistency We directly adopt the geometric regularization framework from PGSR [CLY*24] to enforce geometric consistency, including single-view geometric regularization and multi-view geometric consistency, as

$$\mathcal{L}_{\text{svgeo}} = \frac{1}{W} \sum_{\boldsymbol{p} \in W} |1 - \nabla I|^5 \|\boldsymbol{N_d}(\boldsymbol{p}) - \hat{\boldsymbol{N}}(\boldsymbol{p})\|_1. \tag{18}$$

$$\mathcal{L}_{\text{mvgeo}} = \frac{1}{V} \sum_{\boldsymbol{p} \in V} \phi(\boldsymbol{p}), \quad \phi(\boldsymbol{p}) = \|\boldsymbol{p} - \boldsymbol{H}_{nr}\boldsymbol{H}_{rn}\boldsymbol{p}\|. \tag{19}$$

where ∇I is the image gradient normalized to $[0,1]$, $\boldsymbol{N_d}(\boldsymbol{p})$ is computed from the depth map based on the local plane assumption, and W is the set of image pixels. V is the set of all pixels in the image excluding those with high forward and backward projection errors, and $\boldsymbol{H}_{rn}$, $\boldsymbol{H}_{nr}$ denote the homography matrices between reference and neighboring views.

Uncertainty-aware Normal Regularization To enhance multi-view consistency and reconstruction robustness, we propose an uncertainty-weighted normal regularization strategy. Traditional normal regularization [YPN*22, TRM*25] directly enforces constraints using known normal vectors, but factors such as image noise, blur, and weak textures in multi-view reconstruction can compromise normal accuracy. Thus, uncertainty-weighted regularization improves robustness by adaptively adjusting the influence of normal priors.

Specifically, we screen reliable normal priors using rendered normal uncertainty maps. For each pixel, we compute its corresponding uncertainty value and adjust the weight of its normal constraint accordingly: regions with low uncertainty receive higher weights to strengthen their contributions, while high-uncertainty regions are downweighted to mitigate their impact.

The uncertainty-weighted normal regularization loss is defined as:

$$\mathcal{L}_{\text{un}} = \sum_{\boldsymbol{p} \in \mathcal{R}} \frac{\|\hat{\boldsymbol{N}}(\boldsymbol{p}) - \boldsymbol{N}(\boldsymbol{p})\|_1 + \|1 - \hat{\boldsymbol{N}}(\boldsymbol{p})^\top \boldsymbol{N}(\boldsymbol{p})\|_1}{e^{\hat{U}_n(\boldsymbol{p})}}, \tag{20}$$

where $\hat{U}_n(\boldsymbol{p})$ denotes the normal uncertainty at pixel $\boldsymbol{p}$. This formulation ensures that normal constraints dominate in low-uncertainty regions while being suppressed in high-uncertainty areas, enhancing reconstruction accuracy and robustness.

Total Geometric Regularization Loss The total geometric regularization loss combines all components:

$$\mathcal{L}_{\text{geo}} = \lambda_3 \mathcal{L}_s + \lambda_4 \mathcal{L}_{\text{un}} + \lambda_5 \mathcal{L}_{\text{svgeo}} + \lambda_6 \mathcal{L}_{\text{mvgeo}}. \tag{21}$$

4.4. Uncertainty-Aware NCC Regularization

Multi-view photometric consistency is typically measured via normalized cross-correlation (NCC) between image patches. However, noise, blur, and occlusions can degrade geometric reliability in certain regions. To enhance robustness, we propose an uncertainty-aware NCC regularization that leverages pixel-wise uncertainty to reduce the influence of high-uncertainty areas.

We compute a hybrid uncertainty map $A(\boldsymbol{p})$ like DebSDF [XXYG24] by combining depth uncertainty $\hat{U}_d(\boldsymbol{p})$ and normal uncertainty $\hat{U}_n(\boldsymbol{p})$:

$$A(\boldsymbol{p}) = \hat{U}_d^{1-\lambda}(\boldsymbol{p}) \hat{U}_n^{\lambda}(\boldsymbol{p}), \tag{22}$$

where the blending weight $\lambda = 0.9$ prioritizes normal uncertainty due to its critical role in high-precision surface reconstruction.

To refine optimization, we apply a thresholding strategy. Only pixels with hybrid uncertainty below a threshold τ_h contribute to the NCC loss [CLY*24]:

$$\mathcal{L}_{\text{NCC}} = \frac{1}{|\mathcal{V}|} \sum_{\boldsymbol{p} \in \mathcal{V}} \left(1 - \text{NCC}\left(I_r(\boldsymbol{p}), I_n(\boldsymbol{H}_{rn}\boldsymbol{p})\right)\right), \tag{23}$$

where $I_r(\boldsymbol{p})$ and $I_n(\boldsymbol{H}_{rn}\boldsymbol{p})$ are pixel values from the reference and neighboring views, respectively. The valid pixel set $\mathcal{V}$ is defined as:

$$\mathcal{V} = \{\boldsymbol{p} \mid A(\boldsymbol{p}) < \tau_h\} \cap \{\boldsymbol{p} \mid \mathcal{E}_{\text{proj}}(\boldsymbol{p}) < \tau_e\}, \tag{24}$$

ensuring optimization focuses on geometrically reliable regions with low uncertainty and small reprojection errors $\mathcal{E}_{\text{proj}}$.

Table 2: ***Quantitative Comparison on the Tanks and Temples Dataset [KPZK17].*** *We report the F1 score and average optimization time.Our method achieves F1 score similar to PGSR and Neurlangelo, while significantly outperforming the vast majority of methods in terms of training speed.*

	NeuS	Geo-NeuS	Neurlangelo	SuGaR	2D GS	GOF	PGSR	Ours
Barn	0.29	0.33	0.70	0.14	0.36	0.51	0.66	0.64
Caterpillar	0.29	0.26	0.36	0.16	0.23	0.41	0.41	0.41
Courthouse	0.17	0.12	0.28	0.08	0.13	0.28	0.21	0.18
Ignatius	0.83	0.72	0.89	0.33	0.44	0.68	0.80	0.79
Meetingroom	0.24	0.20	0.32	0.15	0.16	0.28	0.29	0.32
Truck	0.45	0.45	0.48	0.26	0.26	0.58	0.60	0.63
Mean	0.38	0.35	0.50	0.19	0.30	0.46	0.50	0.50
Time	>24h	>24h	>128h	2h	34.2m	2h	1.2h	1.06h

Table 3: ***Quantitative evaluation of novel view synthesis on Mip-NeRF360 [BMV*22] dataset.*** *Our method achieves competitive performance across all scenes.*

	Outdoor Scene			Indoor Scene		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NeRF [MST*21]	21.46	0.458	0.515	26.84	0.790	0.370
Deep Blending [HPP*18]	21.54	0.524	0.364	26.40	0.844	0.261
Instant NGP [MESK22]	22.90	0.566	0.371	29.15	0.880	0.216
MipNeRF360 [BMV*22]	24.47	0.691	0.283	31.72	0.917	0.180
3DGS [KKLD23]	24.24	0.705	0.283	30.99	0.926	0.199
SuGaR [GL24]	22.76	0.631	0.349	29.44	0.911	0.216
2DGS [HYC*24]	24.33	0.709	0.284	30.39	0.923	0.183
GOF [YSG24]	24.76	0.742	0.225	30.80	0.928	0.167
PGSR [CLY*24]	24.45	0.730	0.224	30.41	0.930	0.161
Ours	24.53	0.729	0.259	29.49	0.919	0.184

The total multi-view photometric consistency loss is:

$$\mathcal{L}_{\text{mvrgb}} = \mathcal{L}_{\text{rgb}} + \lambda_7 \mathcal{L}_{\text{NCC}}. \tag{25}$$

where the color loss $\mathcal{L}_{rgb}$ combined with SSIM terms use the exposure compensation defined in PGSR.

Total Loss The overall training loss combines multi-view photometric, geometric, and uncertainty-aware losses:

$$\mathcal{L} = \mathcal{L}_{\text{mvrgb}} + \mathcal{L}_{MM} + \mathcal{L}_{\text{geo}}. \tag{26}$$

By integrating uncertainty-weighted normal regularization and photometric consistency constraints, our framework dynamically adjusts regularization weights to suppress unreliable regions, enhancing geometric accuracy and robustness against noise, blur, and textureless areas.

5. Experimments

We conduct comprehensive experiments of our uncertainty-aware 3D Gaussian Splatting method, focusing on geometric reconstruction, appearance modeling, and novel view synthesis tasks. First, we evaluate geometric and appearance reconstruction capabilities on multiple standard datasets against state-of-the-art methods. Ablation studies then validate the contributions of individual components.

Datasets. Our approach is evaluated on three benchmark datasets: DTU for surface reconstruction (49-69 images/scenes downsampled to 800Œ600 using COLMAP sparse points); Tanks and Temples for robustness of large-scale reconstruction (100-1000+ images in 6 scenes with provided camera poses); Mip-NeRF360 for novel view synthesis in complex real-world environments with varying lighting and scene complexity. All datasets follow established preprocessing protocols, including sparse point cloud generation and resolution adjustment to ensure fair comparisons with previous work.

Evaluation Metrics. Our evaluation follows the usual practice of quantifying surface reconstruction quality using chamfer distance on DTUs, F1 scores (the harmonic mean of precision and recall) on Tanks and Temples, and PSNR (global image fidelity), SSIM (structural similarity), and LPIPS [ZIE*18] on Mip-NeRF360 to evaluate novel view synthesis. The chamfer distance measures the geometric alignment between the reconstructed surface and the ground truth surface, the F1-score balances the completeness and accuracy of the reconstructed geometry, and the image metrics synthesize the pixel-level error, structural consistency, and human-perceived quality of the synthesized views.

Baselines. We compare against two categories of state-of-the-art methods: 1) Implicit neural surface reconstruction: NeRF [MST*21], NeuS [WLL*21], VolSDF [YGKL21], Neuralangelo [LME*23], Deep Blending [HPP*18], Instant-NGP [MESK22], MipNeRF 360 [BMV*22], and Geo-NeuS [FXOT22], which leverage SDF or density fields for volumetric rendering. 2) Explicit Gaussian Splatting-based reconstruction: 3DGS [KKLD23], SuGaR [GL24], 2DGS [HYC*24], GOF [YSG24], and PGSR [CLY*24], ensuring comprehensive comparison across scene representation paradigms.

5.1. Implementation Details

All experiments are conducted in a PyTorch 2.3.1 and CUDA 11.8 environment on NVIDIA RTX 4090 GPUs. Most hyperparameters align with the original 3D Gaussian Splatting [KKLD23], and we adopt the same AbsGS subdivision strategy [YLL*24] as PGSR. Depth and normal priors are generated using the Metric3D V2 model [HYZ*24], with default weights: $\lambda_1 = 0.0125$, $\lambda_2 = 0.03$, $\lambda_3 = 100$, $\lambda_4 = 0.01$, $\lambda_5 = 0.02$, $\lambda_6 = 0.03$, $\lambda_7 = 0.15$. During training, we use third-order spherical harmonic functions to model uncertainty, performing a total of 30k iteration steps. we first optimize 3D Gaussians using only the rendering loss for the initial 7,000 iterations. Depth and normal prior losses are introduced between iterations 7,000 and 9,000, with prior learning halted at iteration 9,000 to initiate uncertainty estimation. Similar to RGB learning in 3DGS, we increment the spherical harmonics (SH) degree every 1,000 iterations. To stabilize learning during SH degree transitions, L1 regularization is applied for 200 iterations before and after each degree change. Uncertainty loss weights follow an exponential decay schedule. To account for instability in early training stages, uncertainty-based normal prior filtering and guidance are activated only after iteration 15,000. Experimentally, thresholds are set as: depth uncertainty $\tau_d = 0.25$, normal uncertainty $\tau_n = 0.4$, and uncertainty-aware NCC $\tau_h = 0.3$.

5.2. Comparison

We first evaluate the geometric reconstruction capability of our method on the DTU dataset. As shown in Table 1, our model out-

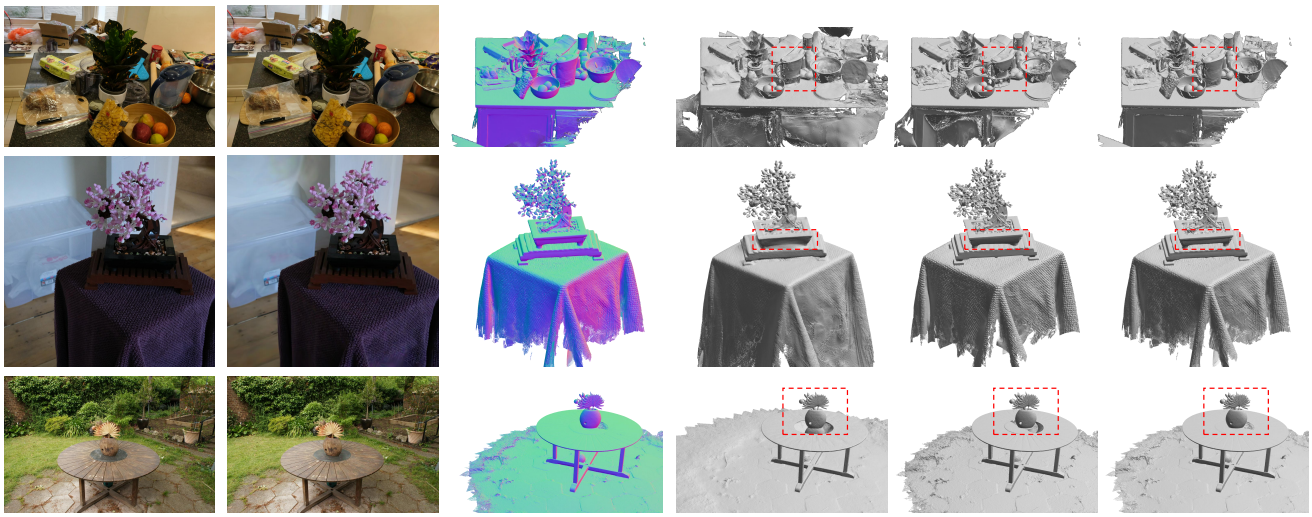

Figure 4: ***Qualitative comparison of reconstruction results for indoor and outdoor scenes in the MipNerf360 [BMV*22] dataset.*** *Our method reconstructs more geometrically accurate surfaces and finer details.*

performs traditional implicit methods in geometric accuracy while achieving 20Œ faster reconstruction speeds. Compared to 2DGS, GOF, and PGSR, our method delivers superior geometric reconstruction results. Figure 3 demonstrates our methods excellence in recovering fine geometric details.

Additionally, Table 2 presents results on the Tanks and Temples dataset, where our method achieves comparable F1-scores to PGSR while surpassing other reconstruction approaches. Our training time is also lower than PGSR. As shown in Figure 5, our method better handles textureless regions and reconstructs finer details compared to PGSR.

For novel view synthesis (NVS) evaluation on the Mip-NeRF360 dataset, Table 3 shows that our method outperforms most existing approaches. While slightly trailing PGSR in rendering quality, it maintains higher geometric accuracy, as visualized in Figure 4.

5.3. Ablations

In this section, we further analyze the impact of different components in our method, including the uncertainty module, normal regularization term, and uncertainty-aware NCC term, on the final reconstruction quality. Ablation studies on the DTU dataset demonstrate the contribution of each component, as listed in Table 4.

Impact of the Uncertainty Module We first evaluate the effectiveness of the uncertainty module. Experimental results in Table 4 show that removing all uncertainty guidance (baseline) degrades reconstruction quality (0.53 vs. 0.47 CD). This validates the necessity of explicit uncertainty modeling for resolving geometric ambiguities. Disabling the normal regularization term further reduces reconstruction accuracy due to erroneous normal orientations.

To illustrate that the improvement in the performance of our method is not all due to the addition of a prior knowledge, we added a prior knowledge to PGSR and compared it with our method, Figure 7 visualizes the reconstruction results for the bonsai and counter scenes in the MipNeRF360 dataset, where our method reconstructs a more correct geometric topology and finer geometric details. To validate uncertainty-aware NCC, we compare it against traditional NCC on the TNT Meetingroom scene. As shown in Figure 6, in weakly textured regions (e.g., white walls, tabletops), traditional NCC without uncertainty awareness converges to local minima, generating spurious planar structures.

Table 4: ***Ablation study on DTU dataset.***

Model setting	CD↓
Baseline (w/o all uncertainty guidance)	0.53
w/o Normal Regularization	0.50
PGSR + prior	0.49
Ours	0.47

6. Limitations and Future Work

Although our UA-GS method improves geometric reconstruction accuracy and completeness, several challenges remain. First, higher uncertainty persists around object boundaries, particularly in outdoor scenes with complex vegetation (e.g., grass, foliage), where intricate uncertainty maps complicate optimization. Generating accurate uncertainty maps during thousands of iterations remains challenging for reconstructing fine structures like leaves and grass. Second, while the AbsGS subdivision strategy [YLL*24] effectively captures fine details, it may lead to exponential growth in Gaussian primitives. Future work could explore uncertainty-aware subdivision, prioritizing high-uncertainty regions to balance preci-

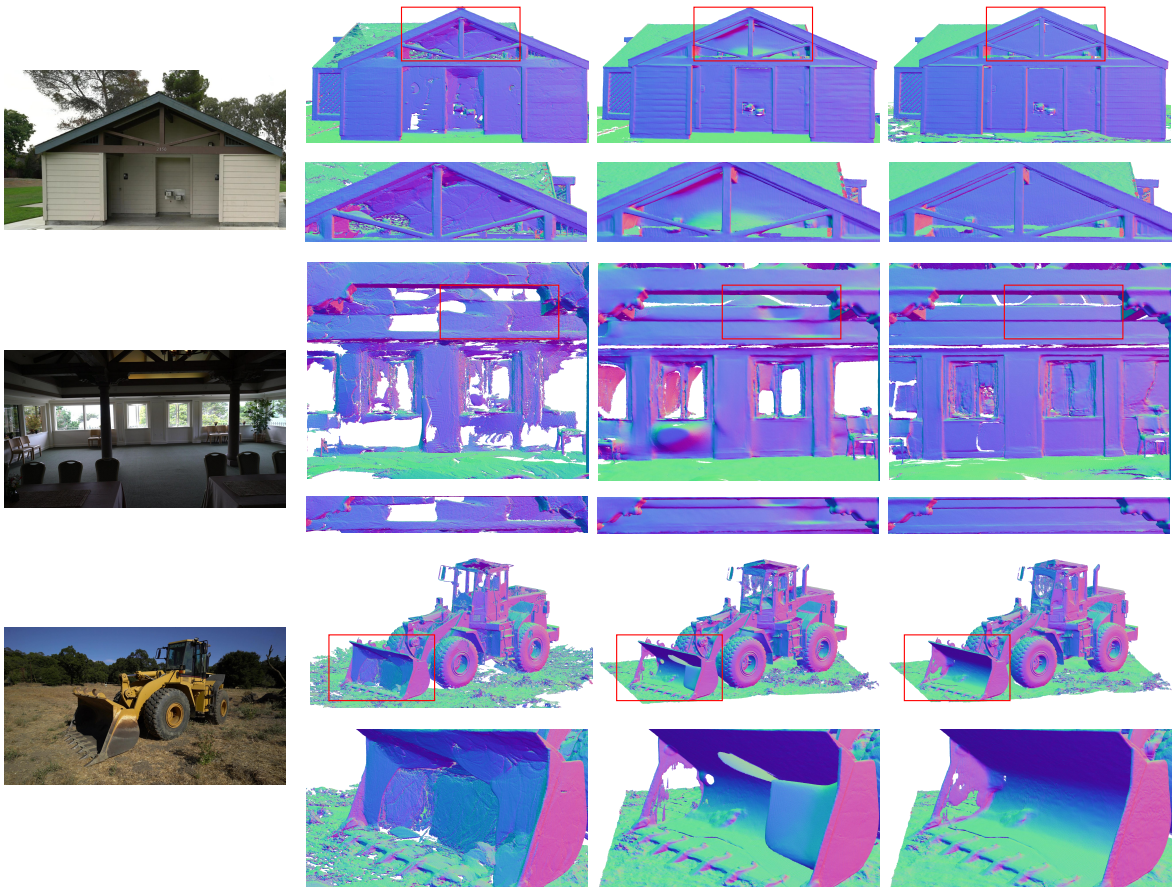

Figure 5: ***Qualitative comparison on TNT dataset [KPZK17].*** *We show normal maps of the reconstructed meshes from our method, 2DGS, and PGSR, as well as ground truth RGB images. Our method can reconstruct more accurate topological structures and yield more complete reconstruction results in textureless regions.*

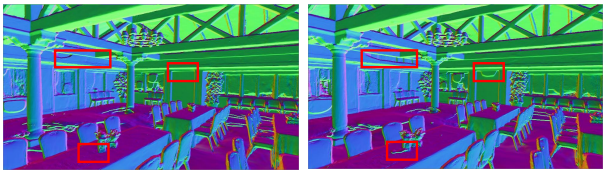

Figure 6: ***Visualizations comparison of uncertainty-aware NCC (ours) vs. traditional NCC in normal maps.*** *Uncertainty guidance mitigates artifacts in textureless regions.*

sion and computational efficiency. Finally, despite introducing geometric priors to reduce mesh artifacts, incomplete surfaces may still occur. In textureless regions, high depth uncertainty can cause erroneous Gaussian stacking, which may be mitigated by increasing the weight of depth uncertainty regularization.

7. Conclusion

We propose an uncertainty-aware 3D Gaussian Splatting framework that demonstrates state-of-the-art performance in geometric reconstruction and novel view synthesis, outperforming existing implicit and explicit reconstruction methods. By integrating depth and normal uncertainty regularization, spherical harmonics-based view-dependent modeling, and optimized mesh extraction strategies, our method achieves breakthroughs in reconstruction quality while maintaining computational efficiency. It enables large-scale, high-fidelity 3D reconstruction of complex scenes with practical applicability.

References

[AAHL20] ANTORÁN J., ALLINGHAM J., HERNÁNDEZ-LOBATO J. M.: Depth uncertainty in neural networks. *Advances in neural information processing systems 33* (2020), 10620–10634. 3

[BBC21] BAE G., BUDVYTIS I., CIPOLLA R.: Estimating and exploiting the aleatoric uncertainty in surface normal estimation. In *Proceedings*

of the IEEE/CVF International Conference on Computer Vision (2021), pp. 13137–13146. 3

[BD24] BAE G., DAVISON A. J.: Rethinking inductive biases for surface normal estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 9535–9545. 1

[BMV*22] BARRON J. T., MILDENHALL B., VERBIN D., SRINIVASAN P. P., HEDMAN P.: Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), pp. 5470–5479. 8, 9

[CLL23] CHEN H., LI C., LEE G. H.: Neusg: Neural implicit surface reconstruction with 3d gaussian splatting guidance. *arXiv preprint arXiv:2312.00846* (2023). 2, 7

[CLS*23] CHEN Z., LI Z., SONG L., CHEN L., YU J., YUAN J., XU Y.: Neurbf: A neural fields representation with adaptive radial basis functions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2023), pp. 4182–4194. 1, 2

[CLY*24] 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. *arXiv preprint arXiv:2406.06521* (2024). 1, 2, 7, 8

[CWL*24] CHEN H., WEI F., LI C., HUANG T., WANG Y., LEE G. H.: Vcr-gaus: View consistent depth-normal regularizer for gaussian surface reconstruction. *arXiv preprint arXiv:2406.05774* (2024). 3

[CXG*22] CHEN A., XU Z., GEIGER A., YU J., SU H.: Tensorf: Tensorial radiance fields. In *European conference on computer vision* (2022), Springer, pp. 333–350. 2

[CXZ*20] CHENG S., XU Z., ZHU S., LI Z., LI L. E., RAMAMOORTHI R., SU H.: Deep stereo using adaptive thin volume representation with uncertainty awareness. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 2524–2534. 3

[DHY*22] DENG N., HE Z., YE J., DUINKHARJAV B., CHAKRAVARTHULA P., YANG X., SUN Q.: Fov-nerf: Foveated neural radiance fields for virtual reality. *IEEE Transactions on Visualization and Computer Graphics 28*, 11 (2022), 3854–3864. 1

[DLW24] DU K., LIANG Z., WANG Z.: Gs-id: Illumination decomposition on gaussian splatting via diffusion prior and parametric light source optimization. *arXiv preprint arXiv:2408.08524* (2024). 2

[EFHP20] ELDESOKEY A., FELSBERG M., HOLMQUIST K., PERSSON M.: Uncertainty-aware cnns for depth completion: Uncertainty from beginning to end. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 12014–12023. 3

[EPF14] EIGEN D., PUHRSCH C., FERGUS R.: Depth map prediction from a single image using a multi-scale deep network. *Advances in neural information processing systems 27* (2014). 4

[ESMZ21] EFTEKHAR A., SAX A., MALIK J., ZAMIR A.: Omnidata: A scalable pipeline for making multi-task mid-level vision datasets from 3d scans. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 10786–10796. 6

[FKYT*22] FRIDOVICH-KEIL S., YU A., TANCIK M., CHEN Q., RECHT B., KANAZAWA A.: Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), pp. 5501–5510. 2

[FXOT22] FU Q., XU Q., ONG Y. S., TAO W.: Geo-neus: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. *Advances in Neural Information Processing Systems 35* (2022), 3403–3416. 1, 2, 8

[GCS06] GOESELE M., CURLESS B., SEITZ S. M.: Multi-view stereo revisited. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (2006), vol. 2, IEEE, pp. 2402–2409. 1

[GKJ*21] GARBIN S. J., KOWALSKI M., JOHNSON M., SHOTTON J., VALENTIN J.: Fastnerf: High-fidelity neural rendering at 200fps. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 14346–14355. 2

[GL24] GUÉDON A., LEPETIT V.: Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 5354–5363. 2, 7, 8

[HPP*18] HEDMAN P., PHILIP J., PRICE T., FRAHM J.-M., DRETTAKIS G., BROSTOW G.: Deep blending for free-viewpoint image-based rendering. *ACM Transactions on Graphics (ToG) 37*, 6 (2018), 1–15. 8

[HSM*21] HEDMAN P., SRINIVASAN P. P., MILDENHALL B., BARRON J. T., DEBEVEC P.: Baking neural radiance fields for real-time view synthesis. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 5875–5884. 2

[HYC*24] HUANG B., YU Z., CHEN A., GEIGER A., GAO S.: 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers* (2024), pp. 1–11. 2, 7, 8

[HYZ*24] HU M., YIN W., ZHANG C., CAI Z., LONG X., CHEN H., WANG K., YU G., SHEN C., SHEN S.: Metric3d v2: A versatile monocular geometric foundation model for zero-shot metric depth and surface normal estimation. *arXiv preprint arXiv:2404.15506* (2024). 4, 6, 8

[JDV*14] JENSEN R., DAHL A., VOGIATZIS G., TOLA E., AANÆS H.: Large scale multi-view stereopsis evaluation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2014), pp. 406–413. 7

[KKLD23] KERBL B., KOPANAS G., LEIMKÜHLER T., DRETTAKIS G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph. 42*, 4 (2023), 139–1. 1, 2, 8

[KKO*22] KAYA B., KUMAR S., OLIVEIRA C., FERRARI V., VAN GOOL L.: Uncertainty-aware deep multi-view photometric stereo. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), pp. 12601–12611. 3

[KPZK17] KNAPITSCH A., PARK J., ZHOU Q.-Y., KOLTUN V.: Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics (ToG) 36*, 4 (2017), 1–13. 8, 10

[LGZL*20] LIU L., GU J., ZAW LIN K., CHUA T.-S., THEOBALT C.: Neural sparse voxel fields. *Advances in Neural Information Processing Systems 33* (2020), 15651–15663. 2

[LME*23] LI Z., MÜLLER T., EVANS A., TAYLOR R. H., UNBERATH M., LIU M.-Y., LIN C.-H.: Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 8456–8465. 1, 2, 7, 8

[LYZ*22] LI H., YANG X., ZHAI H., LIU Y., BAO H., ZHANG G.: Vox-surf: Voxel-based implicit surface representation. *IEEE Transactions on Visualization and Computer Graphics 30*, 3 (2022), 1743–1755. 2

[MESK22] MÜLLER T., EVANS A., SCHIED C., KELLER A.: Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG) 41*, 4 (2022), 1–15. 1, 2, 8

[MST*21] MILDENHALL B., SRINIVASAN P. P., TANCIK M., BARRON J. T., RAMAMOORTHI R., NG R.: Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM 65*, 1 (2021), 99–106. 1, 2, 7, 8

[NMOG20] NIEMEYER M., MESCHEDER L., OECHSLE M., GEIGER A.: Differentiable volumetric rendering: Learning implicit 3d representations without 3d supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 3504–3515. 2

[PLSH22] PAN X., LAI Z., SONG S., HUANG G.: Activenerf: Learning where to see with uncertainty estimation. In *European Conference on Computer Vision* (2022), Springer, pp. 230–246. 3

[QLT21] QU C., LIU W., TAYLOR C. J.: Bayesian deep basis fitting for depth completion with uncertainty. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 16147–16157. 3

[RPLG21] REISER C., PENG S., LIAO Y., GEIGER A.: Kilonerf: Speeding up neural radiance fields with thousands of tiny mlps. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 14335–14345. 2

[SAMNR22] SHEN J., AGUDO A., MORENO-NOGUER F., RUIZ A.: Conditional-flow nerf: Accurate 3d modelling with reliable uncertainty quantification. In *European Conference on Computer Vision* (2022), Springer, pp. 540–557. 3

[SXT22] SU W., XU Q., TAO W.: Uncertainty guided multi-view stereo network for depth estimation. *IEEE Transactions on Circuits and Systems for Video Technology 32*, 11 (2022), 7796–7808. 3

[TLH*24] TONDERSKI A., LINDSTRÖM C., HESS G., LJUNGBERGH W., SVENSSON L., PETERSSON C.: Neurad: Neural rendering for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 14895–14904. 1

[TLY*21] TAKIKAWA T., LITALIEN J., YIN K., KREIS K., LOOP C., NOWROUZEZAHRAI D., JACOBSON A., MCGUIRE M., FIDLER S.: Neural geometric level of detail: Real-time rendering with implicit 3d shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 11358–11367. 2

[TRM*25] TURKULAINEN M., REN X., MELEKHOV I., SEISKARI O., RAHTU E., KANNALA J.: Dn-splatter: Depth and normal priors for gaussian splatting and meshing. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (2025), IEEE, pp. 2421–2431. 7

[TRZ*23] TANG J., REN J., ZHOU H., LIU Z., ZENG G.: Dreamgaussian: Generative gaussian splatting for efficient 3d content creation. *arXiv preprint arXiv:2309.16653* (2023). 2

[WHH*23] WANG Y., HAN Q., HABERMANN M., DANIILIDIS K., THEOBALT C., LIU L.: Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2023), pp. 3295–3306. 1, 2

[WLL*21] WANG P., LIU L., LIU Y., THEOBALT C., KOMURA T., WANG W.: Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *arXiv preprint arXiv:2106.10689* (2021). 1, 2, 7, 8

[WPH*24] WIERSMA R., PHILIP J., HAŠAN M., MULLIA K., LUAN F., EISEMANN E., DESCHAINTRE V.: Fast and uncertainty-aware svbrdf recovery from multi-view capture using frequency domain analysis. *arXiv preprint arXiv:2406.17774* (2024). 3

[XXP*22] XU Q., XU Z., PHILIP J., BI S., SHU Z., SUNKAVALLI K., NEUMANN U.: Point-nerf: Point-based neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), pp. 5438–5448. 2

[XXYG24] XIAO Y., XU J., YU Z., GAO S.: Debsdf: Delving into the details and bias of neural indoor scene reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024). 1, 3, 5, 6, 7

[YCW*23] YANG Z., CHEN Y., WANG J., MANIVASAGAM S., MA W.-C., YANG A. J., URTASUN R.: Unisim: A neural closed-loop sensor simulator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 1389–1399. 1

[YGKL21] YARIV L., GU J., KASTEN Y., LIPMAN Y.: Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems 34* (2021), 4805–4815. 2, 7, 8

[YH09] YOO J.-C., HAN T. H.: Fast normalized cross-correlation. *Circuits, systems and signal processing 28* (2009), 819–843. 1

[YIL*23] YANG J., IVANOVIC B., LITANY O., WENG X., KIM S. W., LI B., CHE T., XU D., FIDLER S., PAVONE M., ET AL.: Emernerf: Emergent spatial-temporal scene decomposition via self-supervision. *arXiv preprint arXiv:2311.02077* (2023). 1

[YLL*24] YE Z., LI W., LIU S., QIAO P., DOU Y.: Absgs: Recovering fine details in 3d gaussian splatting. In *Proceedings of the 32nd ACM International Conference on Multimedia* (2024), pp. 1053–1061. 8, 9

[YLSY19] YIN W., LIU Y., SHEN C., YAN Y.: Enforcing geometric constraints of virtual normal for depth prediction. In *Proceedings of the IEEE/CVF international conference on computer vision* (2019), pp. 5684–5693. 1

[YLZ*24] YAN Y., LIN H., ZHOU C., WANG W., SUN H., ZHAN K., LANG X., ZHOU X., PENG S.: Street gaussians: Modeling dynamic urban scenes with gaussian splatting. In *European Conference on Computer Vision* (2024), Springer, pp. 156–173. 1

[YPN*22] YU Z., PENG S., NIEMEYER M., SATTLER T., GEIGER A.: Monosdf: Exploring monocular geometric cues for neural implicit surface reconstruction. *Advances in neural information processing systems 35* (2022), 25018–25032. 1, 7

[YSG24] YU Z., SATTLER T., GEIGER A.: Gaussian opacity fields: Efficient and compact surface reconstruction in unbounded scenes. *arXiv preprint arXiv:2404.10772* (2024). 2, 7, 8

[YZC*23] YIN W., ZHANG C., CHEN H., CAI Z., YU G., WANG K., CHEN X., SHEN C.: Metric3d: Towards zero-shot metric 3d prediction from a single image. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2023), pp. 9043–9053. 6

[ZFS*24] ZHANG B., FANG C., SHRESTHA R., LIANG Y., LONG X., TAN P.: Rade-gs: Rasterizing depth in gaussian splatting. *arXiv preprint arXiv:2406.01467* (2024). 2, 7

[ZIE*18] ZHANG R., ISOLA P., EFROS A. A., SHECHTMAN E., WANG O.: The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 586–595. 8

[ZLH24] ZHANG W., LIU Y.-S., HAN Z.: Neural signed distance function inference through splatting 3d gaussians pulled on zero-level set. *arXiv preprint arXiv:2410.14189* (2024). 2, 7

[ZLS*24] ZHOU X., LIN Z., SHAN X., WANG Y., SUN D., YANG M.-H.: Drivinggaussian: Composite gaussian splatting for surrounding dynamic autonomous driving scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 21634–21643. 1

[ZXL*24] ZHANG W., XIANG H., LIAO Z., LAI X., LI X., ZENG L.: 2dgs-room: Seed-guided 2d gaussian splatting with geometric constrains for high-fidelity indoor scene reconstruction. *arXiv preprint arXiv:2412.03428* (2024). 2

[ZYL*20] ZHANG J., YAO Y., LI S., LUO Z., FANG T.: Visibility-aware multi-view stereo network. *arXiv preprint arXiv:2008.07928* (2020). 3

[ZZD*24] ZHOU Z., ZHANG C., DONG Z., MARSHALL C., ZHAO S.: Estimating Uncertainty in Appearance Acquisition. In *Eurographics Symposium on Rendering* (2024), Haines E., Garces E., (Eds.), The Eurographics Association. doi:10.2312/sr.20241149. 3

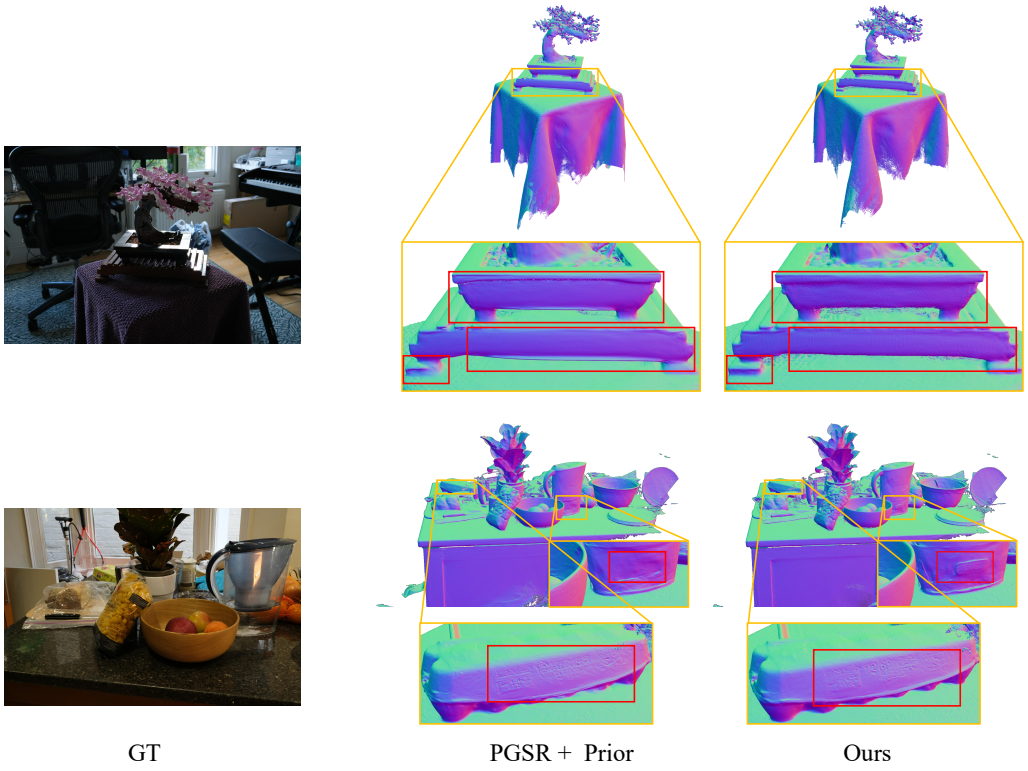

Figure 7: ***Comparison of adding our uncertainty perception vs. just adding a prior PGSR visualization in the normal map.*** *Adding uncertainty reconstructs a more correct geometric topology and finer details.*