GAGS: Gradient-Guided Adaptive Gaussian Splatting for Efficient and Geometry-Regularized Surface Reconstruction

Yongmao Hou, Tengfei Wang,* Xin Wang, Zongqian Zhan

Wuhan University, School of Geodesy and Geomatics, 430079 Wuhan, China PR
yongmaohou@whu.edu.cn, tf.wang@whu.edu.cn, xwang@sgg.whu.edu.cn, zqzhan@sgg.whu.edu.cn

Keywords: Gaussian Splatting, Surface Reconstruction, Adaptive Densification, Geometry Regularization.

Abstract

While 3D Gaussian Splatting (3DGS) has advanced surface reconstruction, existing implementations face critical challenges in memory efficiency and geometric fidelity. Current approaches like PGSR generate excessive Gaussian primitives due to uncontrolled densification, leading to redundant memory consumption while struggling with surface artifacts in boundary regions. This paper introduces GAGS (Gradient-Guided Adaptive Gaussian Splatting), a geometry-regularized framework that addresses these limitations through three technical contributions: photometric gradient-driven adaptive densification that strategically controls primitive subdivision using image gradient analysis, anisotropy-aware shape regularization for adaptive Gaussian scale constraint, and a dual regularization mechanism combining normal self-smoothing with depth-aware correction. Evaluations on the DTU dataset demonstrate the framework's effectiveness in maintaining visual quality while significantly reducing redundant primitives—achieving an 83% reduction compared to PGSR baseline—with improved surface regularity in complex geometric regions. Project web:`https://3241674469.github.io/GAGS-project/`

1. Introduction

Surface reconstruction constitutes a fundamental task in computer vision with wide-ranging applications from robotic navigation to virtual reality systems. Recently, 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) has emerged as a breakthrough paradigm in surface reconstruction, offering fast optimization, explicit scene representation, and high-quality rendering capabilities. This innovative approach optimizes the parameters of anisotropic 3D Gaussian primitives and employs differentiable α-blending for view-consistent rendering, thereby enabling real-time reconstruction at high resolutions with impressive visual fidelity. Nevertheless, two critical limitations hinder its practical deployment: (1) excessive memory consumption from redundant primitives in geometrically complex regions, and (2) surface irregularities along object boundaries caused by unconstrained primitive distributions.

Recent advances in 3DGS-based surface reconstruction, exemplified by PGSR (Chen et al., 2024), propose an unbiased depth rendering framework derived from Gaussian distributions to address topological inconsistencies. This approach computes depth values through camera-to-plane distance measurements and contribution partitioning, while integrating multiview geometric regularization with single-view geometric priors to preserve global accuracy. The framework further incorporates photometric consistency constraints and camera exposure compensation to handle illumination variations. However, despite these advancements, PGSR inherits fundamental limitations from conventional 3DGS methodologies: the uncontrolled densification process generates redundant primitives in texture-deficient regions, and the lack of adaptive shape constraints leads to persistent boundary artifacts due to unregulated primitive distributions.

This paper presents GAGS (Gradient-Guided Adaptive Gaussian Splatting), a novel framework that systematically addresses

---
* Corresponding Author

these challenges through three key technical innovations:

1. **Photometric Gradient-Driven Adaptive Densification**. Unlike conventional heuristic densification strategies employed by 3DGS, GAGS leverages photometric gradients from real images to guide primitive densification. This adaptively suppresses unnecessary splits in flat/low-texture regions, reducing primitives up to 83% compared to PGSR.
2. **Anisotropy-Aware Gaussian Primitive Shape Regularization**. We impose adaptive shape constraints on Gaussian primitives, compressing their scales along tangential directions in flat regions to better align with thin surfaces while preserving geometric fidelity in detailed areas.
3. **Dual Regularization via Normal-Depth Consistency**. A normal self-smoothing loss and depth-aware normal correction module are jointly optimized to eliminate high-frequency noise. By refining primitive positions using depth-derived normals, our method generates structurally coherent surfaces without sacrificing rendering quality.

The remainder of this paper is organized as follows: Section 2 discusses related work. Section 3 provides an overview of the preliminary concepts of PGSR. The details of the proposed method are presented in Section 4. Section 5 reports the performance evaluation and experimental results of our work. Finally, conclusions and an outlook are drawn in Section 6.

2. Related Work

Surface reconstruction serves as a fundamental pillar in computer vision, with the goal of producing precise geometric representations from sparse or noisy input data. This paper focuses on efficient and geometry-regularized surface reconstruction based on PGSR (Chen et al., 2024). In this section, we explore three key research directions: 3D Gaussian compression, neural surface reconstruction, and Gaussian splatting-based surface reconstruction methods.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-1/W2-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China
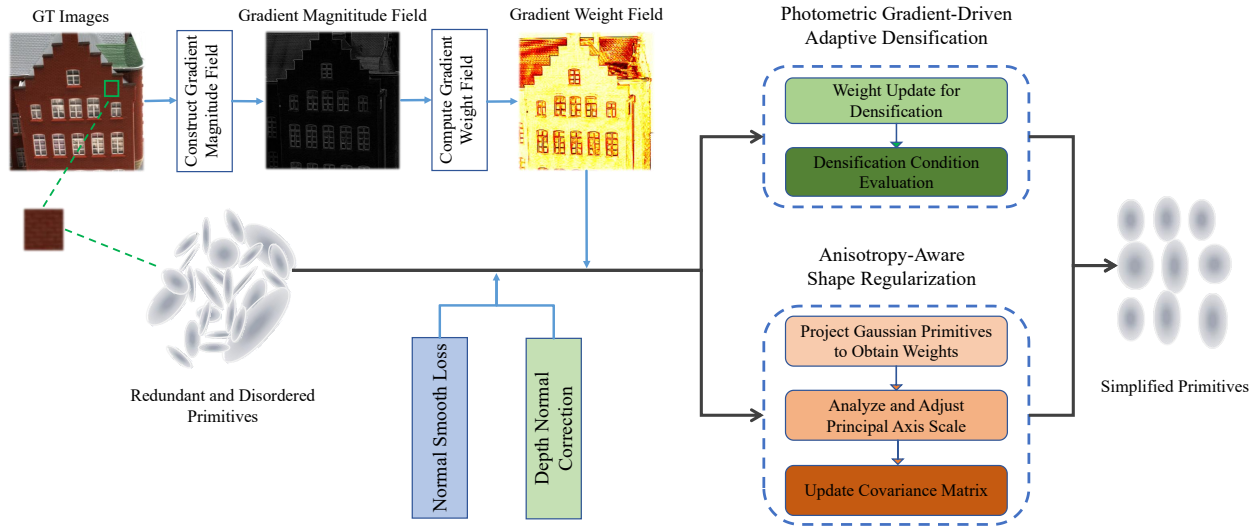
Figure 1. Overview of our GAGS. Guided by image gradients, GAGS effectively reduces redundant Gaussians and optimizes their distribution, leading to geometrically regularized surface reconstructions with minimized memory usage.

2.1 3D Gaussian Compression

Recent research in 3D Gaussian compression has prioritized balanced optimization of memory efficiency and computational performance while maintaining visual fidelity. Papantonakis et al. (2024) developed an adaptive compression framework integrating hierarchical pruning with parameter quantization, achieving substantial memory reduction while preserving rendering quality. Spatial optimization methods have emerged as effective solutions, with Wang et al. (2024) proposing dynamic radius adjustment to optimize Gaussian spatial distributions and eliminate redundant splats. Hybrid compression strategies demonstrate particular effectiveness, exemplified by Fan et al. (2024b)'s entropy-coded parameterization that combines novel encoding schemes with real-time rendering capabilities.

In parameter representation and scene scalability, Lee et al. (2024) introduced a decomposed encoding method separating geometric attributes from appearance features through hybrid spherical-harmonic/hash-grid integration. For large-scale environment processing, Liu et al. (2025) implemented multi-resolution clustering with view-dependent detail control. Complementary approaches like Fang and Wang (2024)'s dynamic primitive allocation and Zhang et al. (2024)'s gradient-aware density modulation further enhance compression through geometric constraints and perceptual optimization. These methodologies collectively advance 3D Gaussian splatting practicality through spatial redundancy reduction, parametric efficiency improvement, and perceptually-guided simplification.

2.2 Neural Surface Reconstruction

Recent advances in neural surface reconstruction have demonstrated the capability of deep learning to overcome fundamental limitations of conventional geometry processing techniques, particularly in managing sparse input data and reconstructing objects with intricate topological structures. Modern approaches can be broadly categorized into explicit and implicit representation paradigms, each presenting unique advantages and limitations in terms of reconstruction accuracy, computational efficiency, and practical applicability.

Explicit Surface Reconstruction: Explicit methods construct discrete geometric representations through structured formats such as voxel grids, point clouds, and polygonal meshes. The seminal work by Choy et al. (2016) proposed 3D-R2N2, a recurrent neural network architecture for multi-view 3D reconstruction using volumetric representations. Subsequent innovations improved scalability through hierarchical learning frameworks like PointNet++ (Qi et al., 2017), which processes unstructured point clouds via spatial partitioning strategies. Mesh-based approaches achieved notable progress with techniques such as Pixel2Mesh (Wang et al., 2018), employing graph convolutional networks (GCNs) to deform template meshes under 2D image guidance. While these explicit representations facilitate direct integration with downstream applications (e.g., CAD/CAM systems), they face inherent limitations: voxel representations scale cubically with resolution, point clouds lack inherent connectivity information, and mesh deformation approaches require careful initialization to avoid topological artifacts.

Implicit Surface Reconstruction: Implicit methods model surfaces as continuous functions represented by neural networks, typically using signed distance functions (SDFs) or occupancy fields. Park et al. (2019) pioneered this approach with DeepSDF, demonstrating continuous shape representation through coordinate-based MLPs. Parallel work by Mescheder et al. (2019) established Occupancy Networks for probabilistic surface estimation. The field experienced a transformative shift with Neural Radiance Fields (NeRF) (Mildenhall et al., 2021), which unified geometry and appearance learning through differentiable volume rendering. Subsequent advancements, such as VolSDF (Yariv et al., 2021) integrated SDF constraints with radiance field formulations to produce watertight surfaces. While demonstrating remarkable capability in recovering intricate geometric details, these approaches typically require dense view sampling and suffer from high computational costs during surface extraction.

Hybridization Strategies: Contemporary research focuses on combining the strengths of both paradigms. Explicit methods offer computational efficiency and direct editability, while implicit representations excel at handling topological complexity

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-1/W2-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China

and providing continuous surface descriptions. Notable hybrid approaches include Neural Parts (Deng et al., 2020), which integrates mesh-based components with neural implicit functions to preserve both interpretability and reconstruction fidelity. Such hybrid architectures demonstrate promising potential for applications requiring both geometric precision and practical usability.

2.3 Gaussian Splatting based Surface Reconstruction

Recent advancements in 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) have expanded its utility beyond radiance field rendering to explicit surface reconstruction. SuGaR (Guédon and Lepetit, 2024) pioneered mesh extraction through regularization terms that align Gaussians with scene surfaces. While their Poisson reconstruction-based approach improves geometric accuracy, inherent limitations persist: irregular Gaussian distributions cause surface fragmentation, particularly in low-texture regions, and discrete sampling leads to incomplete geometry reconstruction. Alternative representations attempt to address geometric consistency. 2DGS (Huang et al., 2024) projects 3D Gaussians onto view-consistent 2D manifolds, enhancing multi-view coherence at the cost of depth estimation bias through planar approximations. GOF (Yu et al., 2024) constructs Gaussian opacity fields for surface extraction via level-set identification, though global geometric consistency remains challenging in complex scenes.

The state-of-the-art PGSR (Chen et al., 2024) introduces planar-aligned Gaussian primitives that explicitly encode surface parameters (normals, origin distances), enabling analytic projection for unbiased depth estimation. Through multi-view geometric regularization, PGSR achieves the highest geometric reconstruction accuracy and rendering quality compared to the current state-of-the-art methods. However, two critical limitations emerge: (1) *memory inefficiency* from redundant primitives in homogeneous regions due to conservative density thresholds, and (2) *boundary artifacts* caused by sparse sampling and unconstrained Gaussian overlaps near occlusion boundaries.

To address these limitations, this work proposes Gradient-Guided Adaptive Gaussian Splatting (GAGS), which introduces photometric gradient-driven densification to adaptively suppress redundant primitive splits in homogeneous regions through image-derived gradient modulation. Anisotropic shape constraints are incorporated to compress Gaussians along planar tangents for thin-surface alignment while preserving isotropic fidelity in detailed areas. Additionally, dual normal-depth regularization is integrated to eliminate boundary artifacts by jointly applying self-smoothing and depth-aware correction.

3. Preliminary of PGSR

Building upon the foundation of 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023), PGSR (Chen et al., 2024) introduces a structured framework for high-fidelity surface reconstruction through three synergistic technical innovations. The methodology establishes geometric-aware constraints while preserving the efficient differentiable rendering pipeline of 3DGS.

3.1 Planar-Aligned Gaussian Representation

The core geometric representation evolves from volumetric 3D Gaussians to surface-oriented planar primitives. Each Gaussian G_i undergoes anisotropic compression along its minimum scaling axis, enforced through a scale-aware regularization term:

$$\mathcal{L}_s = \|\min(s_1, s_2, s_3)\|_1 \quad \text{where } \boldsymbol{S}_i = \text{diag}(s_1, s_2, s_3) \quad (1)$$

This operation effectively collapses 3D ellipsoids into planar elements whose orientation aligns with the surface normal:

$$\boldsymbol{n}_i = \boldsymbol{R}_i[:, k] \quad \text{where } k = \arg\min_j s_j \quad (2)$$

The planar compression enables explicit surface parameterization while maintaining the differentiable rendering properties of original 3DGS.

3.2 Unbiased Depth Rendering

To address the inherent depth bias in conventional Gaussian splatting, PGSR reformulates the rendering pipeline through plane-constrained ray casting. For each pixel $\boldsymbol{p} = (u, v)$, the rendering process first aggregates planar parameters via alpha blending:

$$D = \frac{\sum_i \alpha_i d_i}{\sum_i \alpha_i}, \quad \boldsymbol{N} = \frac{\sum_i \alpha_i \boldsymbol{n}_i}{\sum_i \alpha_i} \quad (3)$$

$$d_i = (\boldsymbol{R}_c^\top (\boldsymbol{\mu}_i - \boldsymbol{T}_c))^\top \boldsymbol{n}_i \quad (4)$$

where d_i denotes the signed distance from camera center $\boldsymbol{T}_c$ to the Gaussian plane. The final depth map $\boldsymbol{Z}$ is analytically derived through planar-ray intersection:

$$z(\boldsymbol{p}) = \frac{D}{\boldsymbol{N}^\top (\boldsymbol{K}^{-1} \tilde{\boldsymbol{p}})} \quad \text{with } \tilde{\boldsymbol{p}} = [u, v, 1]^\top \quad (5)$$

This formulation eliminates cumulative blending errors by enforcing physical consistency between planar orientation and depth estimation.

3.3 Geometric Regularization

The geometric constraints form a multi-scale consistency framework bridging local surface smoothness and global structure coherence.

The *single-view consistency* term leverages image gradients to align rendered normals with local geometry:

$$\mathcal{L}_{sv} = \sum_{\boldsymbol{p}} \|\nabla I(\boldsymbol{p})\| \cdot \|\boldsymbol{N}_r(\boldsymbol{p}) - \boldsymbol{N}_c(\boldsymbol{p})\|_2 \quad (6)$$

where $\boldsymbol{N}_r$ denotes the rendered normal and $\boldsymbol{N}_c$ the geometry-derived normal from depth variations.

For cross-view consistency, the *multi-view photometric* constraint employs normalized cross-correlation (NCC) to enforce appearance consistency under planar warping:

$$\mathcal{L}_{mv}^{ph} = \sum_{\boldsymbol{p}} (1 - \text{NCC}(I_r(\boldsymbol{p}), I_n(H_{rn}\boldsymbol{p}))) \quad (7)$$

Simultaneously, the *multi-view geometric* regularization ensures coherent planar projections across viewpoints:

$$\mathcal{L}_{mv}^{geo} = \sum_{\boldsymbol{p}} \|\boldsymbol{p}_r - H_{nr}H_{rn}\boldsymbol{p}_r\|_2 \quad (8)$$

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-1/W2-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China

These complementary constraints jointly suppress the Gaussian splatting's tendency towards local minima while preserving high-frequency geometric details.

3.4 Exposure-Aware Optimization

To handle real-world illumination variations, PGSR introduces per-image exposure parameters that decouple brightness changes from geometric reconstruction:

$$\mathcal{L}_{exp} = \| \exp(a_i)I_r + b_i - I_{gt} \|_1 + \lambda \text{SSIM}(\exp(a_i)I_r, I_{gt}) \quad (9)$$

This formulation allows simultaneous recovery of scene albedo and transient illumination effects.

The complete optimization objective integrates all components through balanced weighting:

$$\mathcal{L} = \mathcal{L}_{rgb} + 100\mathcal{L}_s + 0.01\mathcal{L}_{sv} + 0.2\mathcal{L}_{mv}^{ph} + 0.05\mathcal{L}_{mv}^{geo} \quad (10)$$

This formulation achieves a Chamfer distance of 0.49 mm on the DTU dataset (Jensen et al., 2014a) while maintaining a rendering speed of 30 FPS. It outperforms 3DGS-based baselines in terms of geometric accuracy.

4. Method

Our goal is to reduce redundancy in Gaussian primitives, particularly in flat 3D scene regions, guided by image gradients while regularizing geometry. To achieve this, we propose GAGS (Gradient-Guided Adaptive Gaussian Splatting), an advanced approach built upon the PGSR (Chen et al., 2024) framework. GAGS exploits photometric gradients extracted from real images to guide primitive densification, thereby significantly reducing the number of redundant Gaussian primitives. Moreover, GAGS imposes adaptive shape constraints on these primitives by compressing their scales along tangential directions in flat regions, resulting in better alignment with thin surfaces and improved geometric fidelity of the reconstructed scene. Finally, a normal self-smoothing loss and a depth-aware normal correction module are jointly optimized to suppress high-frequency noise, refine primitive positions using depth-derived normals, and enhance the overall robustness of 3D scene surface reconstruction. The overview of our GAGS is illustrated in Figure 1, and the following sections provide a detailed explanation of our approach.

4.1 Gradient Feature Extraction

The gradient weighting mechanism establishes photometric-geometric correspondence through continuous image analysis and primitive association. Beginning with gradient field computation, we employ central difference operators to capture horizontal and vertical intensity variations:

$$G_x = \frac{\partial I}{\partial x}, \quad G_y = \frac{\partial I}{\partial y} \quad (11)$$

The resultant gradient magnitude field M, calculated through Euclidean norm $M(p) = \sqrt{G_x(p)^2 + G_y(p)^2}$, encodes local texture complexity. To ensure cross-scene consistency, this magnitude field undergoes max-min normalization followed by quintic scaling:

$$\omega(p) = \left(\frac{M(p) - M_{\min}}{M_{\max} - M_{\min}} \right)^5 \quad (12)$$

This non-linear transformation amplifies weights in high-frequency regions by 5 orders of magnitude while suppressing flat areas, creating a photometric sensitivity map. Each Gaussian primitive G_i then associates with specific weights through screen-space projection:

$$\omega_i = \omega \left(\text{clamp}(\lfloor x_i \rceil, 0, W-1), \; \text{clamp}(\lfloor y_i \rceil, 0, H-1) \right) \quad (13)$$

where (x_i, y_i) denotes the primitive's projected coordinates. The coordinate clamping ensures valid texture sampling and spatial coherence. This weight association bridges image gradients with 3D primitive density requirements - high ω_i values demand precise representation in textured regions while permitting sparsity in homogeneous areas.

4.2 Photometric Gradient-Driven Adaptive Densification

Traditional densification approaches often generate redundant primitives in homogeneous regions due to uniform gradient thresholds. Our adaptive mechanism establishes gradient-aware primitive management through three interconnected operations: accumulated gradient modulation, anisotropic splitting criteria, and geometric consistency filtering.

The process begins with dual-gradient accumulation that combines photometric weights from Sec. 4.1 with geometric gradients:

$$\begin{aligned} \mathcal{A}(G_i) &\propto \|\nabla G_i\| \cdot \omega_i \\ \mathcal{A}_{\text{abs}}(G_i) &\propto \|\nabla G_i\| \cdot |\omega_i| \end{aligned} \quad (14)$$

Normalized gradients $\mathcal{A}/\mathcal{D}$ drive splitting decisions through an anisotropy-modulated threshold:

$$\frac{\mathcal{A}(G_i)}{\mathcal{D}(G_i)} > \tau_s \left(1 + \frac{\|s_i\|_1}{\|s_i\|_\infty} \right) \quad (15)$$

where the scaling ratio $\|s_i\|_1/\|s_i\|_\infty$ penalizes anisotropic primitives. Concurrently, absolute gradient accumulation $\mathcal{A}_{\text{abs}}/\mathcal{D}_{\text{abs}}$ controls cloning frequency to maintain primitive density in critical regions.

Geometric pruning enforces compactness through multi-criteria filtering:

$$\alpha_i < \tau_\alpha \; \vee \; r_i > \tau_r \; \vee \; \|s_i\|_\infty > 0.1 \cdot \text{extent} \quad (16)$$

removing under-contributing, oversized, or degenerate primitives. The complete pipeline (Algorithm 1) periodically executes these operations, achieving:

- Edge-sensitive refinement through gradient-weight coupling

- Anisotropy-aware splitting preventing over-flattening

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-1/W2-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China

- Adaptive density control across texture variations

Algorithm 1 Adaptive Densification Pipeline

Require: Current Gaussian set $\mathcal{G}$, gradient weights $\{\omega_i\}$
Ensure: Optimized Gaussian set $\mathcal{G}'$
  1: Update accumulators via Eq.(1-2)
  2: **if** iteration % update_interval == 0 **then**
  3:     Normalize gradients: $\tilde{g}_i \leftarrow \mathcal{A}(G_i)/\mathcal{D}(G_i)$
  4:     Split primitives satisfying Eq.(3)
  5:     Clone primitives meeting Eq.(4)
  6:     Prune using Eq.(5)
  7: **end if**

4.3 Anisotropy-Aware Shape Regularization

To align Gaussian primitives with underlying surface geometry while maintaining rendering stability, we develop an adaptive scale adjustment mechanism based on photometric gradients. The process begins by identifying each primitive's minimum scaling component $s_j^{\min} = \min(s_{j1}, s_{j2}, s_{j3})$, which corresponds to the principal compression direction. This minimum scale is then modulated according to local texture complexity through gradient-weighted adjustment:

$$\hat{s}_j^{\min} = \max(\tau_m, s_j^{\min}(1 - \alpha\omega_j)) \tag{17}$$

where ω_j from Sec. 4.1 controls compression strength ($\alpha = 0.8$), and $\tau_m = 10^{-4}$ prevents numerical instability. The adjustment creates geometry-aware deformations - strong compression ($\omega_j \to 0$) in flat regions produces planar-aligned Gaussians, while minimal scaling ($\omega_j \to 1$) preserves isotropic shapes in textured areas. The updated scale parameter is propagated through covariance matrix updates using scatter operations, maintaining differentiability during optimization.

This adaptive mechanism achieves dual objectives through gradient coupling: (1) geometric compliance by enforcing surface-aligned flattening in homogeneous regions, and (2) rendering stability via minimum scale constraints. The parallel implementation (Algorithm 2) ensures efficient scale adjustment across all visible primitives while preserving gradient flow for end-to-end training.

Algorithm 2 Adaptive Scale Adjustment

Require: Visible Gaussians $\{G_j\}$, gradient weights $\{\omega_j\}$
Ensure: Adjusted scaling parameters $\{\hat{s}_j\}$
  1: **for** each G_j in parallel **do**
  2:     Find $k = \arg\min(s_{j1}, s_{j2}, s_{j3})$
  3:     Compute $\hat{s}_{jk} = \max(\tau_m, s_{jk}(1 - \alpha\omega_j))$
  4:     Update $s_{jk} \leftarrow \hat{s}_{jk}$
  5: **end for**

4.4 Dual Regularization via Normal-Depth Consistency

To enhance surface continuity and geometric accuracy, we introduce dual regularization that combines normal field smoothing with depth-derived geometric constraints. The joint formulation addresses high-frequency noise while enforcing consistency between view-dependent normals and geometric surface orientation.

The first component applies gradient-weighted normal smoothing to suppress artifacts caused by primitive misalignment:

$$\mathcal{L}_{\text{smooth}} = \mathbb{E}_p\left[\omega(p)\|\nabla n(p)\|_2\right] \tag{18}$$

where the expectation is computed over all pixels Ω, with $\omega(p)$ from Sec. 4.1 emphasizing edges. Finite differences compute normal gradients $\nabla n(p)$, while gradient-guided depth averaging preserves geometric features during preprocessing.

The second constraint aligns rendered normals $n(p)$ with geometry-derived orientations $n_d(p)$:

$$n_d(p) = \frac{(-\partial_x D, -\partial_y D, 1)^\top}{\sqrt{(\partial_x D)^2 + (\partial_y D)^2 + \epsilon}} \tag{19}$$

where depth derivatives $\partial_x D, \partial_y D$ are computed from smoothed depth maps. The alignment loss:

$$\mathcal{L}_{\text{normal}} = \mathbb{E}_p\left[\omega(p)\|n(p) - n_d(p)\|_2\right] \tag{20}$$

penalizes deviations between photometric and geometric normals. The combined loss $\mathcal{L} = \mathcal{L}_{\text{smooth}} + 0.5\mathcal{L}_{\text{normal}}$ balances surface smoothness with geometric faithfulness, where the 0.5 weighting prevents over-constraining from depth noise.

5. Experiments

We conducted a comprehensive evaluation of GAGS by comparing its surface reconstruction performance with state-of-the-art methods. Furthermore, ablation studies were performed to systematically examine the effectiveness of three critical components in our proposed framework.

5.1 Datasets

To validate the effectiveness of our approach, we conducted experiments on 6 object-centric scenes randomly selected from the DTU dataset Jensen et al. (2014b). This benchmark dataset provides high-quality multi-view scans with precise camera calibration and ground truth point clouds for accurate reconstruction quality assessment under various lighting conditions and material properties.

5.2 Implementation Details

The training strategy and hyperparameters are maintained consistent with PGSR (Chen et al., 2024), with all scenes trained for 30,000 iterations. A Photometric Gradient-Driven Adaptive Densification strategy is employed for progressive densification, performed every 300 iterations until iteration 15,000. Depth maps are rendered for each training view, followed by TSDF field generation through the volumetric fusion algorithm (Curless and Levoy, 1996). Final meshes are extracted from the reconstructed TSDF volumes. All experiments were conducted on an NVIDIA RTX 4090 GPU.
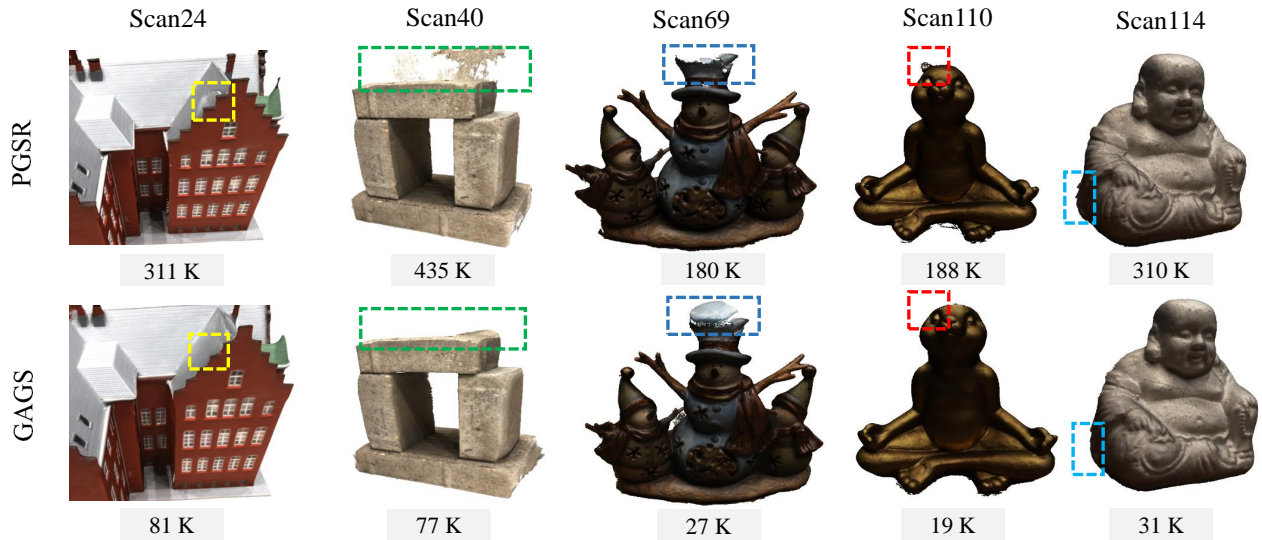

Figure 2. Qualitative comparison results. When comparing the reconstruction results of GAGS with PGSR on the DTU dataset, our GAGS achieves more robust reconstructions with fewer Gaussian primitives.

Table 1. Comprehensive Performance Analysis of 3D Reconstruction Methods on the DTU Dataset

Scene	Method	Gaussian ↓ (k)	Chamfer ↓ Distance
scan24	2DGS	239	0.123
	GOF	569	0.112
	PGSR	311	0.100
	GAGS	81	0.085
scan40	2DGS	297	0.118
	GOF	750	0.108
	PGSR	435	0.103
	GAGS	77	0.101
scan69	2DGS	138	0.108
	GOF	320	0.119
	PGSR	180	0.090
	GAGS	27	0.100
scan97	2DGS	168	0.113
	GOF	461	0.095
	PGSR	291	0.102
	GAGS	53	0.082
scan110	2DGS	66	0.106
	GOF	133	0.108
	PGSR	188	0.115
	GAGS	19	0.112
scan114	2DGS	103	0.120
	GOF	399	0.119
	PGSR	310	0.124
	GAGS	31	0.112
Average			
	2DGS	179	0.115
	GOF	411	0.110
	PGSR	294	0.106
	GAGS	48	0.099

Bold Red: Best Chamfer Distance in scene (lower is better)
Blue: Second best Chamfer Distance
Green: Lowest Gaussian count in scene (lower is better)
Orange: Second lowest Gaussian count
↓: Lower is better

5.3 Surface Reconstruction Evaluation

Our evaluation framework rigorously compares GAGS against state-of-the-art 3DGS variants, including PGSR (Chen et al., 2024), 2DGS (Huang et al., 2024), and GOF (Yu et al., 2024). Inspired by TrimGS (Fan et al., 2024a), this work proposes a dual-metric paradigm that quantifies reconstruction efficiency through Gaussian primitive counts while assessing geometric accuracy using Chamfer Distance (CD) between reconstructed and ground-truth point clouds. This geometric evaluation strategy directly operates on point clouds, effectively circumventing quantization artifacts inherent in mesh-based approaches while maintaining sensitivity to outlier distributions.

Following the standardized DTU evaluation protocol (Jensen et al., 2014b), the assessment pipeline comprises four critical stages. Initial spatial normalization eliminates scale discrepancies across reconstructions. Coarse-to-fine registration is then performed using RANSAC (Fischler and Bolles, 1981) for global alignment and ICP (Besl and McKay, 1992) for local refinement. To ensure density-invariant comparison, normalized point clouds undergo uniform voxel downsampling that retains only the centroid-closest points within each volumetric unit, effectively mitigating density-induced biases. Final metric computation focuses on Chamfer Distance to quantify surface deviation accuracy.

Quantitative results in Table 1 demonstrate GAGS' superior performance, achieving 6.6% lower CD than PGSR with 83.7% fewer Gaussians. Visual comparisons in Figure 2 further validate these findings, showing enhanced geometric regularity in GAGS reconstructions, particularly evident in thin structures where conventional methods exhibit over-densification artifacts. The sparser yet more accurate primitive distributions confirm our method's effectiveness in balancing reconstruction efficiency with geometric fidelity.

5.4 Ablation Study

To rigorously evaluate the contribution of each proposed component in GAGS, we conduct controlled ablation studies on DTU Scan114 under identical experimental settings. We sequentially disable individual modules while maintaining other

Full GAGS   w/o Photometric Gradient-Driven Adaptive Densification   w/o Anisotropy-Aware Gaussian Primitive Shape Regularization   w/o Dual Regularization via Normal-Depth Consistency   Baseline(PGSR)
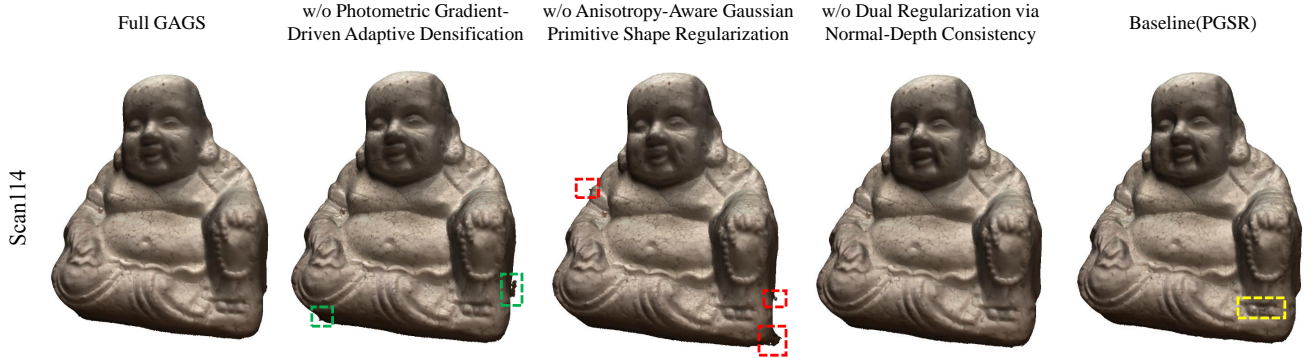

Figure 3. Qualitative comparison of meshes on the DTU dataset.

components to isolate their effects. The baseline PGSR (Chen et al., 2024) is provided as reference for comparison. As shown in Figure 3, qualitative comparisons using mesh visualizations indicate that the surface reconstruction results achieved by GAGS are visually superior. It can also be observed that the *Photometric Gradient-Driven Adaptive Densification* and *Anisotropy-Aware Gaussian Primitive Shape Regularization* modules play a crucial role in shape regularization, significantly contributing to the overall quality of the reconstructed surfaces.

Quantitative evaluation employs two complementary metrics: F1-score (higher better) measuring reconstruction completeness, and Chamfer Distance (CD, lower better) assessing geometric accuracy. As detailed in Table 2, our analysis reveals three key insights:

1. The *Photometric Gradient-Driven Adaptive Densification* module is critical for CD optimization. Removing this module increases CD by 8.93% (0.112→0.122) compared to full GAGS

2. The *Anisotropy-Aware Gaussian Primitive Shape Regularization* module primarily enhances surface quality. Its removal reduces F1-score by 1.59% (0.63→0.62) relative to full GAGS, but still maintains 17.0% improvement over PGSR

3. The *Dual Regularization via Normal-Depth Consistency* module plays a critical role in effectively integrating normal and depth information. Disabling this module causes the largest F1-score drop (11.11%, 0.63→0.56) and CD increase (5.36%, 0.112→0.118) versus full GAGS

5.5 Limitations

While achieving leading performance in geometric reconstruction, the GAGS framework exhibits three critical limitations requiring resolution. Firstly, the system demonstrates compromised effectiveness when reconstructing highly reflective surfaces, as exemplified by metallic materials in scan110, revealing inherent challenges in specular surface reconstruction. Secondly, the substantial 83% reduction in Gaussian primitives (averaging from 294k to 48k compared to PGSR) establishes an essential trade-off between noise suppression and detail preservation. Although effectively eliminating outliers (evidenced by 6.6% CD improvement over PGSR) while enhancing reconstruction precision, this sparsity inevitably induces loss of fine

Table 2. Ablation Study of GAGS Components (DTU Scan114)

Configuration	F1 Score ↑	Chamfer Distance ↓
Baseline (PGSR)	0.53	0.124
w/o Photometric Gradient-Driven Adaptive Densification	0.60	0.122
w/o Anisotropy-Aware Gaussian Primitive Shape Regularization	0.62	0.117
w/o Dual Regularization via Normal-Depth Consistency	0.56	0.118
Full GAGS	0.63	0.112

Bold Red: Best performance in category (F1: higher better, Chamfer: lower better)
↑: Higher is better, ↓: Lower is better

geometric details. Thirdly, parameter sensitivity persists in critical modules - the adaptive densification and shape regularization components rely on empirically tuned hyperparameters (e.g., pruning thresholds), necessitating scene-specific adjustments for optimal performance across diverse environments.

6. Conclusion and Future Work

Conclusion. This paper presents GAGS, a novel gradient-guided adaptive Gaussian splatting framework that addresses critical limitations in 3DGS-based surface reconstruction. By introducing three key innovations—photometric gradient-driven adaptive densification, anisotropy-aware shape regularization, and dual normal-depth consistency regularization—GAGS achieves significant improvements in both efficiency and geometric fidelity. Experimental results on the DTU dataset demonstrate that GAGS reduces the number of Gaussian primitives by an average of 83% while maintaining comparable visual quality, significantly enhancing surface reconstruction quality and accuracy, and showing a clear advantage over PGSR.The gradient-guided densification mechanism effectively suppresses redundant primitives in low-texture regions, while anisotropic shape constraints and dual regularization synergistically eliminate boundary artifacts and enforce surface coherence. These advancements make GAGS a practical solution for high-fidelity, memory-efficient surface reconstruction in real-world applications.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-1/W2-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China

Future Work. Although GAGS has shown promising results in geometric reconstruction, several challenges remain. Future work will focus on improving performance in large-scale, detailed scenes without compromising efficiency. Specifically, we aim to enhance reconstruction quality on reflective surfaces and better balance rendering quality with reconstruction accuracy. Additionally, we plan to develop a robust, scene-adaptive parameter optimization method to reduce reliance on manually tuned hyperparameters. Tackling these issues will help expand GAGS's applicability to a wider range of environments.

Acknowledgment

References

Besl, P. J., McKay, N. D., 1992. A Method for Registration of 3-D Shapes. *TPAMI*, 14(2), 239–256.

Chen, D., Li, H., Ye, W., Wang, Y., Xie, W., Zhai, S., Wang, N., Liu, H., Bao, H., Zhang, G., 2024. PGSR: Planar-based Gaussian Splatting for Efficient and High-Fidelity Surface Reconstruction. *arXiv preprint arXiv:2406.06521*.

Choy, C. B., Xu, D., Gwak, J., Chen, K., Savarese, S., 2016. 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction. *ECCV*, 628–644.

Curless, B., Levoy, M., 1996. A volumetric method for building complex models from range images. *SIGGRAPH*, 303–312.

Deng, B., Genova, K., Yazdani, S., Bouaziz, S., Hinton, G., Tagliasacchi, A., 2020. Nasa: Neural articulated shape approximation. *ECCV*, 612–628.

Fan, L., Yang, Y., Li, M., Li, H., Zhang, Z., 2024a. Trim 3d gaussian splatting for accurate geometry representation. *arXiv preprint arXiv:2406.07499*.

Fan, Z., Wang, K., Wen, K., Zhu, Z., Xu, D., Wang, Z. et al., 2024b. Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200+ fps. *Advances in neural information processing systems*, 37, 140138–140158.

Fang, G., Wang, B., 2024. Mini-splatting: Representing scenes with a constrained number of gaussians. *European Conference on Computer Vision*, Springer, 165–181.

Fischler, M. A., Bolles, R. C., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6), 381–395.

Guédon, A., Lepetit, V., 2024. Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5354–5363.

Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S., 2024. 2d gaussian splatting for geometrically accurate radiance fields. *ACM SIGGRAPH 2024 conference papers*, 1–11.

Jensen, R., Dahl, A., Vogiatzis, G., Tola, E., Aanæs, H., 2014a. Large scale multi-view stereopsis evaluation. *CVPR*, 406–413.

Jensen, R., Dahl, A., Vogiatzis, G., Tola, E., Aanæs, H., 2014b. A new benchmark for stereo with varying illumination. *International Journal of Computer Vision*, 106(2), 92–103.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4), 139–1.

Lee, J. C., Rho, D., Sun, X., Ko, J. H., Park, E., 2024. Compact 3d gaussian representation for radiance field. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 21719–21728.

Liu, W., Guan, T., Zhu, B., Xu, L., Song, Z., Li, D., Wang, Y., Yang, W., 2025. Efficientgs: Streamlining gaussian splatting for large-scale high-resolution scene representation. *IEEE MultiMedia*.

Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A., 2019. Occupancy networks: Learning 3d reconstruction in function space. *CVPR*, 4460–4470.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., Ng, R., 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106.

Papantonakis, P., Kopanas, G., Kerbl, B., Lanvin, A., Drettakis, G., 2024. Reducing the memory footprint of 3d gaussian splatting. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 7(1), 1–17.

Park, J. J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S., 2019. Deepsdf: Learning continuous signed distance functions for shape representation. *CVPR*, 165–174.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *NeurIPS*, 5099–5108.

Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W., Jiang, Y.-G., 2018. Pixel2mesh: Generating 3d mesh models from single rgb images. *ECCV*, 52–67.

Wang, T., Hou, Y., Zhang, Z., Xu, Y., Zhan, Z., Wang, X., 2025a. GS-I^3: Gaussian Splatting for Surface Reconstruction from Illumination-Inconsistent Images. *arXiv preprint arXiv:2503.12335*.

Wang, T., Wang, X., Hou, Y., Xu, Y., Zhang, W., Zhan, Z., 2025b. PG-SAG: Parallel Gaussian Splatting for Fine-Grained Large-Scale Urban Buildings Reconstruction via Semantic-Aware Grouping. *PFG(2025)*. https://doi.org/10.1007/s41064-025-00343-0.

Wang, X., Yi, R., Ma, L., 2024. Adr-gaussian: Accelerating gaussian splatting with adaptive radius. *SIGGRAPH Asia 2024 Conference Papers*, 1–10.

Yariv, L., Kasten, Y., Moran, D., Galun, M., Atzmon, M., Ronen, B., Lipman, Y., 2021. Multiview neural surface reconstruction by disentangling geometry and appearance. *NeurIPS*, 2492–2502.

Yu, Z., Sattler, T., Geiger, A., 2024. Gaussian opacity fields: Efficient adaptive surface reconstruction in unbounded scenes. *ACM Transactions on Graphics (TOG)*, 43(6), 1–13.

Zhang, Z., Hu, W., Lao, Y., He, T., Zhao, H., 2024. Pixel-gs: Density control with pixel-aware gradient for 3d gaussian splatting. *European Conference on Computer Vision*, Springer, 326–342.